

City University of New York (CUNY)

CUNY Academic Works

Dissertations, Theses, and Capstone Projects

CUNY Graduate Center

9-2015

Effectiveness of a Partial Read-Aloud Test Accommodation to Assess Reading Comprehension in Students with a Reading Disability

Michelle Marie Giusto

Graduate Center, City University of New York

[How does access to this work benefit you? Let us know!](#)

More information about this work at: https://academicworks.cuny.edu/gc_etds/944

Discover additional works at: <https://academicworks.cuny.edu>

This work is made publicly available by the City University of New York (CUNY).

Contact: AcademicWorks@cuny.edu

EFFECTIVENESS OF A PARTIAL READ-ALoud TEST ACCOMMODATION TO ASSESS
READING COMPREHENSION IN STUDENTS WITH A READING DISABILITY

By

Michelle Giusto

A dissertation submitted to the Graduate Faculty in Educational Psychology in partial fulfillment
of the requirements for the degree of Doctor of Philosophy, The City University of New York.

2015

© 2015

Michelle Giusto

All Rights Reserved

This manuscript has been read and accepted by the Graduate Faculty in Educational Psychology
in satisfaction of the dissertation requirement for the degree of Doctor of Philosophy

September 11, 2015

Date

Dr. Linnea Ehri

Chair of Examining Committee

September 11, 2015

Date

Dr. Bruce Homer

Executive Chair

Dr. Linnea C. Ehri, Distinguished Professor

Dr. Helen Johnson, Professor

Dr. Lois Dryer, Professor

Supervisory Committee

THE CITY UNIVERSITY OF NEW YORK

Abstract**Effectiveness of a Read-Aloud Test Accommodation to Assess Reading Comprehension in
Students with a Reading Disability****By Michelle Giusto**

Adviser: Dr. Linnea Ehri

This study explored the potential effectiveness of a partial read-aloud accommodation on the reading comprehension scores of third grade students classified as poor decoders. Past research has explored the use of an accommodation in which the test items are read aloud to students. These studies have demonstrated that reading an entire test aloud results in gains for both students with reading disabilities and their peers reading on grade level, thus invalidating this procedure as an appropriate test accommodation. To be appropriate, a test accommodation must benefit only the students with reading disabilities, not their grade level peers. Previous research has not explored the procedure of reading only portions of a test. A partial read-aloud accommodation with pacing requires that the examiner read aloud only directions, question items, multiple choice answers and proper nouns, while students are responsible for reading the passages independently. A field research study showed that struggling third grade readers given the partial read-aloud condition demonstrated greater gains in reading comprehension test scores than their grade level peers. The present study explored the issue further by comparing a partial read-aloud with pacing condition to a standard testing condition and a pacing only condition. The latter condition involved teachers guiding the students through the passages and questions, one by one on a schedule, without reading any items aloud. Participants included 82 third grade students (28 poor decoders and 54 average readers) from two schools and one summer program in Queens, New York (Mean Age = 8 years, 9 months). Results revealed a significant interaction

($p < .01$) between test condition and student classification. Poor decoders showed a greater increase in their test scores under the partial read-aloud with pacing condition than under the pacing only or standard conditions, whereas average readers did not benefit from the partial read-aloud with pacing procedure. This study supports the use of a partial read-aloud with pacing accommodation to help level the playing field for students who struggle with decoding on tests of reading comprehension. Results suggest the value of determining the nature of a student's reading disability before identifying a valid and appropriate test accommodation.

Keywords: test accommodations, learning disabilities, reading disabilities, standardized tests, reading comprehension, listening comprehension, decoding.

Acknowledgments

The completion of my dissertation has been a long, rewarding endeavor. I would like to take this opportunity to recognize all of those people who helped me to successfully reach this goal.

Primarily, I would like to recognize my advisor, Dr. Linnea Ehri. Dr. Ehri is an accomplished, dedicated, inspiring professional. She has helped to keep me focused, grounded, and eager to learn, and for that I am incredibly grateful. I would also like to thank the other members of my committee, Dr. Helen Johnson and Dr. Lois Dreyer. Your assistance and feedback are much appreciated.

I would also like to thank my colleagues, participants, and peers. The administrators, camp directors, teachers and staff members who willingly participated in my research have shown me incredible support, and always welcomed me without hesitation. Additionally, I would like to thank my fellow students at The Graduate Center, who have consistently offered useful feedback to help make my research ideas work more efficiently.

Finally, I am very thankful to my family. My husband, Stephen, and daughter, Maria, were consistently supportive, and never complained about the gratuitous amount of time I spent on coursework and research. I owe immense gratitude to my incredible parents, James and Regina Montefinise, who have offered me so much help along the way that I do not imagine reaching this goal without them. Lastly, thanks to my siblings, grandparents, aunts, uncles, cousins, and friends for your support and unwavering faith in my ability to complete this dissertation. Your support has brought me to where I am today.

Table of Contents

Abstract	i
Acknowledgements	iii
Chapter	
I. Introduction	1
II. Literature Review	6
Terminology and Classification – What is Meant by a “Reading Disability?”	6
The Development of Reading Accommodations and Extended Time	11
The Read-Aloud Accommodations in Tests of Mathematics	14
The Read-Aloud Accommodations in Tests of Reading Comprehension	19
Research Summary	29
III. Field Study, Rationale and Hypotheses	31
Variables, Hypotheses and Research Questions	31
Methods	33
Results	36
Conclusions and Questions for Further Study	40
IV. The Current Study	43
Variables, Hypotheses, and Research Questions	43
V. Methods	46
Participants	46
Pretests	48
Reading Comprehension Test	52
Experimental Treatment	53

Procedures	55
Data Analysis	56
VI. Results	57
Characteristics of Participants	57
Effects of Reading Ability and the Accommodation Treatment Variable on Reading Comprehension	60
Correlations Between Tests	65
Anecdotal Observations	66
VII. Discussion	67
Student Ability Classifications	67
Research Question 1	69
Research Question 2	70
Limitations	73
Educational Implications	76
References	89

List of Tables

Table 1	<i>Characteristics of Participants – Field Study</i>	37
Table 2	<i>Mean Scores of Average Readers, Poor Readers, Those Assigned to Groups 1 and 2 on the Gates-MacGinitie Reading Test on the Field Test</i>	38
Table 3	<i>Analysis of Variance for Reading Group and Ability on the Field Test</i>	39
Table 4	<i>Characteristics and Numbers of Participants Contributed by Each Cohort with Percentages in Parentheses Indicating the Preponderance of Characteristics Within Each Cohort</i>	57
Table 5	<i>Characteristics of Groups of Poor Readers and Average Readers</i>	58
Table 6	<i>Analysis of Variance for Testing, Condition, and Reading Ability of Participants</i>	59
Table 7	<i>Analysis of Variance for Testing Condition and Reading Ability on the Gates MacGinitie Test.</i>	61
Table 8	<i>Mean Performance and Standard Deviations as a Function of Reading Ability on the Gates-MacGinitie Reading Comprehension Test</i>	62
Table 9	<i>Mean Performance of Students from Different Linguistic Backgrounds on the Gates-MacGinitie Reading Comprehension Test</i>	63
Table 10	<i>Correlations Between Pretests and Gates-MacGinitie Reading Comprehension Test</i>	65

List of Figures

Figure 1 <i>Mean Gates-MacGinitie Reading Comprehension Scores across Conditions for Poor Readers and Average Readers on the Field Study</i>	39
Figure 2 <i>Group Means under the Standard, Pacing and Read-Aloud Conditions on the Gates-MacGinitie Reading Comprehension Test, Form S</i>	62

List of Appendices

Appendix A	<i>Information Form for Administrators</i>	80
Appendix B	<i>Parent Permission Form</i>	82
Appendix C	<i>Directions for the Partial Read-Aloud Condition of the Gates-MacGinitie Reading Comprehension Test, Form S</i>	85
Appendix D	<i>Directions for Pacing Conditions of the Gates-MacGinitie Reading Comprehension Test, Form S</i>	87
Appendix E	<i>Student Assent Form</i>	88

Chapter 1 - Introduction

The purpose of this study was to explore the effect of a partial read-aloud with pacing accommodation on the reading comprehension scores of students with weak decoding skills. Effects of this accommodation were compared to two conditions - the standard condition, in which no accommodation was provided, and the pacing only condition, in which the researcher guided students through the reading comprehension test, prompting them when to begin each story and answer test items, but did not read test items aloud.

Since the development of the Individuals with Disabilities Education Act (IDEA), students with learning disabilities have been offered test accommodations in order to help eliminate any unrelated barriers that prevent them from showing their true understanding of the content area. In order for an accommodation to be considered valid, it must provide support to those students who are in need of it, rather than improve scores for all test takers. In this way, it bridges the gap between students with learning disabilities and their mainstream peers rather than enhancing the performance of all test takers and maintaining that gap. (Fletcher, Frances, Boudousquie, Copeland, Young, Kalinowski, & Vaughn, 2006; Thurlow, Lazarus, Thompson, & Morse, 2005).

