RESPONSE OUTCOMES OF LOW ACHIEVING STUDENTS AND STUDENTS WITH LEARNING DISABILITIES

Odell George Vining

University of Southern Mississippi

Follow this and additional works at: https://aquila.usm.edu/dissertations

Part of the Educational Psychology Commons, School Psychology Commons, and the Special Education and Teaching Commons

Recommended Citation

Vining, Odell George, "RESPONSE OUTCOMES OF LOW ACHIEVING STUDENTS AND STUDENTS WITH LEARNING DISABILITIES" (2007). Dissertations. 1328.
https://aquila.usm.edu/dissertations/1328

This Dissertation is brought to you for free and open access by The Aquila Digital Community. It has been accepted for inclusion in Dissertations by an authorized administrator of The Aquila Digital Community. For more information, please contact Joshua.Cromwell@usm.edu.
RESPONSE OUTCOMES OF LOW ACHIEVING STUDENTS
AND STUDENTS WITH LEARNING DISABILITIES

by

Odell George Vining

Abstract of a Dissertation
Submitted to the Graduate Studies Office
of The University of Southern Mississippi
in Partial Fulfillment of the Requirements
for the Degree of Doctor of Philosophy

December 2007
ABSTRACT
RESPONSE OUTCOMES OF LOW ACHIEVING STUDENTS AND STUDENTS WITH LEARNING DISABILITIES
by Odell Vining
December 2007

Response-To-Intervention (RTI) has emerged as a viable, research based alternative to the IQ-Achievement (IQ-A) model for identifying students with learning disabilities. This study compared the treatment effects of a reading intervention package on low achieving (LA) students to students with learning disabilities (LD) as identified by an IQ-A discrepancy. Six, 3rd-grade students were screened and placed into dyads consisting of one student classified as LD and one student categorized as LA. The findings of this study were inconclusive in evidencing either LA or LD students as consistently increasing reading fluency at a greater rate than members of their respective dyad when an empirically supported reading intervention was implemented. The research study is consistent with previous literature suggesting that no uniform and meaningful difference exists between students identified as LA or LD.
ACKNOWLEDGMENTS

The writer would like to thank the dissertation director, Dr. Joe Olmi, for his insight, guidance, and continued support through this research project. It was Dr. Olmi’s techniques, work ethic, and trust that have had the strongest impact on my professional development over the last five years. Thanks also go to former South Forrest Attendance Center principal, Jon Will, for his ongoing support and proactive stance in incorporating Response-To-Intervention and research into his school.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>ABSTRACT</td>
<td>ii</td>
</tr>
<tr>
<td>ACKNOWLEDGMENTS</td>
<td>iii</td>
</tr>
<tr>
<td>LIST OF TABLES</td>
<td>vi</td>
</tr>
<tr>
<td>LIST OF ILLUSTRATIONS</td>
<td>vii</td>
</tr>
<tr>
<td>LIST OF ABBREVIATIONS</td>
<td>viii</td>
</tr>
<tr>
<td><strong>CHAPTER</strong></td>
<td></td>
</tr>
<tr>
<td>I.  INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td>II. REVIEW OF RELATED LITERATURE</td>
<td>3</td>
</tr>
<tr>
<td>Learning Disabilities: Prevalence and History</td>
<td></td>
</tr>
<tr>
<td>Discrepancy Model</td>
<td></td>
</tr>
<tr>
<td>Diagnostic Validity</td>
<td></td>
</tr>
<tr>
<td>Response-To-Intervention (RTI)</td>
<td></td>
</tr>
<tr>
<td>Empirical Support for RTI</td>
<td></td>
</tr>
<tr>
<td>Additional Research</td>
<td></td>
</tr>
<tr>
<td>Purpose of the Present Investigation</td>
<td></td>
</tr>
<tr>
<td>Research Question</td>
<td></td>
</tr>
<tr>
<td>III. METHODOLOGY</td>
<td>18</td>
</tr>
<tr>
<td>Participants and Setting</td>
<td></td>
</tr>
<tr>
<td>Materials</td>
<td></td>
</tr>
<tr>
<td>Dependent Variable</td>
<td></td>
</tr>
<tr>
<td>Design</td>
<td></td>
</tr>
<tr>
<td>Procedure</td>
<td></td>
</tr>
<tr>
<td>Pre-intervention</td>
<td></td>
</tr>
<tr>
<td>Skill Versus Performance Assessment</td>
<td></td>
</tr>
<tr>
<td>Baseline</td>
<td></td>
</tr>
<tr>
<td>Experimental Condition</td>
<td></td>
</tr>
</tbody>
</table>

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
IV. RESULTS

Discussion
Limitations
Future Research
Conclusion

V. DISCUSSION

APPENDIXES

REFERENCES
LIST OF TABLES

Table

1. Student Classification, IQ/Achievement Scores, and Reading Level ..................... 29
2. Pre/Post Intervention CWPM Median ................................................................. 33
## LIST OF ILLUSTRATIONS

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>Trails to Criteria</td>
<td>32</td>
</tr>
<tr>
<td>2.</td>
<td>Dyad 1 – Oral Reading Fluency</td>
<td>34</td>
</tr>
<tr>
<td>3.</td>
<td>Dyad 2 - Oral Reading Fluency</td>
<td>35</td>
</tr>
<tr>
<td>4.</td>
<td>Dyad 3 - Oral Reading Fluency</td>
<td>36</td>
</tr>
</tbody>
</table>

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
LIST OF ABBREVIATIONS

CBM Curriculum-Based Measurement
CWPM Correct Words Per Minute
DD Dual Discrepancy
DIBELS Dynamic Indicator of Basic Early Literacy Skills
ES Effect Size
FDD Frequently Dually Discrepant
IDD Infrequently Dually Discrepant
ISA Inter-Scorer Agreement
IQ-A IQ-Achievement
KBIT-2 Kaufman Brief Intelligence Test - Second Edition
LA Low Achieving
LD Learning Disability
NDD Never Dually Discrepant
ORF Oral Reading Fluency
ORFP Oral Reading Fluency Probes
R-CBM Reading – Curriculum-Based Measurement
RLD Learning Disability in Reading
RTI Response-To-Intervention
RTI-D Response-To-Intervention Discrepancy
WISC-IV Wechsler Intelligence Scale for Children - 4th Edition
CHAPTER I

INTRODUCTION

Reading impairment affects society on both small and large scale levels. To the individual, an inability to comprehend text reduces the probability of long-term, moderate/high income employment. In the community, dysfluent reading may result in additional fund allocation to social programs and correctional facilities. Illiteracy does not cause poverty or criminality, but statistics show that a disproportionate number of poor readers are represented in low income and prison populations.

In 1994, a five-point categorical scale was used to examine literacy of individuals incarcerated in the United States. A skill assessment from Level 1 asked participants to read a story and underline the sentence explicitly answering a specific question. Level 2 required the reader to utilize low level inferences from the reading passage. A task on Level 3 involved organizing information from longer text passages. Level 4, in part, asked participants to read and contrast opposing opinions outlined in the text. Level 5 involved parsing out distracters and making interpretations from dense text. Of the individuals assessed, 68% produced scores low enough to place them at either Level 1 or Level 2 (Haigler, Harlow, O'Connor, & Campbell, 1994). In a report issued by the U.S. Department of Justice three years later it was noted that 39.4% of males and 41.5% of females in state prisons did not have high school diplomas (U.S. Department of Justice, 2000). In 2001, $38.2 billion dollars were spent on state prisons (U.S. Department of Justice, 2004), as opposed to $27.3 billion federal dollars on education during the 1999-2000 school year (U.S. Department of Education, 2002).
The high incidence of illiteracy supports the idea that academic failure affects, in some capacity, the larger society. By improving reading skills in poor readers the likelihood of academic, behavior, and social problems that include peer/adult rejection, dropping out of school, difficulty finding and/or maintaining meaningful employment, and criminal incarceration is reduced (National Institute for Literacy, 1998).
CHAPTER II
REVIEW OF RELATED LITERATURE

Learning Disabilities: Prevalence and History

In the 2003-2004 academic school year, an estimated 5.8% of school aged children were diagnosed with specific learning disabilities (U.S. Department of Education, 2006). Of the nearly 2.9 million special education students in the public school system, 50% are diagnosed with a learning disability (LD) (U.S. Department of Education, 2002), and almost 40% of those with learning disabilities are expected to drop out of school (American Psychiatric Association, 2000).