While extended time has frequently been offered as an accommodation, for students with difficulty decoding, this has often been ineffective, or even counterproductive, resulting in lower test scores (Elliot & Marquart, 2004; Gregg & Nelson, 2012). As an alternative, another accommodation has been considered. This is the read-aloud accommodation, in which struggling readers are read-aloud questions before answering them independently. This accommodation removes the barrier of decoding for students who struggle with it so that they can more readily show their understanding of the content being tested. This has proven to be an

effective and appropriate accommodation in tests of mathematics (Helwig & Tindal, 2003; Helwig, Rozeck-Tedesco, & Tindal, 2002; Schulte, Elliot, & Kratochwill, 2001). However, this accommodation is not currently offered on tests of reading comprehension. Some theorists have argued that reading aloud passages and items on a test of reading comprehension alters the construct being measured, invalidating the test. Specifically, reading aloud the test content would not enable students to truly demonstrate their skill in reading independently since the task of reading is done for them (Fuchs, Fuchs, & Capizzi, 2005; Melov, L, 2002). In addition, research exploring this accommodation has yielded unsuccessful results in some studies (Elbaum, Arguelles, Campbell, & Saleh, 2004; Fletcher et al., 2006; McKevitt & Elliott, 2003). The read-aloud accommodation has been found to benefit both students with a reading disability and their grade level peers equally, invalidating it as an appropriate accommodation.

Further, the nature of a reading disability has been defined very broadly, and many studies have lumped students with reading disabilities in with students having other learning or physical disabilities. Inconsistencies across studies in defining the term “reading disability” have been identified as a contributing factor to the mixed results in current research. (Elbaum, Arguelles, Campbell, & Saleh, 2004; Fletcher et al., 2006; McKevitt & Elliott, 2003).

Meanwhile, the Simple View of Reading (SVR), described by Gough and Tunmer (1986), suggests that the nature of a reading disability varies from one case to the next. According to this view, reading comprehension is a product of decoding skill and linguistic comprehension. This suggests that reading disabilities could include low decoding and average comprehension, average decoding and low comprehension, or poor decoding and low comprehension together (Gough & Tunmer). Therefore, it is important to describe the nature of the reading disability being explored at the onset of any study concerning test accommodations

designed for this disability in order to ensure that the results generalize to the appropriate population.

Past studies have not explored an accommodation that involves reading aloud only portions of a reading comprehension test. In a 2012 field study I conducted, I compared the performance of good and poor third grade readers on the Gates-MacGinitie Reading Comprehension test (MacGinitie, MacGinitie, Maria & Dreyer, 2000) under two conditions – standard administration and a partial read-aloud accommodation. Under the latter condition, only question items, multiple choice answers and proper nouns were read-aloud, and students were left to read the test passages independently. Results demonstrated a significant interaction between ability and test condition. Students who were identified as those with a reading disability showed statistically higher scores under the accommodated condition than the standard condition. In contrast, students who did not struggle in decoding performed almost the same under the partial read-aloud condition and the standard condition. Thus, findings indicated that the partial read-aloud accommodation helped only the struggling readers, not other test takers. However, there was a confound. Under the accommodated condition, students were paced and prompted as to when they should read each item, choose an answer, and move on to the next passage. Thus, it was unclear whether it was the reading aloud of test items or the pacing that led to the improved scores of struggling readers.

Therefore, the current study sought to explore this issue further. This study was designed to investigate a variation of the read-aloud accommodation, specifically a partial read-aloud with pacing condition. Under this test accommodation, poor decoders taking tests in reading comprehension were read test questions and proper nouns and paced through the assessment, but they read the exam passages independently. In this study, three conditions were compared.

Students with poor decoding skill and average readers were given the same test under one of three possible conditions. These included no accommodation, or the standard condition, the partial read-aloud with pacing condition, or the pacing only condition, in which the researcher guided students through the test but questions were not read aloud. The purpose was to see whether the partial read-aloud with pacing condition would improve the reading comprehension test scores of poor decoders compared to the pacing and the standard conditions. It was also of interest to compare struggling readers and average readers' performance in each condition to determine whether the partial read-aloud condition with pacing would benefit struggling readers but not average readers. It was expected that positive findings would serve to identify a better option for poor decoders faced with the struggle of standardized testing.

The research questions guiding the study were:

- Does the partial read-aloud with pacing condition improve performance on a reading comprehension test more than the pacing only condition and a standard condition?
- Do poor decoders show significantly greater improvement in reading comprehension with the partial read-aloud with pacing condition than with pacing only and standard conditions, whereas average readers show little improvement in reading comprehension with the partial read-aloud with pacing condition compared to other conditions?

The hypotheses to be tested were:

- The partial read-aloud with pacing condition will boost the reading comprehension scores of poor decoders compared to the pacing only and standard conditions, whereas the

partial read-aloud with pacing condition will provide little improvement in reading comprehension compared to the pacing only and no accommodation condition for average readers. This will show that the benefit of the partial read-aloud with pacing condition is limited to poor decoders and hence constitutes a valid test accommodation.

- The partial read-aloud with pacing condition will boost the reading comprehension scores of poor decoders more than the pacing only condition, showing that the benefit of partial accommodation is not simply a result of pacing students through the test.

Chapter 2 – Literature Review

Terminology and Classification – What Is Meant by “A Reading Disability?”

The current study focuses on students who are identified by the researcher as having a reading disability. The nature of this classification, which had impacted the sample of participants chosen, is based on past literature exploring this concept and the nature of its formal definition in research and practice.

Research on reading disabilities has explored how to define and distinguish students with a “reading disability.” One issue concerns the inconsistency of characteristics used to identify those who fit this classification. The precise definition of a learning disability varies across states, and from one school district to another (Spear-Swerling & Sternberg, 1996). However, research has shown the significance of precisely identifying what is meant by a “reading disability,” as there are different classifications falling under this umbrella.

The varying nature of reading disabilities is depicted in the Simple View of Reading (SVR), described by Gough and Tunmer (1986). According to this view, two fundamental components of reading contribute to reading comprehension skill – word recognition, or decoding, and linguistic comprehension, which involves using lexical or word level information to achieve sentence and discourse interpretations. Follow-up research studies have measured linguistic comprehension through tests of listening comprehension (Dreyer & Katz, 1992; Aaron, Joshi & Williams, 1999; Nation & Snowling, 2007; Tilstra, McMaster, Van den Broek, Kendeou & Piapp, 2009). Gough and Tunmer devised a formula, $R = D \times L$, with the value of D and L ranging from 0 to 1. Therefore, according to this theory, any individual with a decoding or linguistic comprehension score of 0 will likewise have a reading comprehension score of 0, because it is not possible to comprehend text without each of these components. This also implies that poor reading comprehension results from more than one possible factor; it could be

due to deficiencies in decoding, reading comprehension, or both, which would be classified as the “garden variety” reading disability (Gough & Tunmer).

Current research studies have supported the implications of SVR, illustrating that there is more than one type of reading disability. A study by Aaron, Joshi and Williams (1999) was designed to explore the varied nature of what constitutes a reading disability. The SVR served as a basis for their study. They reasoned that if poor reading comprehension can result from either poor decoding, poor listening comprehension, or both, then there should be some reading disabilities involving insufficient decoding skill but adequate listening comprehension, and others with adequate decoding but poor listening comprehension. The researchers worked with 139 children in grades 3, 4 and 6. The children were administered several pretests. These included the Reading Comprehension and Listening Comprehension tests from the Woodcock Reading Mastery Test (WRMT-R) (Woodcock, 1987), a vocabulary subtest from the Gates-MacGinitie Reading Test (MacGinitie et al., 2002), and a series of tests devised by the researchers. These included a list of nonwords, a test of orthographic processing, a list of irregular words, a list of content words, and a list of function words. “Poor readers” were those who scored below one standard deviation of the mean on the WRMT reading comprehension test. Further, in all subtests, those identified as poor in a particular skill were students who scored one standard deviation below the mean in the corresponding subtest. After analyzing the data, it was found that those classified as “poor readers” overall varied in their areas of weakness. Some were poor in decoding alone, others struggled only with comprehension, some were identified as poor in a combination of decoding, orthography and fluency, while a few did poorly in all measures. Therefore, this research supports the implications of the Simple View;

reading disabilities vary in their nature, and, while some struggle in all areas, others have difficulty with decoding or listening comprehension alone.

Dreyer and Katz (1992) also conducted a study to assess the generality and predictive validity of the SVR. To do so, they worked with 137 monolingual students in their third and fifth grade years. They gave each student a 60-item decoding test with low frequency, phonetically regular, single-syllable words. They also used data taken from a standardized test given to the students, the Educational Records Bureau Comprehensive Testing Battery (Educational Testing Service, 1987). This test included both reading comprehension and listening comprehension components. Pearson product moment correlations showed that decoding and listening comprehension were highly related to reading comprehension. Hierarchical multiple regressions also showed that decoding and listening comprehension combined accounted for 43.9% of the variance in reading comprehension. These findings support the Simple View, which claims that both decoding and linguistic comprehension can be used to predict reading comprehension. Further, this study yielded another interesting finding – the correlation between listening comprehension and reading comprehension increased as students went from third grade to fifth grade. This is an important factor to consider for the current study, which is seeking to explore an accommodation for struggling readers with poor decoding skills. If listening comprehension becomes more highly correlated with reading comprehension over time, then this accommodation would be especially helpful to students capable of showing comprehension while listening, but struggle to do so while reading due to impaired decoding skills.

A third study yielded similar results. Tilstra, McMaster, Van den Broek, Kendou, and Piapp (2009), conducted a study to review the predictive value of the SVR. Tilstra et al. studied

271 children in 4th, 7th and 9th grades. They gave all participants a series of measures in different reading tasks – the Gates-MacGinitie Reading Comprehension test, the listening comprehension subtest of the Iowa Test of Basic Skills, the Curriculum Based Measurement (CBM) Maze task, which asked students to fill in incomplete sentences with one appropriate word from a list of choices, the CBM oral reading task, which required students to model how many words they could read in a short time frame, and the Word Attack subsection of the Woodcock Johnson III. Results of the study showed that 66.88% of the variance in reading comprehension scores was accounted for by decoding and listening comprehension scores in grade 4. This further supports the SVR model. In addition, this study also showed the increased significance of listening comprehension as children got older. From fourth to seventh grades, the predictive validity of listening comprehension increased, while that of decoding decreased. Therefore, listening comprehension impacts overall comprehension somewhat independently of decoding. This justifies exploring accommodations that would help readers capable of listening comprehension while struggling in decoding.

From the research listed above, several conclusions can be drawn. For one, both decoding and listening comprehension are strong predictors of overall reading comprehension. Further, decoding and listening comprehension are seen to contribute independently to reading comprehension, suggesting that each has a distinct impact on reading comprehension ability. Finally, as the SVR implies, there is more than one reading disability. Some students are poor in decoding, others struggle in comprehension, and some have difficulty with both. For the purposes of the current study, I am interested in students who struggle in decoding, but who show grade level listening comprehension skill

The current study sought to explore poor decoders, often identified in research as “dyslexic,” for several reasons. Primarily, as the research above demonstrates, as students get older, the importance of linguistic comprehension becomes more significant. According to Nation and Snowling (2007), “Dyslexic children may perform poorly on sentence completion tasks not because they do not understand the sentences, but because of the heavy demands placed on their decoding skills. In such a case, it would be incorrect to interpret poor performance as evidence of comprehension weakness.” Therefore, the current study intended to see if these students could better show their comprehension under a test condition that would remove the barrier of decoding.

Further, research has shown that a reading disability involving poor decoding is most prominent. Meta-analyses of past research studies have demonstrated that most reading disabilities are associated with deficiencies in word-recognition processes and phonological awareness, which often lead to problems in comprehension as children get older (Spear-Swerling & Sternberg, 1996; Stanovich, 1991). If this accounts for such a great percentage of those currently classified as having a reading disability, then an accommodation designed to support this group would be important to explore and identify.

Therefore, for the purpose of the current study, I have recruited those students most often classified in research as dyslexic. This group was classified as “poor decoders” for the purpose of this study. While there are some poor decoders who also struggle in comprehension, for this particular study, the term “poor decoders” was used to identify students who have average comprehension skills while struggling in decoding. I chose to include students with or without formal Title I Special Education Services in my group of poor decoders for several reasons. A student who is identified as having a reading disability through an academic evaluation may not

necessarily struggle in decoding. The nature of that child's disability may also include low comprehension, either on its own or coupled with poor decoding (Spear-Swerling & Sternberg, 1996). The accommodation explored in the current study is intended to help students who struggle on standardized tests because of their deficiencies in decoding, particularly if they score high on measures of listening comprehension. Finally, other factors may impact whether or not students with reading disabilities receive special education services. For instance, parents have a legal right to refuse services for their children, or some teachers may not feel that certain characteristics merit a formal academic evaluation. Thus, it is possible that there are students with reading disabilities who are enrolled in general education classrooms and have either slipped through the cracks of the evaluation system, have not been referred, or have been denied special education services by their parents or legal guardians. I do not wish to omit these students from my study, as I hope to determine whether or not this accommodation benefits students with reading disabilities that fit a wider classification.

For the purposes of the current study, therefore, students identified as poor decoders are those who demonstrate grade-level listening comprehension, but below-grade level decoding. While past research described in the next sections has yielded mixed results for the read-aloud accommodation, the expectation is that a more clearly defined reading disability can better assess its effectiveness.

The Development of Reading Accommodations and Extended Time

It is only recently that students with disabilities were mandated to take standardized tests, and that their performance on high stakes tests was taken into account. In the past, students with

disabilities were not tested, but the No Child Left Behind Act of 2002 changed that. Since this act has been established, students with disabilities must be tested “to the fullest extent possible” (Fuchs & Fuchs, 2004). However, because these students have disabilities, they are unable to participate in assessments in the same fashion as their nondisabled peers. Therefore, accommodations are necessary to “level the playing field,” removing an irrelevant barrier that hinders a student’s ability to show his or her true knowledge of skills being tested (Tindal & Fuchs, 1999).

However, it is also important that the accommodations provide fair assessments of the construct being measured. According to Fuchs et al. (2004), valid accommodations “help students with disabilities demonstrate their knowledge and produce scores that evaluate the same constructs that are intended with standardized measurements of nondisabled peers.” If these accommodations either withhold or overestimate a student’s abilities, then they can be considered unfair. Accommodations that overestimate abilities are those which are equally beneficial to everyone rather than specifically helpful to students with disabilities. Therefore, in exploring the partial read-aloud accommodation as well as alternatives, the current study sought to determine whether it would benefit everyone in the same way or level the playing field for poor decoders. This is particularly important since current research demonstrates that other accommodations currently in place do not always benefit struggling readers as well as they should.

One accommodation which is widely used as an alternative to the partial read-aloud condition is extended time. Students with reading disabilities are often given extra time on exams, which is suggested as something that will help them more than their nondisabled peers. In a study by Elliot et al. (1998), the researchers sought to investigate the fairness of this

accommodation by testing 97 eighth grade students, 23 receiving special education services for learning, emotional, behavioral, physical or speech and language disabilities, and the remainder without any disabilities. The independent variables were disability status and test time allowed. All students completed two equivalent forms of a standardized mathematics test developed from TerraNova Level 18, each under a different set of conditions. One group of students took the standard, 20-minute administration for Form A and an extended time accommodation for 40 minutes using Form B. The second group took Form A with extended time and Form B under a standard format. It was predicted that disability status and mathematical skill level would interact, with the effectiveness of extended time operating as a function of a student's disability status. However, this hypothesis was not supported by the data; a MANOVA was used to compare the difference in performance of students with and without disabilities across the testing conditions. There was a main effect for test condition, but no significant interaction between testing condition and student group. In fact, the increase in scores obtained by students with disabilities was slightly less than the increase in students without a reading disability. Because this commonly given assessment was found to benefit everyone, extended time may not be a valid accommodation for students with disabilities.

A recent meta-analysis further calls into question the validity of extended time (Gregg & Nelson, 2012). This meta-analysis analyzed the effect sizes of 132 studies. Studies included involved participants in at least 9th grade, with 9 or more learning disabled students in the sample population, in tests of reading, writing, and math. Comparisons between groups included those between learning disabled students and typically achieving students in standard conditions and extended time conditions. Findings yielded mixed results; one common feature was that, under all conditions and subject areas, typically achieving participants outperformed learning

disabled students. Thus, the validity of extended time was called into question, since it consistently benefitted all students, rather than just improve the scores of the learning disabled students for whom it was intended. (Gregg & Nelson, 2012).

The mixed results and questioned validity of the extended time accommodation seems problematic, particularly since this is such a widely used accommodation. One possible explanation is that extending time does not facilitate test taking for a group of students with reading disabilities. These students, particularly those who have difficulty decoding words, may experience fatigue at being asked to decode a series of questions, and this may hinder their performance. Extending the time may further prolong this fatigue, which can exacerbate a decline in performance. (Fletcher, Francis, Boudousquie, Copeland, Young, Kalinowski & Vaughn, 2006). Therefore, for this group of students, it has been suggested that reading questions aloud can enable them to show their understanding of a content area without the barrier of decoding standing in their way. This is particularly emphasized in the field of mathematics.

The Read-Aloud Accommodation in Tests of Mathematics

Research has shown that the read-aloud accommodation given during mathematics tests has leveled the playing field for students with reading disabilities without having a similar effect on their nondisabled peers. This finding has largely influenced the hypotheses tested in the current study. It has demonstrated that reading questions aloud has been particularly helpful to struggling readers and prompted curiosity regarding whether or not the same outcome would result if a similar test accommodation were given on reading comprehension exams.

A study by Schulte, Elliot and Kratochwill (2001) examined the impact of reading items aloud to students compared to other accommodations given during mathematics exams in order

to determine which would be most effective. The researchers worked with a sample of 85 4th graders from Wisconsin and Iowa. Of these students, 43 had learning disabilities, mild cognitive disabilities or speech and language impairments. Each student with disabilities was randomly matched with a student without disabilities. The purpose of the random matching was to determine whether or not students who have not been mandated to receive an accommodation would benefit just as well from it as those who have. Each member of the pair took two forms of the TerraNova Multiple Assessment Battery Mathematics Subtest under different conditions - a standard format without any testing accommodations, and an accommodation format, in which the specific accommodation given to each pair was determined by the individual needs of the student with disabilities as mandated by his or her IEP. These accommodations included one or more of the following: extra time, reading test items aloud, testing in a separate location, positive praise given during testing, frequent breaks, paraphrasing test items, defining vocabulary and the use of manipulatives. The most common accommodation was extended time accompanied by the reading aloud of test items. The researchers sought to determine whether or not the accommodations benefited students with disabilities more so than their nondisabled peers.

An ANOVA and effect sizes were used to analyze the effect of the accommodations on student test scores. Results indicated that, while both groups benefited under the accommodated conditions, for most accommodations, the scores of students with disabilities benefited more than those of students without disabilities. Surprisingly, the exception to this existed within the most common accommodation - extended time coupled with reading test items aloud. Within this condition, there was no differential impact of accommodations for students with disabilities compared to students without disabilities.