A brief review on the history of LD provides the premise of this study. On April 6, 1963, Samuel Kirk coined the phrase “learning disability,” which was used to identify children who would be eligible for additional services (Kirk, 1963). Later, in 1975, The Education of All Handicapped Children Act was passed (Jacob & Hartshorne, 2003). Confronted with a Congressional deadline, the Bureau of Education for Handicapped Children had not selected a LD model that could identify and serve only those children who would benefit most from additional services. The discrepancy model, supported by a 1975 article (Rutter & Yule, 1975) on the findings from the Isle of Wight studies was adopted despite heated debate regarding the model’s validity and appropriateness. During this study, IQ scores for children living on the Isle of Wight were measured and regressed onto each child’s respective reading score. Results showed a disproportionate number of scores at lower levels of the distribution; indicating a higher instance of poor reading ability than would be expected by chance alone, given the scores come from a normal distribution and the null hypothesis is true (Rutter & Yule, 1975).
The conclusion drawn was that the students with the IQ/Reading discrepancy differed from non-discrepancy children in that children with a discrepancy had a better prognosis in non-reading/spelling subjects, while children without a discrepancy were given the better prognosis in subjects weighing heavily on reading and spelling. This contributed to the belief that children with significantly lower achievement to IQ scores suffer from some form of neuro-developmental deficit (Rutter & Yule, 1975).

Discrepancy Model

Relying heavily on the Isle of Wight project, despite the absence of replicable findings, schools have for 30 years classified students as learning disabled using the discrepancy model (Kavale, 2001). Based on the premise that IQ is static and can gauge one’s capacity for learning (Share, McGee, & Silva, 1989), an IQ score often acts as a “gatekeeper” for identifying low achieving students with average to high IQ, and parsing out students whose IQ is not high enough to allow a 15-point discrepancy (Kavale & Forness, 2000). The validity of the Isle of Wight study remains an area of contention some 30 years later (Fletcher et al., 1998).

It should be noted that criteria for diagnosing LD varies by state and by the magnitude of difference between either IQ-A or Achievement-Achievement scores. The common 15-point IQ-A discrepancy is the criteria selected for the current study. When using the IQ-A discrepancy model, children are not simply administered standardized tests, identified as having a 15-point discrepancy, and placed in special education. Instead, evidence must first demonstrate a failure by the child to respond positively to systemic implementation of pre-referral interventions. Only after one or more interventions have proven to be ineffective with the child will formal assessment through
standardized testing be considered (MacMillan & Siperstein, 2001). Advocates of the IQ-A discrepancy model point to research that shows .3 to .7 correlations between IQ and reading scores; suggesting that a relationship exists between intelligence and reading ability (Carver, 1990; Stanovich, Cunningham, & Freeman, 1984). Critics of the IQ-A discrepancy model counter by pointing out that no mandate exists that requires pre-referral interventions to be empirically supported. It can then be argued that a failure to improve may be due to poor selection of prescribed intervention(s) and instruction rather than a disability within the child.

Diagnostic Validity

Failure to address inadequate instruction (Fletcher et al., 2001), absence of effective treatment interventions (Gresham & Witt, 1997), and single point assessment determinations (Reschly & Ysseldyke, 2002) are issues of concern regarding diagnostic validity of the IQ-A model. The issue with single point assessment determination is that a discrepancy between achievement and intelligence scores typically reaches the 15-point threshold only when task demands are increased in later elementary grades (e.g., 3rd-grade and above). Students in lower elementary grades are assessed on a more narrow range of behaviors. Furthermore, cognitive tests frequently contain a poor floor which places limits on the minimum score possible. These factors decrease the likelihood that a discrepancy threshold will be met. Jenkins and O’Connor (2001) suggest that waiting for an IQ-A discrepancy forces children to fail before academic services are provided.

Given the prominent role of estimated intellectual functioning, IQ has been subject to review and later criticism from various sources. Willson and Reynolds (1985) noted intra-subject inconsistency in the quantity of the discrepancy between IQ-A scores
over time. As new IQ tests such as the Wechsler Intelligence Scale for Children – 4th Edition (WISC-IV) are introduced, smaller IQ-A discrepancies are noted. This is explained by a WISC-IV mean IQ that is 2.5 points lower than the mean IQ on the Wechsler Intelligence Scale for Children – 3rd Edition (Wechsler, 2003). This is not to suggest that society is becoming less intelligent. On the contrary, the Flynn effect is noted when scores on cognitive tests increase when children are given previous versions of tests. However, due to variation in test construction, test items may be tapping into different components of intelligence resulting in the variation of reported scores. For students evaluated using the discrepancy model, smaller differences may be obtained between IQ and achievement scores, resulting in many students ruled ineligible for special education services when new IQ tests are introduced to the market.

Dehn (2004) furthered this line of thought by noting that the WISC-IV might suppress IQ scores in students with LD due to its inclusion of a subtest, working memory, which is noted as a weakness in this population. Avis (2002) compared memory performance between students diagnosed with Attention-Deficit/Hyperactivity Disorder, students diagnosed with LD in reading (RLD), and a control group. Assessment tools for measuring attention included a memory scale and attention-concentration index. Results indicated that students with RLD experienced greater difficulty with working memory than the ADHD and control groups. Harcourt Assessment, Inc., previously known as
Psychological Corporation, now provides an alternative to generating the General Ability Index without using the Working Memory Index and Processing Speed Index.

Due, in part, to dynamic intra-student discrepancies over time, deflated IQ-A discrepancy with the introduction of new IQ tests, the potential suppression of IQ scores for students with learning disabilities, and/or the similarities between low achieving and students with LD on pre-reading and reading measures (Fuchs, Fuchs, Matthes, Lipsey, & Roberts, 2002; Steubing et al., 2002), many educational professionals do not support an IQ-A model for identifying learning disabilities. This may be best evidenced by a 2002 survey of 218 editorial board members of professional journals in the learning disabilities field with 70% reporting that an IQ-A discrepancy should play no role in the identification of children with LD (Speece & Shekitka, 2002). Wade Horn, a member of the President’s Commission on Excellence in Special Education (here to referred to as the President’s Commission) may have foreshadowed current/future events in stating “…drive a stake through the heart of this over reliance on the discrepancy model for determining the kinds of children that need services” (President’s Commission, 2002).

The United States Department of Education has reviewed LD research to assess the viability of the current identification process. Based largely on findings that IQ is not appropriate in identifying LD (Bradley, Danielson, & Hallahan, 2002; Donovan & Cross, 2002; Lyon et al., 2001; President’s Commission, 2002; Stuebing et al., 2002), The Individuals With Disabilities Education Improvement Act (P.L. 108-146) was passed on May 13, 2004. The bill allows educational agencies to make LD decisions based on student response to evidenced-based interventions rather than IQ-A discrepancies.
Response-To-Intervention (RTI)

The National Institute of Child Health and Development was responsible for much of the research conducted on LD in the 1990s. After a decade of research the organization concluded that the IQ-A criteria for LD lacked empirical support. In addition, they recommend broadening service delivery models by providing support to all students struggling academically (Fuchs et al., 2002). The RTI model, gaining support from an increasing number of school-based practitioners and researchers, allows teachers and school psychologists using evidenced-based interventions to identify, isolate, and monitor specific skill or sub-skill deficits early in the child’s school experience. The value of early identification is noted in research that shows a .88 probability that poor readers in 1st-grade remain poor readers in 4th-grade when students fail to receive appropriate interventions (Juel, 1988).

Gresham (1991) defined RTI as the change in academic performance resulting from an intervention. The RTI model has deep roots in learning theory and for over twenty years, RTI and curriculum-based measurement (CBM) have been used to identify, progress monitor, and reduce deficits of at-risk students by screening, providing interventions, and frequently monitoring student performance (Deno & Gross, 1973; Fuchs, 2004). Support for RTI has increased and the President’s Commission now recommends RTI as a primary component in the diagnostic criteria for identifying LD (U.S. Department of Education, Office of Special Education and Rehabilitation Services, 2002).

Variations of the RTI model exist, but prevention, early identification, a comprehensive evaluation that includes assessment in appropriate areas, implementation
of evidenced-based interventions, and progress monitoring are common elements across models. The process focuses on observable behaviors exhibited in excess or as a deficit when compared to same age/grade peers, and feedback to teachers regarding location, severity, and recommended treatment for skill deficits.

As a preventative measure, the RTI model is a conceptual framework that can take various forms. One form includes the identification of elementary students at-risk of reading failure using literacy tools such as the Dynamic Indicators of Basic Early Literacy Skills (DIBELS) (Good & Kaminski, 2002) and CBM. The DIBELS is a commercially available variation of CBM and most often used for benchmarking. Traditional CBM procedures allow for progress monitoring of level and slope of reading performance in the classroom. Benchmarking provides evidence on the performance of a student, class, school, and/or district at designated times (e.g., beginning, middle, and end of academic year). However, benchmarking is most often used to rank order student performance to allow for identification of the bottom 15-20% of students considered “at-risk.” Progress monitoring occurs between benchmarks and provides ongoing information on the level at which the student is currently performing, and the rate at which a student or group of students are learning, the data are valuable for screening, identification, monitoring progress, and evaluating effectiveness of instruction/intervention with a single or multiple students.

Criteria for meeting the RTI’s discrepancy (RTI-D) is the presence of a continued discrepancy in growth rate between peers, despite implementation of evidenced-based interventions (Gresham et al., 2004). Children showing a RTI-D are provided evidence-based intervention(s), progress monitored, and their scores compared to the mean peer
performance in a control group. When the RTI-D is resistant to interventions, the child is considered eligible for an LD diagnosis, and more intensive services are offered through special education (Compton, 2003). Benefits of the RTI model include reduction of false-positives (incorrectly identifying) and false-negatives (failing to identify), objective data-based decision making, comparison of both level and growth of reading performance, minimized need to administer intelligence tests, and the opportunity to improve education for all children (Speece, Case, & Molloy, 2003).

Used properly, RTI allows educators to monitor growth rates, implement classroom interventions, and make referrals based on treatment outcomes. After identifying students who are at-risk for academic failure, the focus shifts to alleviating the inter-student performance gap (Barnett, Daily, & Jones, 2004). The alternative approach to intervening with these students is not intervening and/or not referring children for evaluation until they have already experienced failure. The inherent flaw with the latter option is that when a small deficit(s) is/are ignored, the situation is exacerbated and the discrepancy(ies) increase(s) while the child experiences academic frustration and failure.

The RTI-D provides a more relevant assessment and intervention package than the 15-point difference identified by the IQ-A discrepancy model. The fundamental question is whether IQ-A discrepant children respond differently from non-discrepant, low achievers when empirically-validated interventions are implemented (Fletcher et al., 2001).

*Empirical Support for RTI*

*Meta-analyses.* In an attempt to differentiate cognitive functioning of LA children versus children with LD in the areas of phonological processing and reading, Hoskyn and
Swanson (2000) conducted a meta-analysis that included 19 studies. Phonological processing, vocabulary, reading, and syntax were dependent variables. Differences in treatment gains or effect sizes (ES) were noted in vocabulary (ES = .55) and syntax (ES = .87). However, ES for phonological processing and reading varied from -.02 to .29. The scores from the two groups fell at or approached the same level, indicating no meaningful differences in intervention effects were consistently found between groups. The conclusion was that the two groups were more alike than different and that cognitive functioning, as used for identification purposes in LD assessments, was not adequate in differentiating students with LD from LA children (Hoskyn & Swanson, 2000).

Researchers conducting a 2002 meta-analysis examined the validity of the IQ-A model by determining whether ES was consistent across areas more and less related to core reading skills, and by comparing the ES for treatment of students with IQ-A discrepancies to treatment size on non-discrepant, LA students. In reviewing the response-to-treatment between IQ-A and LA children, results were not significant (i.e., ES of .135). ES less than .20 suggests notable overlap in results of scores obtained by students classified as LA and LD. Examination of performance revealed skills most closely related to successful reading, generally produced small ES (i.e. verbal short-term memory (.10), phonological awareness (-.13) and vocabulary/lexical skills (.10)) for the group of students classified as LD (Stuebing et al., 2002). The ES for rapid naming was large (1.12), but phonological awareness, a better predictor of reading success, suggested higher reading outcomes for LA children. The authors concluded that placing a child in special education based on an invalid procedure (e.g., discrepancy model) would arbitrarily identify students eligible for services (Stuebing et al., 2002).
Therrien (2004) furthered Stuebing et al.’s (2002) research by examining whether a difference exists between the performance of students classified as LD and non-disabled students when administered a repeated reading intervention. Dependent variables included fluency and comprehension on both a repeated reading passage and on a generalized passage. Administration of repeated reading procedure involved having a student read and re-read an instructional level passage multiple times. When the reader fluently read the passage to a specified level of competency, a new passage was introduced. The repeated reading procedure has empirical support as being effective in increasing reading fluency (National Reading Panel, 2000).

The Therrien (2004) study reported repeated reading passages with non-disabled students resulted in a large ES (.85) in fluency with a moderate ES (.64) in comprehension; on novel passages, a moderate ES (.59) in fluency and low ES (.18) in comprehension. Students with LD showed moderate ES (.75, .73, .79, & .41) in all areas measured. The study suggested that empirically-supported interventions (i.e. repeated reading) are effective in remediating deficits in IQ-A discrepant students and improves skills of non-disabled peers (Therrien, 2004). With ES approaching .80, the research contributes to a growing body of literature that points to large intervention effects as evidenced by students with learning disabilities showing notable gains in reading performance (Alexander, Anderson, Heilman, Voeller, & Torgesen, 1991; Truch, 1994; Wise, Ring, & Olson, 1999).

*Longitudinal Studies.* Findings from longitudinal studies provided additional support for the conclusion drawn from meta-analyses suggesting that students, regardless of the presence or absence of a 15-point IQ-A discrepancy, possess the capability of
showing meaningful improvements in reading. To determine if IQ, even when in excess of a standard deviation, makes a significant difference in differentiating the ability of an individual to learn to read or respond to an intervention, the Connecticut Longitudinal Project in 1983 followed 445 children from 1st-graders through their ninth grade year. Twenty one students, identified as reading disabled, or frequently below peers in reading ability, were compared to thirty-five average readers and thirty-nine students classified as reading in the superior range. During the students' ninth grade year, each was given an assessment in academic, cognitive, and language skills. Growth rates of LA students and students with LD were compared. Despite a mean IQ difference of 18 points between groups, no significant differences were found in any grade. These findings suggest that when students are provided effective instruction, no meaningful difference exists when comparing growth rates between children diagnosed as LD and LA children. This is an important consideration when determining a criteria by which state and local education agencies assess learning disabilities (Francis, Shaywitz, Stuebing, Shaywitz, & Fletcher, 1996).

A second longitudinal study examined the reading development of kindergarten and first grade students. The purpose of the study was to determine whether different interventions were warranted dependent upon the LD or LA classification of the student. Kindergarten students were assessed on letter sound association, perceptual discrimination, phonological processing, rapid naming of letters and objects, and visual-motor integration. First graders were measured on perceptual discrimination, phonological processing, rapid naming of letters and objects, spelling, visual-motor integration, and word reading (O'Malley, Francis, Foorman, Fletcher, & Swank, 2002).
When comparing LA students to students with LD, no significant difference was reported in letter sound association, perceptual discrimination, phonemic awareness, rapid object naming, spelling, visual motor integration, and word reading. Findings suggest that there was no meaningful difference between the two reading impaired groups (O’Malley et al., 2002).

In a third longitudinal study, dual discrepancy (DD) was used to test the diagnostic validity of RTI. Requirements for DD were that both slope and level of a student’s performance be a minimum of one standard deviation below the class average. Over a three-year period, a DD group was compared to a control group. Children were placed in one of three groups: never dually discrepant (NDD), infrequently dually discrepant (IDD), or frequently dually discrepant (FDD), which required meeting the dual discrepant criteria more than three times in the nine to ten opportunities over a three-year period that they were assessed. FDD had less developed skills than both IDD and NDD in Letter Word Identification with a large ES (> .80). Beginning in year 2, the NDD and IDD groups each performed significantly better ES (> 1) than the FDD group. The results of the study suggest that students meeting the FDD criteria had fewer reading skills, exhibited fewer acceptable behaviors, and used more school resources. These findings support claims that RTI is socially valid (Case, Speece, & Molloy, 2003). These claims of validity are further supported in the works of Siegel, Jastak, and Wilkinson.

Additional Research. Research (Siegel, 1992) has examined whether, in reading, meaningful differences exist between non-disabled, LA students and students identified as LD; as defined by obtaining a score in the bottom 25th percentile on the Wide Range Achievement Test (Jastak & Wilkinson, 1984). A battery of standardized tests measured
student performance in reading. Phonological tasks included reading of nonsense words, spelling of nonsense words, recognition of visual form of sounds, and visual-phonological reading tasks. Additional tests measured reading, spelling, and arithmetic. No meaningful differences were reported between the groups in areas required for successful reading (Siegel, 1992).

Stanovich and Siegel (1994) continued Siegel's line of research by assessing whether characteristics of reading sub-skills differ between children classified as LD and LA. Phonological, orthographic, memory, and language processing skills were measured and compared between groups. Students were administered spelling and symbols, pseudo word reading subtests from the Goldman-Fristo-Woodcock Sound-Symbol Test (Goldman, Fristoe, and Woodcock, 1974), word attack subtest from the Woodcock Reading Mastery Test-Revised (Woodcock, 1987), and sets of phonetically decodable and non-decodable words. No meaningful differences were identified in the processing techniques that allow for recognition of words (Stanovich & Siegel, 1994).

Speece and Case (2001) sought to identify the type of student exhibiting the greatest reading deficit. They compared students with a DD of level/growth rate to students classified as LD, LA, and children with no reading problems (purposive sample). In addition, examiners sought to assess whether current procedures, CBM and phonological awareness tests, were good predictors for identifying DD students.

The following ES were noted from group comparisons: DD group to LD group on Total Phonological Awareness ES of -.33 in Fall, and ES of -1.39 in Spring (grade 2) with age ES of 1.16. DD to LA on Total Phonological Awareness for grade 2 ES of -.64 in fall, and ES of -.84 in spring with Age ES of -.69. CBM oral reading fluency (ORF)
was superior in identifying students with the greatest reading deficit (Speece & Case, 2001).