This research, however, does not necessarily challenge the benefits of a read-aloud accommodation, for there are several factors to consider. Primarily, the accommodations granted were those recommended on the IEPs of the participants with disabilities; there is no indication that those recommendations are the most effective. In fact, research by Spear-Swerling and Sternberg (1996) shows that these recommendations may not be best suited for students with reading disabilities. Further, the study by Elliot et al. (1998) demonstrated that a commonly recommended accommodation is not always the most beneficial. The Elliot et al. study also indicated that extended time may be frustrating for students with reading disabilities. Finally, this study put all learning disabled students together, and did not distinguish between the varying needs of each student. A read-aloud accommodation, like the one being explored in this study, would likely be more beneficial to students who struggle specifically with decoding than it would for students who demonstrate low levels of overall comprehension. Therefore, students with a reading disability taking only the read-aloud accommodation may perform better than those who take this accommodation paired with extended time. Since the Elliot et al. study showed the potentially negative impact of giving students with reading disabilities extended time, the current study will not pair that condition along with the read-aloud accommodation beyond the incidental time increase it will take to read the items aloud.

One study exploring the sole impact of the read-aloud accommodation given during math tests was conducted by Helwig and Rozek-Tedesko (2002). This study examined the performance of 1,343 general and special education students from 8 states broken up into two cohorts – middle school and elementary school. The middle school students were in grades 7 and 8, while the elementary school students were in grades 4 and 5. All students were predominantly white, while 40% qualified for free or reduced price lunch.

Within each cohort, students were given two 30-item forms of a mathematics test devised from one of the participating state's standardized math exams. Test items included basic operations, whole numbers, fractions and decimals, and graphic and geometric functions. All students took each form of the test in separate sessions 1-4 days apart. One form was taken according to the standard set of instructions, and the other form was taken with the read-aloud accommodation. This method included questions that were shown on a video monitor. As students viewed the monitor, the questions were read aloud by a prerecorded voice. Words were only shown on the screen as they were read. Within this condition, students were given booklets with one test item per page. Students were given the option of following along with the monitor, following along in the booklet, or ignoring both and going at their own pace for each item, only going to the next item when the proctor instructed them to do so. Students were randomly assigned to one of two groups, each of which took each form under a different administration: some took form A under standard administration and B as the read-aloud form, while the other half did the opposite. A 2-factor ANOVA was run, using reading status as a between-participants variable and test format as a within participants variable. Results of the study showed a significant main effect for student type, no main effect for test format, and a significant interaction between student type and test format. Students with disabilities improved more under the video accommodation than students without disabilities.

This study was extensive, involving a very large sample, which makes its results more generalizable and a noteworthy contribution to the study of test accommodations. The results of this study showed not only that the read-aloud accommodation was effective for students with disabilities, but also that it did not equally help their mainstream peers. Therefore, unlike the

extended time accommodation, this seems to be a more valid accommodation for students with reading disabilities.

However, there are a few factors to consider. The use of video presentation, while helpful in standardizing the test, can be costly. It is more common and feasible for accommodations to have the proctor read the questions aloud. It is not known whether some benefit will result from this procedure. In addition, this study explored a wide range of disabilities, rather than reading or math disabilities in particular. Therefore, it is unclear whether the benefits of this accommodation may differ depending on the specific needs of each child. Finally, these students were given accommodations based upon their IEPs, which are granted after an evaluation prompted by an educator's referral. One may wonder whether this accommodation might benefit other students who struggle in reading, but who have not been formally identified as having a reading disability on an IEP. It is for this reason that the study being proposed here specifically narrows in on students who struggle with decoding.

A second study by Helwig and Tindal (2003) addressed some of these points. This study used the same sample population and test forms as the 2002 study done by Helwig et al. However, this study incorporated the use of teacher rating surveys, evaluating the performance of students in reading and math. Teachers were given a Likert-based scale to rate how well they could evaluate the suitability of accommodations for individual students. After teachers evaluated the needs of their students for receiving accommodations, classrooms were randomly assigned to conditions described in Helwig et al.'s 2002 study. Results of the study showed that the teachers' predictions were ineffective; ratings coincided with actual performance of students on tests only half of the time. More specifically, 14% of the elementary school sample and 88% of the middle school sample performed better under the read-aloud condition, whereas their

teachers predicted that this accommodation would make very little difference for these students. It is evident that a teacher's recommendation alone may not lead to the most appropriate testing accommodations for students with disabilities. Therefore, while the read-aloud accommodation may not be as widely recommended in student IEPs as extended time, it is still possible that this is the most effective accommodation for students with reading disabilities.

The research reported on the read-aloud accommodation in math supports several points. Primarily, it shows that this accommodation does not uniquely benefit students with disabilities when it is paired with extended time. Further, it demonstrates that, quite often, this accommodation is not appropriately recommended by teachers; it can either be overlooked as a helpful tool, or can be suggested when it is not necessary. Finally, the study by Helwig (2002) showed that the use of the read-aloud accommodation in a large-scale study, when used on its own, had a differential impact for students with disabilities. This is relevant to the current study, as it demonstrates that the read-aloud accommodation has been proven to benefit those with disabilities more so than grade level readers.

The Read-Aloud Accommodation in Tests of Reading Comprehension

Since the read-aloud accommodation has been found effective on tests of mathematics, this study sought to explore whether or not it would also benefit students on tests of reading comprehension. However, some have argued that this could invalidate the measure; in reading aloud a reading test, the nature of the exam is being changed from reading comprehension to listening comprehension, which could possibly undermine the purpose of the assessment to the extent that listening comprehension is a different process from reading comprehension (Hollenback, 2002).

A study used factor analysis in order to further explore the effectiveness of the read-aloud accommodation in math and reading for students with disabilities (Bolt & Ysseldyke, 2006). Researchers analyzed data from a large-scale achievement test administered as part of a statewide assessment program, although the particular state in question was not identified in the course of the report. This test offered students with disabilities the read-aloud accommodation. While the precise nature of the read-aloud accommodation as it was given on this test was not described, the researchers defined this accommodation as “having a test assistant read-aloud test directions, items, and item responses.” This data set included information on item-level performance made available for 3rd, 7th and 11th grade students in Reading and Language Arts (R/LA) and 4th, 8th, and 12th grade students in math (MA). The researchers used differentiated item functioning (DIF) on this data set to compare the measurement of two groups of students with disabilities, including those given (RA) and not given (NSD) a read-aloud accommodation, to a reference group of non-accommodated students without disabilities (SS). The purpose was to determine if there would be more DIF identified for the students given the read-aloud accommodation than not, particularly on R/LA tests. This would indicate that the read-aloud accommodation is associated with greater measurement problems than no accommodation, and that this is particularly the case for R/LA as compared to MA. These findings would support the claim that such an accommodation is questionable for tests of reading comprehension.

To carry out this analysis, researchers selected three groups to analyze according to student disability status and accommodations provided, identified by the coding system used on test forms. The first group consisted of 5,000 randomly selected non-accommodated students without disabilities (SS), the second included all non-accommodated students with disabilities (NSD), and the final group included students with disabilities offered the read-aloud

accommodation, often accompanied by another accommodation (RA). Data was separately analyzed by grade and content area. The DIF analysis conducted was an application of item response theory, involving the estimation of item parameters for a focal group, either RA or NSD, and a reference group, SS. Analyses showed that the NSD group had lowest percentage of items showing moderate to large DIF. Meanwhile, larger DIF was identified among the RA group as compared to the SS group. Finally, for all dataset sections, the proportion of items identified with moderate to large DIF was greater in R/LA than in math. Thus, the researchers concluded that the read-aloud accommodation is associated with greater measurement problems than no accommodation, and that this is more pronounced on tests of R/LA than in mathematics. This suggests that this is not a valid accommodation on tests of reading comprehension.

However, various limitations warrant further investigation. Primarily, Bolt et al. (2006) indicate that a different group of students took the test under each form, as this was a preexisting data set and they had no control over group assignment. It is difficult to conclude whether the results gathered can be attributed to the nature of the accommodation or the students taking the exams. Further, the study does not clearly define what is meant by “read-aloud accommodation”; it is not clear how it was administered, how teachers were trained, or how many confounds existed in the presence of other accommodations accompanying it. Finally, the nature of the test given is somewhat limited as well; it is only stated that the tests were multiple choice and approached a variety of topics. However, the state in which it was given, student demographics, precise style of questioning and overall focus is still questionable. Therefore, these findings leave in doubt whether a read-aloud accommodation is valid for students with a disability in tests of reading comprehension.

Other researchers have more loosely examined the nature of the read-aloud accommodation in order to draw conclusions about its possible effectiveness. Some have argued that the inclusion of listening comprehension as part of an accommodation would not invalidate a test of reading comprehension, because listening comprehension and reading comprehension are very highly correlated and, therefore, can be used interchangeably. Joshi, Williams and Wood (1998) explored this further. Within their study, 60 third graders and 60 fifth graders completed the reading and listening comprehension subtests of the Wechsler Individual Achievement Test (Wechsler, 2005) and the Woodcock Reading Mastery Test-Revised (Woodcock et al., 1987). Correlations for measures on these tests ranged from 0.40 to 0.90, with the highest correlations existing in grade 5. It was concluded, therefore, that there is a significant correlation between listening and reading comprehension, and that the strength of the relationship increases over time. Furthermore, reading comprehension tests are not designed to be tests of decoding, but, rather, tests of higher level thinking and processing that could very similarly be assessed through listening comprehension. Therefore, Joshi et al. concluded that it may be feasible to use a read-aloud accommodation on a test of reading comprehension.

However, a study by Elbaum, Arguelles, Campbell and Saleh (2004) demonstrated that, in studying the read-aloud accommodation, it is very important to consider the nature of the accommodation given. In this study, Elbaum et al. worked with 456 students in grades 6 through 10 in an urban, southeastern U.S. school. Of these students, 230 had learning disabilities and the remaining students did not. All students were given one of two forms of a reading test, composed of reading passages and answering comprehension questions. Under one form, students silently took the test under standard conditions as a group within their classrooms. Students were given the second form of this test under accommodated conditions individually in

the school library. They were asked to read the passages aloud to themselves while testing. A 2x2 ANOVA showed that neither the main effect for test accommodation nor the disability x accommodation interaction was statistically significant. In fact, the accommodation only boosted performance for 17% of the LD students. For 20% of these students, the read-aloud accommodation impaired performance.

This research shows the importance of specifying what is meant by the nature of the read-aloud accommodation. In having students with disabilities read passages aloud to themselves, it is quite possible the researchers caused frustration or fatigue in these students, which hindered their performance. In addition, the administration of the accommodated condition may have also affected the results of the study. Each student in the class likely had a different experience reading the passage aloud than the next, depending upon that student's accuracy, fluency, confidence, and support given by the test administrator. Therefore, the nature of this accommodation varied from one participant to the next. In determining whether or not an accommodation is effective, it is important to ascertain that all students receiving the accommodation are being given the condition in as equivalent a manner as possible to remove the possible confounding effects of any environmental variables. Finally, this study does not indicate the precise nature of the reading disability characterized by each student. If students have strong decoding but poor comprehension ability, then they may not benefit as strongly from an accommodation designed to facilitate decoding. In contrast, students who demonstrate good comprehension but poor decoding may benefit more from this accommodation. Therefore, while the results of this study did not support the read-aloud accommodation, more research is necessary to draw any valid conclusions. As noted previously, the current study was specifically designed for students who struggle in decoding to avoid this problem.

A second study by McKeivitt and Elliot (2003) explored the effects of the read-aloud accommodation in which the passages and questions were read-aloud by the researcher. The purpose of this study was to determine whether the read-aloud accommodation or teacher-recommended accommodations would be a more effective testing environment for students with learning disabilities. In order to test this, McKeivitt and Elliot worked with 79 eighth grade students, 40 of whom were selected based upon having an educationally defined disability and special education services mandated on their IEPs. In addition, the teachers of these students participated in the study as well. Teachers filled out an Assessment Accommodations Checklist (AAC) in order to recommend testing accommodations that they felt would best fit the needs of their students. Once the researchers had collected the AACs, they randomly assigned students to one of two conditions. Students in each condition took two alternate forms of the TerraNova Multiple Assessments Reading Test. However, students in condition one took one form with no accommodation, and the second with a teacher-recommended accommodation. In condition two, students took one form with no accommodations and the second form with the read-aloud accommodation. Results of the study showed that, for condition one, neither the main effect of the condition nor the interaction between condition and group were significant. In contrast, for condition two, there was a main effect of the test condition, but no significant interaction of condition x group. All of the students improved under the read-aloud accommodation; therefore, the researchers suggested that this could not be a valid accommodation on a reading test.

Again, this study demonstrated that teacher recommended accommodations are not always indicative of the best modifications available. Further, when the entire test and the passages were read-aloud to the students on this test, it overestimated everyone's performance, for the scores of both disabled and nondisabled students improved.

The same results were obtained in a similar study by Melov (2002). Within this study, Melov worked with 260 students in grades 6-8. Of these, 62 were identified as having learning disabilities in reading according to their IEPs. Melov administered several sections of the Iowa test of Basic Skills to the students, including science, usage and expression, mathematical problem solving and data interpretation and reading. A repeated measures design was used, in which each student took one form of the each test under a standard administration format and a second form within the read-aloud accommodation. Melov hypothesized that the read-aloud accommodation would be an inappropriate choice for a test in reading comprehension, for it would give both students with and without disabilities a boost in performance. To test this, Melov read four level 12 test passages aloud during the read-aloud accommodation. Each item and choice was read once, as well as the passage. Scripts were devised and given to the test takers.

Results of the test showed that the mean scores of students without learning disabilities were higher than for those with learning disabilities. In addition, both groups of students scored higher on all tests under the read-aloud accommodation than under standard administration. An ANOVA showed no significant interaction effect between ability and condition. Therefore, Melov's findings were consistent with those of McKeivitt and Elliot (2003); they concluded that the read-aloud accommodation is not specifically helpful to students with learning disabilities in reading. Rather, it benefits everyone equally, and may not be a valid testing accommodation.

In contrast, the findings of Crawford and Tindal (2004) challenge these conclusions. Despite evidence to the contrary, Crawford and Tindal remained convinced that the read-aloud accommodation could lead to a differential boost for students with disabilities. To test this, they worked with 338 fourth and fifth grade students from Oregon and North Carolina. Of these, 89

received Title I support services. They created two forms of a test, which was drawn from a larger sample of items developed by one of the participating states for test practice. Students were randomly assigned to one of two conditions. Within condition one, students took form A of the test under standard administration and took form B under a video format. Within this form of testing, students were shown a video cassette displaying each item on the test, which played an audio recording of the researcher reading each passage aloud. As students took the test, they were asked to look at the monitor or follow along with the passages in their booklets. Students in condition B took form A under the video format and form B using standard conditions.

Statistical analyses were done on this study, including a two way ANOVA. Results showed that there were no order or form effects, nor was there a main effect of ability, but there was a significant interaction between student classification and administration format. Students receiving Title I special education services benefited the most from this accommodation.

In evaluating this study, one may ponder what it is about the nature of the design that led these results to be different from the others. One possibility may be the nature of the test itself. Another may lie in the fact that each item was graphically displayed for the students to see, and that, through the video format, they were consistently advised to follow along with either the words or the screen. In particular, an emphasis on the items, both in video and booklet form, may have made students more attentive to precisely what was being asked of them, which could have enabled them to more accurately reflect on what was read and answer with greater accuracy.

To add further support to Crawford and Tinder's (2004) findings, a 2012 study (Kim, 2012) also found a differential effect of the read-aloud accommodation, favoring disabled students. This study compared the performance of ten South Korean middle school students with

visual impairments to 10 students without disability matched for grade level; two were in Grade 1, 4 in Grade 2 and 3 in Grade 3. All students took both a standard and a read-aloud form of the 2003 National Assessment of Educational Achievement exam in reading. This was a 20 item test, broken into two forms. Students with visual impairment were also given accommodated test booklets, written in braille or large print. Results, while not statistically significant, showed that students with visual impairments scored higher under the read-aloud accommodation ($M=55.56$) than the standard format ($M=45.56$), while those without disabilities scored lower under the read-aloud accommodation ($M=77$) than the standard accommodation ($M=81$). While the focus of this exam was visual impairments, it also underscored a significant point: read-aloud accommodations do not necessarily benefit all students equally. In this particular example, this accommodation had adverse effects on students without disabilities.

Finally, a recent meta-analysis explored the validity of the read-aloud accommodation in both mathematics and reading comprehension. The results of this study support the benefits of utilizing a read-aloud accommodation in tests of reading comprehension (Li, 2014). Studies in the meta-analysis included those in which the read-aloud accommodation was the sole strategy used for students with disabilities, those with and without statistically significant results, experimental and quasi-experimental designs, and tests of math and reading. Variables identified as closely related to the effects of a read-aloud accommodation included disability status, subject criteria, delivery method, grade level, and extra time. Disability status included learning disabled students and their typically achieving peers, subject areas were limited to mathematics and reading, delivery method involved a human proctor or computerized recording, grade level was broken down into elementary, middle school, and high school levels, and extra time was

only explored as an incidental factor of having questions read-aloud; studies intentionally offering extra time in addition to the read-aloud condition were not included.

Findings of this study showed a statistically significant effect of disability status, with LD students who received accommodations scoring .27 standard deviations higher under the accommodated condition than a standard format, while those without accommodations scored 0.14 standard deviations higher under the accommodated condition. This illustrates that the read-aloud accommodation is particularly effective for students with disabilities, which may make it a valid accommodation. However, another interesting result involved the subject area; the effect size for math was significantly smaller than that for reading. In tests of mathematics, participants collectively scored 0.14 standard deviations higher under accommodated conditions than standard conditions, compared to reading, in which they scored 0.42 standard deviations higher under the accommodated conditions. Therefore, while this accommodation has been widely used in tests of mathematics, this study also leads one to wonder if it could possibly be used in tests of reading.

The findings of this study prompt further exploration. While both students with and without disabilities benefitted from reading comprehension when the subject matter was reading, the effect size was considerably stronger for students with disabilities. In the discussion, there is reflection on how limited understanding of the testing characteristics and settings impacted the results of the study. For instance, Li notes that some students were individually tested, while others were in small groups. Therefore, specifying a particular test setting may shed more light on the effectiveness of the read-aloud accommodation under a set of clearly described procedures. Further, in contrast to the findings of Crawford and Tindal (2004), this study found

a larger effect size for human proctors than computer recordings of the test items. Thus, it may be important to follow up with research using a human proctor.