In summary, a growing body of literature shows: (a) results of IQ-A discrepancy alone are not valid measures for differentiating students who can or cannot benefit from appropriate, evidenced-based interventions, and (b) there is no significant difference in the manner that discrepant/non-discrepant children respond to interventions (Stanovich & Siegel, 1994; Fletcher et al., 1994; Foorman, Francis, & Fletcher, 1995; Francis et al., 1996). Current legislation now provides greater latitude to local education agencies in determining how to identify students with LD. As stipulated in IDEIA, 2004, IQ-A discrepancy may be used in the LD identification process, but is no longer required. A child’s ability or inability to respond to evidenced-based instruction or intervention may now weigh heavily in determining special education eligibility.

If the RTI model fails to identify, prevent, and successfully remediate at-risk students, the long favored IQ-A model is likely to return to prominence. School psychologists and other assessment specialists would again be little more than test administrators and “gatekeepers” to special education. However, if RTI proves successful, national education is likely to follow states such as Iowa in the screening, benchmarking, intervening, and progress monitoring of students. Administration of IQ tests may then be best suited for identifying students as gifted or cognitively impaired.

Purpose of the Present Investigation

The purpose of this study is to compare the reading growth rate, or the increase in correct words per minute over time, of ORF between three dyads of LD and LA students. The primary component of the intervention package selected, repeated reading, has been
recommended by the National Reading Panel (2000) as an effective treatment for improving oral reading fluency. Secondary components, phase drill error correction and performance feedback, also have empirical backing (Eckert, Ardoin, Daly, & Martens, 2002; Rosenberg, 1986). Progress monitoring and analysis of performance within the dyads will contribute to a body of literature examining the reading growth rates or response to interventions of different subgroups within the school population.

Research Question

Do response outcomes (i.e., reading level and growth rate) of low achieving 3rd-graders differ from response outcomes of learning disabled 3rd-graders after implementation of an evidence-based treatment package?
CHAPTER III

METHOD

Participants and Setting

Participants included six, 3rd-grade students (three general education, three classified as learning disabled in reading) referred by the school's teacher support team (TST). Three participants were female and three participants were male. Ages ranged from nine years three months to ten years four months. Criteria for admission in the LA group were: (a) falling in the bottom 20th percentile of students benchmarked on the DIBELS ORF subtest during previous spring; (b) identified as pure skill or skill and performance deficit, but at or above a 2nd-grade instructional level on oral reading fluency; (c) <15-point discrepancy between obtained scores on the Kaufman Brief Intelligence Test, Second Edition (KBIT-2) (Kaufman & Kaufman, 2004) and the Woodcock-Johnson III Tests of Achievement (WJ-III-A) Broad Reading Score (Woodcock, McGrew, & Mather, 2001). A criterion for admission into the IQ-A discrepancy group required a diagnosis of learning disability in basic reading as outlined in eligibility guidelines of the Mississippi Department of Education (2003). This discrepancy was often interpreted as a 15-point difference between performance on norm referenced, standardized tests of intelligence as compared to performance on a standardized measure of basic reading. The second criteria included confirmation of a 15-point IQ-A discrepancy using the KBIT-2 and WJ-III-A.

Materials

Kaufman Brief Intelligence Test - Second Edition (KBIT-2). The KBIT-2 is a brief, intelligence test used to estimate an individual's cognitive ability. The K-BIT-2 is
typically used as a screening measure and was not used in this study for official diagnostic purposes, but instead to confirm that students already diagnosed through comprehensive evaluations maintained a 15 point IQ-A discrepancy. In addition, the test was one of the assessment tools used to estimate the IQ-A discrepancy in students not diagnosed with LD. The KBIT-2 has a mean of 100 and a standard deviation of 15, provides nonverbal (fluid) and crystallized (verbal) measures, and age-related percentile ranks. It is designed for individuals aged 4-90, and administration time is approximately 20 minutes. Internal-consistency reliability was calculated using a split-half procedure. Internal consistency for the Verbal score has a mean of .90 and a range from .86 to .96. The reliability for the Nonverbal score has a mean of .88 and a range from .78 to .93. The reliability for the Composite IQ has a mean of .93 and a range from .89 to .96 (Kaufman & Kaufman, 2004).

Test-retest for the Verbal score has a mean of .91 and a range from .88 to .93. Test-retest for the Nonverbal score has a mean of .83 and a range from .76 to .89. Test-retest for Composite IQ has a mean of .90 and a range from .88 to .92. The standard error of measurement for Verbal score ranges from 4 to 5, Nonverbal ranges from 4 to 7, and Composite IQ ranges from 3 to 5 points (Kaufman & Kaufman, 2004).

Examination of correlations between the KBIT-2 and other cognitive tests aide in evaluating the validity of the KBIT-2. The KBIT-2 was correlated with the KBIT, Wechsler Abbreviated Scale of Intelligence (WASI; The Psychological Corporation, 1999), Wechsler Intelligence Scale for Children- Third Edition (WISC-III; Wechsler, 1991), and the Wechsler Intelligence Scale for Children – Fourth Edition (WISC-IV; Wechsler, 2003). Correlations between the KBIT-2 and the KBIT ranges from .77 to .84.
on Composite IQ, .81 correlation with the WASI Full Scale IQ, .76 with the WISC-III
Full Scale IQ, and .77 with the WISC-IV Full Scale IQ. Correlations between the KBIT
and KBIT-2 ranged from .80 to .86 for IQ Composite score, .77 to .84 for the Verbal
portion, and an adjusted correlation value of .78 to .79 for the Nonverbal portions
(Kaufman & Kaufman, 2004).

*Woodcock-Johnson III Tests of Achievement (WJ-III-A).* The WJ-III-A has a mean
of 100 and a standard deviation of 15, with median reliability measure at .83 for
individuals five to nineteen years of age. The Passage Comprehension subtest has a
median reliability coefficient of .88, and a standard error of measurement of 5.12. The
Broad Reading Cluster from the WJ-III-A (McGrew & Woodcock, 2001) has a .93
median alternate-form reliability for individuals ages 5 to 19 and provides an estimate of
an individual’s reading achievement. The cluster includes subtests from Letter-Word
Identification, Reading Fluency, and Passage Comprehension. Collectively, these subtests
measure reading decoding, fluency, and comprehension skill.

*Reading-Curriculum Based Measurements (R-CBM).* As early as 1985, CBM was
identified as a measurement tool that allows progress monitoring and instructional
decision-making (Deno, 1985). Over the last twenty years, research has supported CBM
usage for tracking student progress, influencing instruction, and improving academic
performance (Christ & Vining, 2006; Fuchs & Fuchs, 1986; Fuchs, Fuchs, Hamlett, &
Allinder, 1991; Fuchs, Fuchs, Hamlett, & Walz, 1993; Fuchs, Fuchs, & Stecher, 1989).
Studies have also proven CBM to be a valid measure of reading performance (Fuchs,
Deno, & Mirkin, 1984; Good & Jefferson, 1998). CBA, the umbrella term used for CBM
has shown an ES of .70 in controlled studies (Fuchs & Fuchs, 1986). R-CBM were
correlated with norm-referenced tests of reading that included (MacGinite Reading Tests; MacGinitie, Kamons, Kowalski, MacGinitie, & McKay, 1978; Metropolitan Achievement Tests; Prescot, Balow, Hogan, & Farr, 1984), maze, and teacher prediction. The correlations between R-CBM and the norm-referenced tests ranged from .80 to .88 (Hintze, Shapiro, & Conte, 1997).

As research on CBM has evolved, literature has shown that 3rd grade controlled reading passages significantly reduce measurement error from 15.96 for uncontrolled reading passages to 13.99 for controlled reading passages; resulting in an increased sensitivity of observed scores and reliability of the measurement tool (Hintze & Christ, 2004). During this study, CBM was used to identify reading levels and monitor growth rates of participants’ ORF in response to the intervention strategy.

*Oral Reading Fluency Formulas.* A limitation with using CBM procedures is the failure to address and/or account for error attributable to probe. This study attempted to reduce measurement error by indexing probes using two readability formulas. Research has identified Spache (Spache, 1953) as relatively effective in determining grade level, but less efficient at determining difficulty within grade level. FORECAST (Sticht, 1973) is better suited for providing rank order of difficulty (Ardoin, Suldo, Witt, Aldrich, & McDonald, 2005). Determining grade level with Spache and rank ordering within grade with FORECAST may allow for variation in scores to be more heavily influenced by student performance and less the result of variation of passage difficulty.

Stated differently, the Spache provides a good estimate of the grade level for a reading probe. However, the Spache does not perform as well when asked to identify, within-grade probes that are of less, equal, or greater difficulty. Knowing within-grade
difficulty is important because variability of scores can be reduced when probes approximate the same level of difficulty. The FORECAST is more effective than the Spache at rank ordering. By first determining grade level with the Spache and then using the FORECAST for within-grade rank ordering, the researcher sought to more effectively identify grade level and within-grade difficulty of the probe set used in this study.