Research Summary

From the literature review, several major conclusions can be drawn, which lead to the current study. Primarily, in studies of test accommodations for students with reading disabilities, it is important to clarify the precise nature of the reading disability being explored. As indicated by past research, there are different classifications of reading disabilities; some students struggle in decoding, others in linguistic comprehension, and some in both (Gough & Tunmer, 1986; Dreyer & Katz, 1992; Aaron, Joshi & Williams, 1999; Nation & Snowling, 2007; Tilstra, McMaster, Van den Broek, Kendeou, & Piapp, 2009). Thus, it is important to clarify that the current focus is on poor decoders, who show grade level linguistic comprehension, but below grade level decoding skill.

Further, past literature has explored the validity and usefulness of several accommodations for students with reading disabilities. While extended time is commonly used in practice, research has shown that it may not be a helpful accommodation for students with reading disabilities (Elliot et al, 1998; Fletcher et al., 2006). The read-aloud accommodation, in which all directions, test items and multiple choice answers are read aloud, has been proven to be a valid accommodation for students with reading disabilities in the area of mathematics (Schulte et al., 2001, Helwig & Rozek-Tedezko, 2002; Helwing & Tindal, 2003). However, research studies exploring the validity of this accommodation on tests of reading comprehension have had mixed results. Some studies have shown that this has not benefitted students with reading disabilities any more so than their grade level peers (Bolt et al., 2006; Elbaum et al., 2004,

McKevitt & Elliot, 2003; Melov, 2004). Other studies have shown that students with reading disabilities benefitted more from the read-aloud accommodation than their grade level peers (Crawford & Tindal, 2004; Kim, 2012).

The mixed results of the read-aloud accommodation may be attributed to several factors. For one, while research has shown the significance of differentiating between various types of reading disabilities, studies rarely if ever, indicate the precise nature of the reading disabilities characterizing the participants. Often, participants were those with Title I services, and pretests were not given to further classify them. If those tested had other disabilities, or general comprehension problems, they may not have benefitted from a read-aloud accommodation designed to facilitate test taking for poor decoders. In addition, the nature of the read-aloud accommodation varied from study to study. Some researchers had students read aloud passages on their own, others had a test administrator read to the students, and some showed a video monitor displaying each item being read aloud. Finally, no studies reported using the read-aloud accommodation for only portions of the test. On tests of mathematics, students were read questions, but were expected to show their own work and approach each task independently. The effectiveness of this approach in math suggests that it may work also on tests of reading comprehension. If students with reading disabilities are read question items, and are then expected to explore the reading passages on their own, perhaps they can more accurately show their comprehension of the passage without having their abilities overestimated.

Therefore, taking account of the strengths, limitations, and questions prompted by past research, a field study and dissertation were designed and conducted to further explore the partial read-aloud accommodation for tests of reading comprehension.

Chapter 3 – Field Study, Rationale and Hypotheses

A successful field study provided support for the use of a partial read-aloud accommodation for students with reading disabilities taking tests of reading comprehension. This 2012 field study compared the performance of good and poor third grade readers on the Gates-MacGinitie Reading Comprehension test (MacGinitie et al, 2002) under two conditions – standard administration and a partial read-aloud with pacing accommodation. Under the latter, only question items, multiple choice answers and proper nouns were read aloud, and students were asked to read the test passages independently. Results revealed a significant interaction between reading ability and test condition. This study is described in greater depth in the section below.

Variables, Hypotheses and Research Questions. The field research study considered three independent variables. One of these was test condition, with the partial read-aloud with pacing condition (PRAP) compared to a standard condition in a repeated measures design. Under the partial read-aloud with pacing accommodation, the test administrator began by showing students index cards displaying any proper nouns in the text given that these are generally quite difficult for students with poor decoding skills to pronounce. The administrator then said the word, pointed to it, and asked the students to practice saying it so that they were enabled to decode it when they read the text. The students were then asked to read the test passage silently and independently and to give the test administrator a signal by placing their pencils along the edges of their desks to indicate when they had finished. After all students completed the passage, the administrator read aloud each test item and multiple choices to the students. Once each question and choices were read, students selected their answers and gave the administrator the signal to read the next item. In the standard condition, the reading

comprehension test was administered in a standard format. Students were asked to read all directions, passages and questions independently without any assistance. The same students completed tests in both conditions.

The second independent variable was reading ability. This study was designed to determine if students would respond to the accommodations differently based upon whether or not they were classified as having a reading disability. Therefore, participants were either classified as poor decoders or average readers. For the purpose of this study, students labeled poor decoders were identified as those who demonstrated average listening comprehension ability, but below average decoding skill. This was assessed with teacher recommendations and pretests. The test accommodation was regarded as effective if poor decoders showed better performance with than without the accommodation whereas good readers showed little difference in performance.

A third independent variable was test form. To control for order effects, students in each classroom completed each condition with one of two equivalent test forms. Students in one group were given one form of the reading test (Form S or T) given as a partial read-aloud accommodation, while the other form was taken using the standard directions. Students in the second group used the opposite forms for each condition. The dependent variable in this study was the score attained on each reading assessment.

It was expected that the partial read-aloud with pacing accommodation would prove to be a more valid accommodation compared to the full read-aloud accommodation investigated in past research. Therefore, the research questions addressed in this study were:

- Does the partial read-aloud with pacing accommodation improve performance on a reading comprehension test more than no accommodation for poor decoders?

- Is there an interaction between reading ability and accommodation such that poor decoders show significantly greater improvement in reading comprehension with than without the partial read-aloud with pacing accommodation, whereas average readers show little improvement in reading comprehension with the accommodation?

The hypothesis tested was that the partial read-aloud with pacing accommodation would improve the reading comprehension scores of poor decoders more than average readers compared to no accommodation.

Methods

This study was conducted within a public school in Queens, New York. According to the school's 2009-2010 School Demographics and Accountability report, which was the most current reflection of student demographics at the time of the study, this was a predominantly middle class school (with a poverty rate listed as 33.7%) and the majority of students were Asian American (66.1%), followed by Caucasian (21.5%), Hispanic (8.6%) and African American (1.6%). Prior to the study, the researcher obtained IRB approval as well as written consent from school administrators to show that they understood the nature of the study and were willing to participate. Once granted IRB approval, the researcher asked three participating third grade teachers to identify students they would classify as "average" or "poor" readers on the basis of their classroom assessments. This was done as a recruiting measure, but was not a means of classifying students. The children recommended by their teachers were given permission forms for their parents to fill out and return. All students who returned permission slips were also asked to provide oral assent. Forms were distributed to 45 children. Once the sample of

participants was formally established, the researcher worked one-on-one with students to administer pretests in decoding and listening comprehension.

Decoding was measured using the Letter and Word Identification and Word Attack subtests of the Woodcock Johnson III Reading Test (Woodcock, McGrew & Mather, 2001). The Letter and Word Identification subtest asked students to read aloud a series of letters and words. The Word Attack subtest asked students to do the same with a series of nonsense words to ascertain students' decoding skill.

Listening comprehension was measured using two subtests of the Clinical Evaluation of Language Functions 4th Edition Test (CELF-4) (Semel, Wig, & Secord, 2003). These subtests included the Concepts and Directions and Understanding Spoken Paragraphs subtests. The Concepts and Directions Subtest asked students to interpret spoken directions of increasing length and complexity by responding to a verbal request. Students were shown a line of pictures and were directed as to which picture they should point. On some items this involved asking them to interpret ordinal words ("point to the third car") or to differentiate direction or size ("to the right of" "the little car"). The Understanding Spoken Paragraphs subtest asked students to listen to short passages and answer comprehension questions.

Once pretests were completed, the researcher examined performance to determine which students would be classified as poor decoders and average readers. The CELF-4 included standardized norms for each age group, including a mean and standard deviation. All of the students who scored more than half a standard deviation below the indicated mean were removed from the study to eliminate the possibility that their difficulty was in comprehension itself rather than decoding. Scores on the WJIII were used to determine which of the remaining students were classified as poor decoders. Those attaining scores at least one standard deviation below

the established grade level norms fit the criteria, for they showed stronger listening comprehension than decoding skill.

Once each student's reading ability level was determined, a list was devised for the poor decoders and average readers. It was determined how many students would be assigned to group one and group two. A list of the numbers 1 and 2 were repeated and randomly ordered. The numbers corresponded to the predetermined number of participants to be assigned to groups 1 and 2 (1, 2, 2, 1, 1, 1, 2...). The list was placed alongside a list of participants randomly ordered. This enabled the researcher to randomly assign students to the two testing groups, each of which would have a roughly equal number of average readers and poor decoders.

All students took both Form S and Form T of the Level 3 Gates-MacGinitie Fourth Edition reading comprehension subtest. This is a test of reading comprehension using a series of short reading passages accompanied by a total of 48 multiple choice questions. Students in the first group took Form S of this exam by following the standard administration directions specified in the test manual. All passages and questions were read and answered independently by the students. The same group took Form T by following the partial read-aloud with pacing accommodation procedure. Under this accommodation, the researcher showed students index cards displaying any proper nouns appearing in a passage. She modeled how these nouns were pronounced. Students then read the passage independently. Following this, the researcher read each question and its corresponding answer choices students before they answered them. Upon completing a passage or an item, students were asked to signal when they were ready to move forward to the next question or passage. This involved placing their pencils on the edges of their desks.