Probes were administered in order of increasing difficulty to reduce false positives, or Type I errors. High Type I error rates delay or prohibit students from receiving the more intensive academic services they need to increase the likelihood of academic success. Using the conservative approach described in this paper for evaluation and/or progress monitoring should result in fewer reports of students making acceptable progress when such progress is not actually occurring.

*Oral Reading Fluency Probes (ORFP).* Research from Fuchs and Deno (1994) provides evidence that material from the student’s school curriculum is unnecessary in obtaining relevant data from ORFP. The text probes used in the present study were 150-200 word passages, constructed by students from The University of Southern Mississippi Ph.D. school psychology program. English teachers and journalism professors edited probes, prior to their use, for syntax, flow, and content for the respective grade levels. The use of self-constructed, rather than commercial probes was selected for three reasons. First, with limited monetary resources available, teachers may be unable to purchase probes. Although there are non-commercial probes available for use (e.g., DIBELS Oral Reading Fluency Probes) many teachers are unaware of their availability and/or have difficulty locating such material. Second, though companies selling probes commercially boast of low standard error, seldom do independent parties replicate results. Third,
readability formulas used to assess intra and inter-grade level probe difficulty are imperfect and inconsistent instruments. To enhance the ability to accurately define grade level and passage difficulty; each passage passed through a series of readability formulas. By using the strengths of specific formulas, attempts were made to produce estimates approaching true readability levels.

Dependent Variables

ORF is a measure of accuracy and fluency of connected text. It was selected as the dependent variable because small gains in student growth are detectable over relatively short periods of time, it is a valid measure of basic reading skills as evidenced by correlations as high as .91 with comprehension, and because fluent reading is important for generalization to comprehension (Eckert, Ardoin, Daisey, & Scarola, 2000; Fuchs, Fuchs, & Maxwell, 1988; Shapiro, 1996). ORF has been defined as the number of words correctly read aloud in one minute (Shinn, 1998), or the ability to verbalize an assortment of words with speed, accuracy, and proper expression (National Reading Panel, 2000). The average increase in oral reading fluency for 3rd-graders is 1.18 word gain/week (Deno, Fuchs, Marston, & Shinn, 2001).

Design

A multiple baseline across participants design with replications across dyads was used to evaluate reading gains. The staggering of Dyad 2 was in this study. The treatment phase lasted five weeks with progress monitoring occurring on two non-consecutive days each week. This design allowed for the recommended collection of two data points per week, which met the suggested minimum collection of 10 data points for R-CBM analysis (Shinn, 2002). The multiple baseline across participants design was selected.
because the design allows intervention results to be extrapolated from multiple pairings containing participants categorized as LD and LA.

**Procedure**

*Consent/Assent.* Informed consent was obtained from the students’ parents and teachers (Appendices A and B), and assent was obtained from the students (Appendix C). Reading levels, baseline, and intervention administration took place in an academic intervention room.

*Pre-intervention.* The instructional reading level for each student was determined by comparing the median correct words per minute (CWPM) to the recommended instructional reading level (Howell, 2005) on three, 3rd-grade probes. Stated differently, the level at which a child was considered instructional was determined by taking the middle of rank ordered scores from the number of words that the child read correctly during timed, one-minute intervals. Students scoring below the cutoff were administered three, 2nd-grade probes using the above procedure. If any student had presented with scores indicating frustration at Howell’s 2nd-grade (40-60 CWPM) and 3rd-grade (70-100 CWPM) range (Howell, 2005), the student would have been replaced by a student meeting study participation criteria. The student/students not meeting study criteria due to frustrational reading levels were to be referred to the school’s TST with data collected during the screening process and recommendations for intervention. However, no student failed to meet the recommended instructional level on 2nd-grade or 3rd-grade passages. Next, the KBIT-2 and WJ-III-A were administered to each child to determine intelligence and achievement estimates. Within one week of identifying the instructional reading level for each child the KBIT-2 and WJ-III-A were administered by the researcher in an
academic intervention room. All students were tested to provide support that they met/failed the meet the state requirement for LD. Students were then randomly paired into dyads containing one student identified as meeting the LD criteria and one student categorized as LA.

*Skills Versus Performance Assessment.* Students identified as frustrational on 3rd-grade reading probes entered a skills versus performance assessment phase to determine if low scores were attributable to performance or skill deficits. In this phase, students were presented four different 3rd-grade probes. Scoring followed standard CBM procedures.

*Baseline.* The median CWPM from three instructional level reading probes provided one datum point per day/per child during the baseline phase. A standardized CBM procedure was used to assess reading fluency. The administrator laid the reading passage on the table/desk in front of the student and read aloud a set of standardized directions before each student was instructed to read aloud from ORFP. The instructions were as follows:

When I say “begin,” start reading aloud at the top of the page. Read across the page (DEMONSTRATE). Try to read each word. If you come to a word you don’t know, I’ll tell it to you. Be sure to do your best reading. (National Association of School Psychologists, 2005).

The student was instructed to “begin” and the one-minute timing started when the student read the first word. If a student took more than three seconds on a word, the word was provided to the student. A slash was placed across omissions and words verbalized incorrectly. After 60-seconds, the student was instructed to stop, and a bracket was placed
after the last word read. Total words read, errors, and CWPM were recorded. The median CWPM from that day was selected and plotted by the examiner. CBM procedures are included in Appendix D. A minimum of three data points in baseline were collected to improve the likelihood of obtaining a more accurate level of student performance on oral reading fluency. Determining level with fewer measures subjects the study to risk associated with the possibility of increasing trends and/or outliers.

Phase drill error correction is a supplemental acquisition component that uses stimuli overcorrection on misstated or omitted words. The administrator highlighted or underlined words in texts that the reader did not correctly verbalize. Between passage readings, the student viewed his/her ORFP. Instructions were as follows:

The first word that you missed is [pointed to and verbalize first word missed]. Read the sentence [pointed to the first word in the sentence that included the missed word] that included the word [verbalized missed word].” The student read the sentence that included the missed word. [Examiner provided verbal praise (e.g., “good,” “nice job,” etc.) [Examiner pointed again to the sentence that included the missed word.] Read the sentence a second time. The student read the sentence a second time. Verbal praise was provided again. The examiner pointed to the sentence again. “Read the sentence a third time.” The student read the sentence a third time. Verbal praise was provided. If the student misread the word originally missed; he/she was corrected and the phase drill procedure was repeated. This process continued until sentences containing all missed words had been modeled by the examiner and correctly reread by the student (Daly, Murdoch, & Lillenstein, 2002).
The final component, performance feedback, acted as a motivational element that has been shown to enhance the effects of repeated reading (Eckert et al., 2002). Use of this procedure includes having the student record (e.g. color bar charts, draw line graph) CWPM on the first and last intervention reading probe administered each day. Instructions were as follows:

"Today, on your first reading; you read [stated CWPM for first passage]. I want you to color in the rectangle that shows your score on today’s first reading. [A vertical rectangle was drawn that represented the student’s score on the day’s first reading probe, and then given to the student to pick a colored marker/pen/color.] Now, color in the rectangle that shows your first reading score. [The child was allowed to color the rectangle, positive verbal praise was provided, and a second vertical rectangle was drawn beside the first rectangle.] This rectangle represents your last reading score. [The child was instructed to select a different color, and issued verbal praise]. Now, color in the rectangle that shows your second reading score."
CHAPTER IV
RESULTS

Students’ performance on the KBIT-2, WJ-III-A, score discrepancy between the KBIT-2 and the WJ-III-A, and instructional reading levels are listed in Table 1. Average scores for students with LD were 102 (IQ) and 82 (Achievement) and 95 (IQ) and 84 (Achievement) for students categorized as LA. When presented 3rd-grade reading passages, two students (Participants 5 & 6) performed below the instructional range (70-100 CWPM). Both students entered a skill versus performance assessment phase that identified each student’s low performance as a result of a skill deficit, as evidenced by failure to obtain a 20% increase in CWPM on paired passages that offered a tangible reinforcer. Participant 5 and Participant 6 were then administered 2nd-grade reading passages with obtained scores placing them at instructional level (40 to 60 CWPM) on 2nd-grade reading material.

Third grade probes 1 and 2 were paired, and 3rd-grade probes 3 and 4 were paired. Probes were administered in sequential order. Students were given the opportunity to select a reward if performance scores on probe 2 and probe 4 were each 20% higher than on the paired non-reinforced probe. If a student scored 20% higher (Duhon, Noell, Witt, Freeland, Dufrene, & Gilbertson, 2004) on both probe 2 and probe 4 as compared to his/her respective probe pairing, and one of the scores on probe 2 or probe 4 fell within the instructional range for 3rd-grade, the student would have been identified as exhibiting a performance deficit and removed from the study. The child would have then been referred to TST with data collected during the screening process and recommendations.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
for intervention. However, no student in this study was identified as exhibiting a performance deficit.

Table 1

*Student Classification, IQ/Achievement Scores, and Reading Level*

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>01 (LD)</td>
<td>91</td>
<td>75</td>
<td>16</td>
<td>75 (3rd)</td>
</tr>
<tr>
<td>02 (LA)</td>
<td>94</td>
<td>82</td>
<td>12</td>
<td>71 (3rd)</td>
</tr>
<tr>
<td>03 (LD)</td>
<td>116</td>
<td>86</td>
<td>20</td>
<td>76 (3rd)</td>
</tr>
<tr>
<td>04 (LA)</td>
<td>94</td>
<td>84</td>
<td>10</td>
<td>80 (3rd)</td>
</tr>
<tr>
<td>05 (LD)</td>
<td>99</td>
<td>84</td>
<td>15</td>
<td>68 (2nd)</td>
</tr>
<tr>
<td>06 (LA)</td>
<td>96</td>
<td>85</td>
<td>11</td>
<td>79 (2nd)</td>
</tr>
</tbody>
</table>
Students identified as LD included participants 1, 3, and 5. These participants obtained IQ scores of 91, 116, and 99 respectively. Their Achievement scores were 75, 86, and 84, respectively. Differences between IQ and Achievement scores for students diagnosed as LD were 16, 20, and 15. Two participants (i.e., Participants 1 and 3) diagnosed as LD exhibited an instructional reading level at third grade. Participant 5 demonstrated a second grade instructional reading level.

Students identified as LA included participants 2, 4, and 6. These participants obtained IQ scores of 94, 94, and 96 respectively. Their Achievement scores were 82, 84, and 85, respectively. Differences between IQ and Achievement scores for students diagnosed as LA were 12, 10, and 11. Two participants (i.e., Participants 2 and 4) diagnosed as LA exhibited an instructional reading level at third grade. Participant 6 demonstrated a second grade instructional reading level.

During the baseline phase, students were assessed twice per week, void of intervention, to determine current reading performance. A minimum of three data points were collected in each pairing’s baseline phase. Each dyad began baseline at the same time. In Dyads 1 and 3, students with LD entered the intervention phase when they showed stability within their respective baselines. Visual inspection, rather than a statistical criterion, was used to make determinations on stability. The LA students entered the intervention phase after stability was shown within the LA student’s baseline, and when any improvement was noted in the paired LD student’s interventions score beyond the median of baseline data points. In Dyad 2, the LA student entered the intervention phase when stability was noted in the baseline phase. The student with LD entered the intervention phase after stability was shown within the LD student’s baseline,
and when any improvement was noted in the paired LA student’s interventions score beyond the median of baseline data points.

*Intervention Condition.* Two researchers listened to participants read from ORFP developed by doctoral students in the manner described earlier. A word-by-word inter­scorer agreement (ISA) procedure was used by dividing the sum of agreements by the total number of words (Fiala & Sheridan, 2003). Calculated on 20% of CBM probe administrations, mean agreements were 98% with a range of 92% to 100%. A checklist of procedural components [Appendix E] was completed by the researcher on 50% of R-CBM administrations and used as integrity checks. Procedural integrity checks measured the extent that the procedure was performed as intended. One hundred percent compliance with prescribed procedural implementation was reported on the self-reporting instrument.

During the intervention phase, 3rd-grade students received 10 to 15 min of the intervention package (i.e., repeated reading [Appendix F], phase drill error correction, and performance feedback) three days per week and progress monitoring after intervention administration on two days per week. Administration of repeated reading involved having a student read and re-read an instructional level passage until he/she was able to read the passage at a rate of 100 CWPM. A novel, instructional level passage was then introduced and the child again read and re-read the passage until the 100 CWPM criterion was met. This process continued for the duration of the intervention session.

*Data Analysis.* The frequency of readings per intervention passage(s) varied, but ranged between two and six trials per student. The determination to use a 100 CWPM criterion for each student was based on the upper limit of Howell and Nolet’s (Howell,
2005) recommended instructional criteria for passage difficulty. The frequency of readings per intervention passage(s) varied, but ranged between two and six trials per student as noted on Figure 1.

![Figure 1. Oral reading fluency trials to criteria.](image)

Participants' performance on ORF during baseline and intervention phase is noted in Table 2. Figure 2 illustrates that in Dyad 1, Participant 1 (LD) demonstrated a 2.40 word increase per week, and Participant 2 (LA) exhibited a 2.75 word increase per week. Figure 3 shows that in Dyad 2, Participant 4 (LA) demonstrated a 2.2 word increase per week, and Participant 3 (LD) exhibited a 3.1 word increase per week. Figure 4 illustrates that in Dyad 3, Participant 5 (LD) demonstrated a 2.0 word increase per week, and Participant 6 (LA) exhibited a 2.67 word increase per week.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
Table 2

*Pre/Post Intervention CWPM Median*

<table>
<thead>
<tr>
<th>Participant/Classification</th>
<th><em>Baseline</em></th>
<th><em>Intervention</em></th>
<th>Rate of Growth Per Week</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dyad 1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>01 (LD)</td>
<td>58</td>
<td>70</td>
<td>2.40</td>
</tr>
<tr>
<td>02 (LA)</td>
<td>61</td>
<td>72</td>
<td>2.75</td>
</tr>
<tr>
<td>Dyad 2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>03 (LD)</td>
<td>76</td>
<td>90</td>
<td>3.10</td>
</tr>
<tr>
<td>04 (LA)</td>
<td>81</td>
<td>92</td>
<td>2.20</td>
</tr>
<tr>
<td>Dyad 3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>05 (LD)</td>
<td>75</td>
<td>86</td>
<td>2.00</td>
</tr>
<tr>
<td>06 (LA)</td>
<td>87</td>
<td>99</td>
<td>2.67</td>
</tr>
</tbody>
</table>

* - Median of last three data points per phase
Figure 2. Oral reading fluency during baseline and intervention for students in Dyad 1.
Figure 3. Oral reading fluency during baseline and intervention for students in Dyad 2.
Figure 4. Oral reading fluency during baseline and intervention for students in Dyad 3.
CHAPTER V
DISCUSSION

To identify, if any, the classification/category of students benefiting more from delivery of validated reading treatments, tightly controlled research must systematically group and compare students' performance with empirically-supported interventions. The purpose of this study was to address the question of whether an empirically-supported reading intervention package would result in students diagnosed with a LD experiencing enhanced, equivalent, or a decreased growth rates in ORF when compared with same grade LA peers receiving the same intervention.

The primary intervention component selected, repeated reading, has been recommended by the National Reading Panel (2000) as an effective treatment for improving reading fluency. Progress monitoring occurred twice weekly, after implementation of the reading intervention package, during the intervention period. To minimize variability and reduce standard error, probes were entered into a series of readability formulas specializing in identifying and rank ordering grade level passages.

Intervention data do not support the position that reading level and growth rate/slope of IQ-A discrepant 3rd-graders consistently differed from those of non-discrepant LA peers after implementation of an evidence-based treatment. These findings are based on examination of ORF levels and growth rates of randomly assigned participants in dyads when comparing the median from the last three data points in baseline to the median of the last three data points in intervention phase.

The results of initial analyses indicate that 5 of 6 students in the study exhibited an increase in ORF greater than would be expected by instruction in general education
alone. Growth rate for 3rd-grade students is reported at less than two word gains per week. The average increase in oral reading fluency for 3rd-graders is 1.18 words gained per week (Deno, Fuchs, Marston, & Shinn, 2001). In this study, within group comparisons did not reveal consistency of superior performance gains for students classified as LD or students categorized as LA. In two of the three dyads, students identified as LA increased ORF at a greater rate than students identified as LD. This is not to suggest that students classified as LA would benefit more from empirically-validated treatments, but instead lends credence to the body of evidence supporting no meaningful difference in the potential benefit of administering empirically-supported interventions to either students considered LD or LA. In an applied setting, this translates into additional support for incorporating a RTI model that includes universal screenings, benchmarking, and identification/intervention/progress monitoring of at-risk students. If students respond in similar fashion to proven interventions, it is not conducive to the child nor the learning environment to provide support for one child and not another.

Limitations

Limitations should be acknowledged while interpreting the findings of this study. First, classical test theory suggests that error exists with imperfect measurement instruments. R-CBM, though reportedly sensitive to short-term growth, has been shown to use probes with error rates as high as 15-16 for decisions when using uncontrolled passage(s). An analysis was not conducted on the probe sets/reading passages to determine standard error of measurement, and Generalization Theory using G and D tests did not evaluate the source of error or the number of probes that if administered would reduce error while minimizing administration time. Such psychometric information may
have proven useful in reporting a range reduction in which the true oral reading fluency scores could be stated with greater confidence. To partially offset the impact of this limitation, passages were entered into a series of readability formulas, with respective strengths in identifying grade level and determining difficulty within grade. An additional means of reducing error involved administering and recording the median of three probes rather than graphing a datum using a single probe.