Students in the second group followed the same procedures, except that they completed Form T under standard conditions first followed by the partial read-aloud condition of Form S. Therefore, in each group, the standard accommodation preceded the accommodated condition, but the forms were switched. After each student's final session, they were offered a small prize for their participation in the study.

After the completion of all testing, raw scores were calculated as well as difference scores between the partial read-aloud and the standard administrations. These were used to run a 2x2 (ability x condition) ANOVA with an alpha level at .05 in order to determine main effects of each independent variable as well as interaction effects.

Results

Of the 45 children who were recommended by teachers and received permission slips, 25 responded with parental permission. From that sample, two did not give assent and one was pulled from the study early. The 22 children who participated for the duration of the study included 9 boys and 13 girls. The majority of students were Asian American (15), including students from Korean, Chinese and Indian backgrounds. Five children were Caucasian, one was Hispanic, and one was African American. All but one child were fluent in more than one language. Therefore, the sample selected was a fair reflection of the school's demographics. Characteristics are reported in Table 1.

Results of the pretests were used to group students into poor decoders and average readers. All of the students scored no lower than half a standard deviation below the mean for their age group on the CELF-4 tests, so none had to be eliminated from the study since all demonstrated grade level listening comprehension. Then, scores on the WJIII were analyzed to determine a sample mean and standard deviation, which are shown in Table 1. All those

students who scored at or greater than one standard deviation below the mean calculated on the group were classified as poor decoders, while those who scored higher than one standard deviation below the mean were classified as average readers. Therefore, any students who identified fewer than 9 words on the Word Attack task or 42 on Word identification were classified as poor decoders. Of the students tested, 7 were identified as poor decoders after pretesting, while 15 were average readers. More specific demographics of each group are depicted in Table 1.

Table 1.
Characteristics of Participants in the Pilot Study

Characteristics	Poor Readers	Average Readers
Gender (M/F)	3 / 4	6 / 9
Ethnicity		
Asian	3	12
Caucasian	2	3
African American	1	0
Hispanic	1	0
Pretest [Mean (<i>Range</i>)]		
Letter/Word Identification	42.71 (42-48)	48.07 (43-53)
Word Attack	9.57 (1-20)	20.00 (12-29)
Concepts/Spoken Directions	47.86 (44-52)	51.33 (46-54)
Understanding Spoken Paragraphs	12.86 (9-15)	12.80 (9-15)

Descriptive statistics in Table 2 reveal that the mean score of average readers on the reading comprehension test did not improve significantly under the partial read-aloud accommodation. Average readers attained a mean of 39.00 questions correct under the standard administration and 39.53 under the partial read-aloud accommodation. In contrast, poor decoders showed a far greater increase. These participants attained a mean score of 28.42 under the standard condition and 37.86 under the partial read-aloud accommodation. Moreover, the

superior performance of average readers over poor decoders was substantially narrower under the partial read-aloud with pacing accommodation than under the standard condition.

Table 2

Mean Scores of Average Readers, Poor Readers, Those Assigned to Groups 1 and 2 on the Gates-MacGinitie Reading Comprehension Test in the Field Study

Test	<u>Average Readers</u> <i>M (SD)</i>	<u>Poor Readers</u> <i>M (SD)</i>	<u>Group 1</u> <i>M (SD)</i>	<u>Group 2</u> <i>M(SD)</i>
Standard Test Administration				
Raw (Max 48)	39.00 (6.38)	28.42 (8.14)	35.91 (6.07)	35.36 (10.63)
Grade Equivalent	5.1	3.2	4.4	4.2
Read-Aloud Accommodation				
Raw (Max 48)	39.53 (5.84)	37.86 (4.56)	38.64 (6.56)	39.36 (4.27)
Grade Equivalent	5.1	4.5	4.8	5.2

Note. There were 15 average readers and 7 poor readers. There were 10 students in Group 1 and 12 students in Group 2.

Mean performance was compared statistically in a repeated measures 2x2 ANOVA shown in Table 3. A significant main effect was detected for reading ability, with average readers scoring significantly higher overall than poor decoders. There was a statistically significant main effect of reading condition as well as a statistically significant interaction between test condition and reading ability. As displayed in Figure 1, the mean score of poor readers increased substantially from the standard condition ($M = 29.43$) to the partial read-aloud with pacing accommodation ($M = 37.86$) whereas the mean score of average readers increased very little from the standard condition ($M = 39.00$) to the partial read-aloud with pacing accommodation ($M = 39.53$). These results show that the partial read-aloud with pacing accommodation did not benefit everyone equally.

Table 3

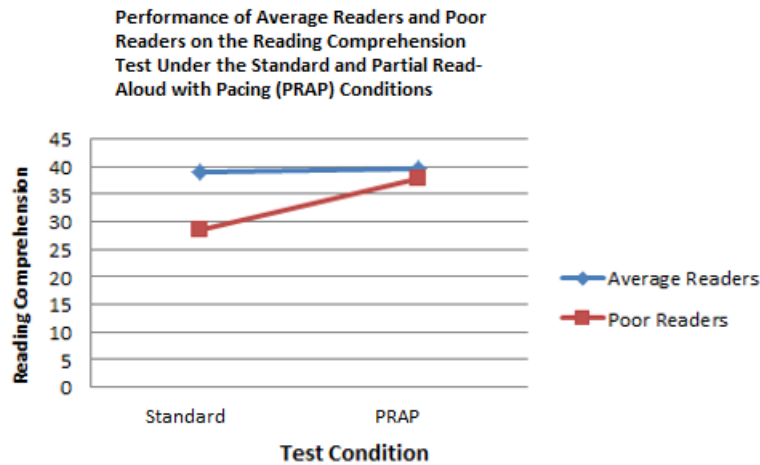
Analysis of Variance for Reading Treatment Condition and Reader Ability Level on Students' Reading Comprehension in the Field Study

Source	<i>df</i>	<i>MSE</i>	<i>F</i>	<i>p</i>	<i>Partial Eta Squared</i>
Reader Ability (RA)	1	357.97	5.10	.035*	0.20
Error	20	70.13			
Reading Condition (RC)	1	236.82	28.24	.00**	0.59
RC x RA	1	188.82	22.52	.00**	0.53
Error	20	8.39			

* $p < .05$; ** $p < .01$

Figure 1

Mean Gates-MacGinitie Reading Comprehension Scores Across Test Conditions for Poor Readers and Average Readers on the Field Study (Maximum Score = 48)



To examine whether the particular form of the Gates-MacGinitie test (S or T) yielded similar scores, mean scores of average readers in Groups 1 and 2 were examined. Results revealed a greater difference in Group 1 than in Group 2. The mean score of Group 1 was 35.89 under the standard condition of Form S and 37.56 under partial read-aloud with pacing

accommodation using Form T. Group 2 scored on average 43.67 under the standard condition using Form T and 42.50 under the partial read-aloud with pacing accommodation with Form S. Therefore, the differences between the accommodation and standard test condition means for each group are small and similar, although Group 1 scored consistently lower on both forms than Group 2.

Conclusions and Questions for Further Study

The data resulting from this study provided support for the hypothesis that the partial read-aloud with pacing accommodation would boost the scores of poor decoders on a reading comprehension test compared to standard administration of the test but it would not boost scores of average readers. This contrasts with previous studies reporting that all students improved equally under a full read-aloud accommodation. In the current study, the significant interaction showed that the partial read-aloud with pacing accommodation did not favor everyone similarly. The students who benefitted most were the poor decoders.

However, there are some limitations to the study that need to be addressed. The sample was small and was drawn from a predominately Asian community, so it may not be representative of a wider range of students. In addition, while giving the partial read-aloud with pacing accommodation, the researcher observed that some of the average readers became very frustrated and bored with the task of sitting and waiting for their slower-reading peers to complete the questions. In the current study, therefore, students who were classified as poor decoders were tested separately from those classified as average readers. Finally, the school population studied was high functioning, with many students scoring above grade level in all of the pretests. By the end of the study, all third graders, even those classified as poor decoders,

scored on or above the third grade level. Finally, while there were two forms, a small sample of students and limited scheduling prevented the order of test condition to be counterbalanced.

Group one completed the standard administration of Form S first, followed by the partial read-aloud accommodation using Form T. Group two completed the standard administration of Form T first, followed by the partial read-aloud condition of Form S. This may lead one to question whether or not exposure to the standard accommodation first facilitated improvement in the partial read-aloud with pacing condition, which was given second. However, since superior performance occurred only among poor decoders and not average readers, general test practice does not seem to explain improved test performance. In the dissertation study, to eliminate this possibility, independent groups assigned to each accommodation were used rather than repeated measures.

In considering which properties of the partial read-aloud with pacing accommodation enabled the poor decoders to improve their reading comprehension scores, the favored explanation is that it alleviated the difficulty of having to decode the test questions and thus gave poor decoders a greater chance of showing their understanding of the passages they had read. However, there are other possible factors to consider. For one, it may have been the fact that the researcher regulated students' speed in completing the test in the accommodation condition but not in the standard condition. The researcher observed that when students' progress through the test was not regulated. The poorest readers were often the first ones to complete the test, suggesting that they may have been speeding through the test without fully processing the information. Under partial accommodation, however, they were not allowed to rush because progress through the items was controlled by the researcher. This difference may have caused

depressed scores among poor readers in the standard condition and hence may explain the benefit of the partial accommodation.