A second limitation involved selection of participants, in part, based on the previous year’s Spring DIBELS ORF subtest data. A more desirable procedure would have been to use Fall or Winter data DIBELS ORF subtest data. The negative impact of this limitation may have been lessened as Spring data from the previous year is expected to have correlated to a high degree with the subsequent Fall data. However, no data was analyzed to substantiate this hypothesis.

A third limitation included implementation of the intervention package for a maximum of five weeks; thus reducing the ability to indicate, with a high degree of confidence, students’ rate of growth. The inclusion of standard error of measurement would in a short 5-week period, given the reported findings of this study, make it difficult to state with certainty how much, if any, students’ oral reading fluency improved. A more desirable timeline for measuring growth would have been a nine to twelve week period.

Unequal quantities of intervention between participants within each dyad compounded this limitation. The study was designed to collect 10 data points per participant per intervention phase. However, the study terminated after collection of 10 data points in Participant 1’s intervention phase; resulting in one to two fewer days exposure to the reading intervention package for three participants in this study. Much of
the concern posed in this study by unequal quantities of intervention is neutralized by calculating rate of ORF growth per week rather than simply stating the overall gain in CWPM for participants.

The fourth limitation pertained to use of a cognitive screener (i.e., KBIT-2) rather than a more comprehensive assessment tool. The universe of behaviors sampled on the screener was less than if a complete IQ test were used. Therefore it can not be stated with certainty that students falling in the LA group would not have demonstrated a 15-point discrepancy had a different, more comprehensive, cognitive assessment tool been selected and used.

The fifth limitation is the use of a self-administered checklist for implementation integrity. Though designed for efficiency of integrity recording, the tool contains a bias for over reporting implementation integrity. An alternative, though more labor intensive, option is the more traditional and objective manner of integrity checks that entail the use of an independent observer. Additionally, this study reports an atypical 100% compliance on integrity checks. It is not expected that, in practice, educators lacking extensive training and knowledge of CBM procedures will, over the course of five weeks, implement CBM procedures with this level of integrity. Additionally, ISA measures conducted only during the intervention phase. A better practice would have been conducting ISA during both intervention and baseline phase.

A sixth limitation involves a slight to modest increasing trend in the scores reported in baseline phase of Participant 2. Extending baseline and allowing a higher degree of stability may have allowed for a stronger argument that the intervention package was effective. However, it can not be stated, without caution, that Participant 2’s
increase in ORF during the intervention phase was due primarily to the reading intervention package.

**Future Research**

**Student Outcomes.** This study is but one piece of a larger puzzle that seeks to provide definitive answers as to whether students diagnosed, using an IQ-Achievement discrepancy model, as learning disabled consistently and meaningfully differ from those students considered low achievers or “slow learners.” Researchers have a number of avenues to pursue as the scientific field continues to address this issue. Possibly most important would be the replication and enhancement of reading studies reporting significant, nominal, or no differences between the growth rates of children identified as LD and LA. These replications, submitted to professional journals and critically reviewed by peers would substantiate or refute earlier findings, and thus guide evolving research.

This is particularly important as the RTI movement continues to spread its influence across education agencies nationwide. A foundational concept of RTI is that appropriate, empirically-based interventions are generally effective with students regardless of the students’ classification (e.g. learning disabled, low achieving, etc.). Opposition can be reduced if research continues to bolster this claim. However, diverging results from students with different classifications may lend credence to revisiting a model where IQ-A discrepancy is used to determine eligibility. A third possibility is the emersion of a new model for identifying the group(s) of students who would benefit most from supplemental and/or modified instruction. Such a model may come to prominence if RTI and IQ-A discrepancy models are rejected. The likely success and/or future of one or
more of these models is contingent upon the body of scientific research and the effectiveness of RTI within school systems.

*Research Based Intervention(s).* Research may focus on confirmation that this intervention package produces reading gains of such magnitude as reported in this study. As stated in the limitations, actual results may have been greater, less than, or the same as reported in this document. Using this intervention package with students of various ages, and with LD, LA, and non-diagnosed students will provide/fail to provide supporting evidence for using or choosing not to use this intervention.

*Reading Probes.* Within the school psychology and education fields, a debate has grown over the development and adoption of a state or national set of standardized reading probes/passages. The wheels of educational change are slow and companies selling commercial probes may oppose a universal probe set(s) if they stand to lose substantial monetary gains. Given these potential hindrances, determining more efficient and inexpensive means of developing probes should be examined. Ideally, entire probe sets would be normed for various populations. However, norming procedures can be time intensive and impractical in an academic environment where increasing demands are placed on each day's finite academic time.

In lieu of norming probes, indexing probes via entry into multiple readability programs may be a viable alternative. Procedures using combinations of multiple formulas may reveal the best option for determining probe difficulty. This is important in that it may aid in reducing the variability and standard error of reading passages. In turn, collapsing the range of possible scores for each datum point and allowing for better decisions to be made by teachers, psychologists, and local/state educational agencies.
Conclusion

The current study provides further support that IQ-A discrepancy model should not be the sole or primary determinant for identification of students eligible for supplemental reading interventions under the guise of special education. Greater than expected gains in ORF were observed for all study participants. However, the study revealed inconsistent findings for determining the classification of students (e.g., LD or LA) benefiting most from intervention.
APPENDIX A

The University of Southern Mississippi Consent Document for Research Participants: Teacher Consent

Title of Study: Examination of Level and Growth Rates: Identification of Learning Disability

Name of Participant: ______________________________ Date: _______________

Purpose of Study. You are being asked to take part in a study that assesses the validity of IQ-Achievement Discrepancy and Response to Instruction Models in identifying Learning Disabled students in Reading.

Who can participate? To be eligible for this study, students must be enrolled in third grade and fall into one of two categories. (A) Have a Special Education ruling of Learning Disability in Reading, based on the IQ-Achievement Discrepancy Model, and show a current discrepancy on IQ-Achievement screening, OR (B) Fall in the bottom 20th percentile of students administered the DIBELS Oral Reading Fluency subtest, demonstrate a skill or skill/performance deficit that suggests that the student is reading below a second grade level, and after referral to the study, have less than a 15-point discrepancy between IQ-Achievement scores. Stated differently, students without a Learning Disability in Reading must not meet criteria for reading disability after screening. Those students identified, during the screening process, as meeting the criteria will be referred to the teacher support team at their school, with their screening scores and recommendations for appropriate interventions. This does not mean that they will be diagnosed as Learning Disabled, only that they will be identified as a student that might need a more in-depth assessment for his/her reading difficulties.

Methods and Procedures. If you agree to participate, and your student is selected for the study, you will be asked to excuse the student 15-20 minutes per day, 3 days per week, during non-testing periods for approximately eight weeks. This time will be used to determine your child’s current reading level, baseline construction, and implementation of a reading intervention. The total anticipated time for completion of this study is approximately 10 weeks.

Risks and Discomfort. This study has very few risks for your student because the short period of time that he/she is removed for the classroom will be replaced by a 1:1 individually administered reading intervention. The intervention to be used has been deemed appropriate, by The National Reading Panel (2000), for remediating reading deficits. It is possible that your student will experience some level of frustration if he/she does not feel that he/she is making adequate progress. However, this frustration is
expected to be no greater than had the child experienced reading difficulty within the classroom. Attempts will be made to minimize distractions that may occur during the removal and return of the student to your classroom.

Benefits. The results of this procedure may benefit you and your student by (a) enhancements in the child’s reading fluency, (b) determination of reading level and growth rate, and (c) assessment of whether the reading performance of your student and the group that he/she is placed in differs from either children with IQ-Achievement discrepancies or from non-discrepant, low achieving students.

Confidentiality of Records. All information obtained during this study will be kept confidential. This means that your name and your student’s name and any other identifying information will be withheld from all persons not connected with the study. There are circumstances in which we are obligated to release information about you and your student. Those circumstances are if your student tells us, or if there is a suspicion, that he/she is planning to harm him/herself or someone else, if there is a suspicion that your student is abused or neglected, if we are ordered by the court to release information, or if there is a medical emergency in which the release of information is necessary in ensuring safety.

Voluntary Participation. Your participation in this study is voluntary. You may withdraw from being a research participant at any time without penalty or prejudice.

Participant’s Consent. I have had the purposes and procedures of this study explained to me and have had the opportunity to ask questions. My questions have been answered to my satisfaction, and I am voluntarily signing this form for myself to participate in this research study. My signature shows my willingness to participate in this study under the conditions stated. If I have any questions about this study, I can contact Odell Vining or Dr. Joe Olmi, at (601) 266-5255. This project and this consent form have been reviewed by the Human Subjects Protection Review Committee, which ensures that research projects involving human subjects follow federal regulations. Any questions or concerns about rights as a research subject should be directed to the Institutional Review Board Office, The University of Southern Mississippi, Box 5147, Hattiesburg, MS 39406-5147, (601) 266-6820. I have received a copy of this consent.

__________________________________________
Teacher Signature                                    Date

__________________________________________
Investigator Signature                              Date
APPENDIX B

The University of Southern Mississippi Consent Document for Research
Participants: Parent Consent

Title of Study: Examination of Level and Growth Rates: Identification of Learning Disability

Purpose of Study. You are being asked to allow your child to take part in a study that will look at whether students diagnosed with Learning Disability in Reading respond differently to intervention than Low Achieving, undiagnosed children receiving the same treatment. Oral reading fluency, or how quickly the child reads is the variable/item of interest during this study.

Who can participate? To be eligible for this study, students must be enrolled in third grade and fall into one of two categories. (A) Have a Special Education ruling of Learning Disability in Reading, based on the IQ-Achievement Discrepancy Model, and show a current discrepancy on IQ-Achievement screening, OR (B) Fall in the bottom 20th percentile of students administered the DIBELS Oral Reading Fluency subtest, demonstrate a skill or skill/performance deficit that suggests that the student is reading below a second grade level, and after referral to the study, have less than a 15-point discrepancy between IQ-Achievement scores. Stated differently, students without a Learning Disability in Reading must not meet criteria for reading disability after screening. Those students identified, during the screening process, as meeting the criteria will be referred to the teacher support team at their school, with their screening scores and recommendations for appropriate interventions. This does not mean that they will be diagnosed as Learning Disabled, only that they will be identified as a student that might need a more in-depth assessment for his/her reading difficulties.

Methods and Procedures. If you agree to allow your child to participate, and he/she is selected for the study, your child’s reading level will be assessed. He/she will be screened with an IQ and Achievement test. If study criteria is met; your child will be excused from class 15-20 minutes per day, 3 days per week, during non-testing periods for approximately eight weeks. This time will be used to determine your child’s current reading level, baseline construction, and implementation of a reading intervention. The total anticipated time for completion of this study is approximately 10 weeks.

Risks and Discomfort. This study has very few risks for your child because the short period of time that he/she is removed from the classroom will be replaced by a 1:1 individually administered reading intervention. The intervention to be used has been deemed appropriate, by The National Reading Panel (2000), for remediating reading deficits. It is possible that your child will experience some level of frustration if he/she does not feel that he/she is making adequate progress. However, this frustration is expected to be no greater than had the child experienced reading difficulty within the
classroom. Attempts will be made to minimize distractions that may occur during the removal and return of the child to his/her classroom.

**Benefits.** The results of this procedure may benefit to your child and his/her teacher by (a) enhancements in the child's reading speed/fluency, (b) determination of reading level and growth rate, and (c) assessment of whether the reading performance of your child and/or the group that he/she is placed in differs from either children with IQ-Achievement discrepancies or from non-discrepant, low achieving students.

**Confidentiality of Records.** All information obtained during this study will be kept confidential. This means that your name and your child's name and any other identifying information will be withheld from all persons not connected with the study. There are circumstances in which we are obligated to release information about you and your child. Those circumstances are if your child tells us, or if there is a suspicion, that he/she is planning to harm him/herself or someone else, if there is a suspicion that your student is abused or neglected, if we are ordered by the court to release information, or if there is a medical emergency in which the release of information is necessary in ensuring safety.

**Voluntary Participation.** Your child's participation in this study is voluntary. You may withdraw him/her being a research participant at any time without penalty or prejudice.

**Participant's Consent.** I have had the purposes and procedures of this study explained to me and have had the opportunity to ask questions. My questions have been answered to my satisfaction, and I am voluntarily signing this form for myself to participate in this research study. My signature shows my willingness to participate in this study under the conditions stated. If I have any questions about this study, I can contact Odell Vining or Dr. Joe Olmi, at (601) 266-5255. This project and this consent form have been reviewed by the Human Subjects Protection Review Committee, which ensures that research projects involving human subjects follow federal regulations. Any questions or concerns about rights as a research subject should be directed to the Institutional Review Board Office, The University of Southern Mississippi, Box 5147, Hattiesburg, MS 39406-5147, (601) 266-6820. I have received a copy of this consent.

<table>
<thead>
<tr>
<th>Investigator Signature</th>
<th>Date</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Parent/Guardian Signature</th>
<th>Date</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>
APPENDIX C

The University of Southern Mississippi Consent Document for Research Participants: Student Assent

Title of Study: Examination of Level and Growth Rates: Identification of Learning Disability

I am working on a school project. In some ways it is similar to the homework given at your school. I am hoping to find out if different types of students do better with certain kinds of reading help. You have been picked as one of the children to help me with this project. If you want to help, you will be excused from class for 15-20 minutes per day, during non-testing periods on three days per week. The project will probably take between eight to ten weeks for us to finish. On days that you are taken from class, we will practice reading, and then I will check to see if you are reading any faster. I do not give you a grade, and you stop at any time that you like.

If you have question about working in the project, please let me, Odell Vining, or my boss, Dr. Joe Olmi, at The University of Southern Mississippi School Psychology Service Center at (601) 266-5255 know that you have a question We will do our best to answer any question that you, your teacher, or your parents have about our reading project. Thank you for helping me with this project.

Odell Vining, M.S.
School Psychologist-In-Training

Assent Signature:

______________________________
Name

______________________________
Date
Appendix D

**CBM Administration Procedures**

The administrator will read aloud a set of standardized directions before each student is instructed to read aloud from ORFP. The instructions are as follows:

"When I say “begin,” start reading aloud at the top of the page. Read across the page (DEMONSTRATE). Try to read each word. If you come to a word you don’t know, I’ll tell it to you. Be sure to do your best reading." (National Association of School Psychologists, 2005).

The student will be instructed to “begin” and the 1-minute timing will begin. If a student takes more than three seconds on a word, the word will be provided to the student. A slash will be placed across omissions and words verbalized incorrectly. After sixty seconds, the student will be instructed to stop, and a bracket will be placed after the last word read. Total words read, errors, and CWPM will be recorded. The median CWPM from that day will be selected and plotted by the examiner.
**APPENDIX E**

**Integrity Checklist**

Examiner: Complete this checklist after completed administration of CBM and before returning the student back to his class. Integrity checks should be completed on no less than 50% of CBM administrations.

<table>
<thead>
<tr>
<th>Yes</th>
<th>No</th>
<th>NA</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>1. Examiner places probe flat on the table in front of student.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>2. Directions are read verbatim.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>3. Stopwatch is started when the examiner tells student to begins reading.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>4. Misread and omitted words are marked as incorrect.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>5. The correct word is provided after a 3-second delay by the student.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>6. Student is instructed to stop reading after 60-seconds [+- 2 seconds].</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>7. A bracket is placed after the last word read.</td>
</tr>
</tbody>
</table>
APPENDIX F

Repeated Reading Administration Procedures

Administration involves having a student read and re-read an instructional level passage multiple times. When the reader is able to fluently read the passage to a specified level of competency, a new passage is introduced. This study will use the upper limit of Howell and Nolet's (Howell, 2005) recommended instructional criteria for progressing to a more difficult passage, (i.e. 100 words with less than seven errors per minute).

Table 3 provides a baseline reading datum as derived from the median of the last three scores recorded in the baseline phase, the datum for each child taken from the median of the last three recorded data points in the intervention condition, and respective growth rates for each student.
APPENDIX F

HUMAN SUBJECTS PROTECTION REVIEW COMMITTEE
NOTICE OF COMMITTEE ACTION

The project has been reviewed by The University of Southern Mississippi Human Subjects Protection Review Committee in accordance with Federal Drug Administration regulations (21 CFR 26, 111), Department of Health and Human Services (45 CFR Part 46), and university guidelines to ensure adherence to the following criteria:

- The risks to subjects are minimized.
- The risks to subjects are reasonable in relation to the anticipated benefits.
- The selection of subjects is equitable.
- Informed consent is adequate and appropriately documented.
- Where appropriate, the research plan makes adequate provisions for monitoring the data collected to ensure the safety of the subjects.
- Where appropriate, there are adequate provisions to protect the privacy of subjects and to maintain the confidentiality of all data.
- Appropriate additional safeguards have been included to protect vulnerable subjects.
- Any unanticipated, serious, or continuing problems encountered regarding risks to subjects must be reported immediately, but not later than 10 days following the event. This should be reported to the IRB Office via the "Adverse Effect Report Form".
- If approved, the maximum period of approval is limited to twelve months. Projects that exceed this period must submit an application for renewal or continuation.

PROTOCOL NUMBER: 26021302
PROJECT TITLE: Response Outcomes to Evidenced-Based Treatments: Learning Disability Versus Low Achieving
PROPOSED PROJECT DATES: 02/13/06 to 12/23/06
PROJECT TYPE: Dissertation or Thesis
PRINCIPAL INVESTIGATORS: Odell Vining
COLLEGE/DIVISION: College of Education & Psychology
DEPARTMENT: Psychology
FUNDING AGENCY: N/A
HSPRC COMMITTEE ACTION: Exempt Approval
PERIOD OF APPROVAL: 02/13/06 to 02/12/07

Lawrence A. Hosman, Ph.D.
HSPRC Chair

Date
REFERENCES


Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.


Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.