To address this possibility in the current study, a third experimental condition was included in which students were paced through the test. That is, they were told when to read each part and when to answer each question but none of the parts was read aloud to them. This insured that students completed each question on the test, and it duplicated the pacing used in the partial accommodation condition. If pacing alone explained the advantage of the partial read-aloud with pacing accommodation, then poor decoders who were paced were expected to perform as well as poor decoders given partial accommodation and to outperform poor decoders in the standard testing condition.

Chapter 4 - The Current Study

Variables, Hypotheses and Research Questions

In the current study, effects of two independent variables on reading comprehension were explored. One was test condition, with the partial read-aloud with pacing accommodation (PRAP) compared to the pacing only (PO) accommodation and standard conditions. In the partial read-aloud with pacing condition, the test administrator began by showing students index cards with any proper nouns in the story, as these are generally quite difficult for poor decoders to read. The administrator then said the word, pointed to it, and asked students to practice saying it so that the word would become familiar before it was encountered in the text. The students were then asked to read the passage silently and independently and give the test administrator a signal by placing their pencils along the edges of their desks to indicate when they had finished. After all students completed the passage, the administrator read aloud each test item and multiple choices to the students. Once each question and choices were read, students selected their answers and gave the administrator the signal to read the next item.

In the pacing only accommodation, no vocabulary words were presented or questions read. However, the administrator told students when to read the passage and each question, when to answer, and when to move on to the next question, which is what was done in the partial read-aloud with pacing accommodation. Students were told to give a signal by placing their pencils on the edge of the desk each time they completed a passage or found the answer to an item.

In the standard condition, the reading comprehension test was administered in its typical format. Students were asked to read all directions, passages and questions independently without any assistance.

The second independent variable was reading ability. The question of interest was whether students would respond to the accommodations differently based upon whether or not they had a reading disability. Therefore, participants were classified as either poor decoders or average readers. For the purpose of this study, students labeled poor decoders were those who demonstrated average listening comprehension ability but below average decoding skill. This was assessed through pretests.

It was hypothesized that the partial read-aloud with pacing condition would prove to be a valid accommodation. Prior research has indicated that reading entire passages and corresponding test items aloud to students is not a valid accommodation because all students improve in their performance. However, research has not been done on reading aloud the test items alone. Students with poor decoding skills, who may become frustrated when reading silently to themselves for extended periods of time, may be better able to show their ability to comprehend passages independently when questions and proper nouns present in the passage are read aloud. In this manner, the students are still responsible for reading and interpreting each test passage independently, but they do not have to devote extra time and energy to decoding each question and its multiple choice answers as well.

The research questions addressed in the current study were as follows:

1. Does the partial read-aloud with pacing condition improve third graders' performance on a reading comprehension test more than the pacing only condition and more than the standard condition?
2. Do poor decoders comprehend text significantly better in a partial read-aloud with pacing condition than in a pacing only or standard condition, whereas average readers show little difference in reading comprehension across these conditions?

The hypotheses to be tested were:

1. The partial read-aloud with pacing accommodation will boost the reading comprehension scores of poor decoders compared to the pacing only and standard conditions whereas the partial read-aloud condition will provide little improvement in reading comprehension compared to the pacing only and standard conditions for average readers. This will show that the benefit of the partial read-aloud condition is limited to poorer decoders and hence constitutes a valid type of test accommodation.
2. The partial read-aloud with pacing accommodation will boost the reading comprehension scores of poorer decoders more than the pacing condition, showing that the benefit of partial read-aloud with pacing accommodation is not simply a result of guiding students through the test.

Chapter 5 - Methods

Participants

All participants in this study were third grade students attending one of three sites, either one of two public schools in Queens, New York, or a summer camp accessible to New York City Public school students. Each site is described below.

School 1: This was a public school in Queens, New York. According to the school's online profile, this is a predominately middle class school, with 22% of students qualifying for free or reduced-price lunch. The majority of students are Asian American (66%), predominantly composed of students from Chinese and Korean backgrounds, followed by Caucasian (20%), Hispanic (11%), and African American (3%). This school is typically high performing, and ranked "Good" or "Excellent" in most components of its 2013-2014 Quality Snapshot. In 2014, 63% of the students met grade level standards on the State English Test, compared to a city average of 30%. Of the student population, 14% are identified as students with special needs.

Two cohorts of students from this school participated in the study. The first cohort was in third grade during the 2013-2014 school year. This cohort included 32 students. The second cohort was in third grade during the 2014-2015 school year, and also consisted of 32 students.

School 2: This was a public school in Queens, New York. According to the school's 2010-2011 School Demographics and Accountability report, this is a predominantly middle class school, with 36% of students qualifying for free or reduced lunch. The majority of students are Asian American (69%), primarily composed of students from Indian and Middle Eastern backgrounds, followed by Hispanic (12%), White (11%) and African American (5%). The administrator indicated that both types of classrooms included many struggling readers. Like School 1, this school appeared to be high performing, based on its 2013-2014 Quality Snapshot

that reported ratings of “good” or “excellent” in nearly all components. 60% of students met grade level standards on their State English Language Arts exams, compared to a city average of 30%. Students from an after school program participated in the current study.

Summer Camp Program: This was a privately funded program housed in a New York City high school. Students attending are drawn from a variety of public schools in a Queens, NY district. However, the camp director indicated that most participants come from the same public school in Queens, NY. According to the school’s 2013-2014 Quality Guide, this is a predominately middle class school, with 33% of students qualifying for free or reduced lunch. The majority of the students are Asian American (45%), primarily composed of students from Indian and Middle Eastern backgrounds, followed by White (33%), Hispanic (18%) and Black (3%). This school appeared to be high performing, as it received ratings of “good” or “excellent” in nearly all components. 64% of students met grade level standards on their State English Language Arts exams, compared to a city average of 30%.

For the purpose of this study, each group is identified as a different cohort. Cohorts are listed in the order in which they participated in the study. Cohort 1 included the first group of students from School 1, Cohort 2 included participants from school 2, Cohort 3 consisted of the summer camp participants, and Cohort 4 was the second group of students from School 2. Collectively, summed across cohorts, the sample included 35.4% (29) boys and 64.6% (53) girls. The majority of students were Asian American (59.8%, N = 49), followed by Caucasian (24.4% , N=20), Hispanic (9.8% , N =8), and African American (6.1%, N =5). The mean age was 8 years, 9 months.

Once granted IRB approval, the researcher distributed information about the study to all principals and camp directors participating (See Appendix A). Then, permission forms were

distributed to all third graders in each site (See Appendix B). During the 2013-2014 school year, 102 permission forms were given in School 1, 50 in School 2, and 40 in the summer camp program. 32 signed permission forms were returned in School 1, 6 in School 2, and 11 in the summer camp program. In order to seek out more participants, an additional 100 permission forms were distributed in School 1 during the 2014-2015 school year, and 40 were returned. Of those, 34 participated in the study and 8 did not. The remaining 6 did not remain in the study due to lack of assent, parental request, or moving out of the school. After pretests, an additional 2 students were classified as poor decoders and eliminated from the study, leaving 32 students in this cohort. The mean age of all participants in the study was approximately 8 years, 9 months.

Pretests.

The following pretests were administered individually to identify average readers and poor decoders. Students who demonstrated average listening comprehension but below-average decoding skills were classified as poor decoders. Students who scored on grade level in both listening comprehension and decoding tests were classified as average readers.

Clinical Evaluation of Language Functions, Fourth Edition (CELF-4)

Understanding Concepts and Spoken Directions Subtest (Semel et al., 2003). This was used as a measure of listening comprehension. This task requires students to interpret spoken directions of increasing length and complexity by responding to a series of verbal requests. Students were shown a line of pictures and directed as to which pictures they should point. This task requires students to follow directions in the proper order, interpret ordinal words (“point to the third car”) or show differentiation in direction or size (“to the right of the little car”). Students were given three practice trials before completing the task. All directions followed accurately without experimenter prompting were marked as correct, and a raw score was

determined by adding together the number of correct trials. Results were compared to reported grade level means supplied by the test. Those who scored lower than one standard deviation below the standardized mean for ages 8-9 were eliminated from the study, as they showed evidence of difficulties in overall comprehension. Those who scored higher than half a standard deviation below the standardized mean were considered to demonstrate grade level comprehension skills and were included in the study.

CELF-4 Understanding Spoken Paragraphs Subtest (Semel et al., 2003). This subtest served as an additional measure of listening comprehension. The experimenter read aloud three short passages to each participant and asked a series of 5 open-ended comprehension questions for each passage. Two short practice trials were attempted before the participants carried out the task. A point was given for each correct answer, and a raw score determined the number of correct responses out of fifteen possible points. Again, those who scored below 1 standard deviation of the mean reported for the measure by children ages 8-9 were removed from the study, including 2 students in Cohort 4.

Woodcock Reading-Mastery Test-Revised (WRMT-R) Word Reading Subtest (Woodcock, 1987). This test was used to assess each participant's knowledge of sight words and decoding ability. Participants were measured on their ability to read a list of words ordered by increasing difficulty. All words read promptly and accurately without experimenter prompting were recorded as correct. The split-half reliability coefficient reported in the manual is .97.

WRMT-R Word Attack Subtest (Woodcock, 1987). This served as an additional measure of decoding, in which students were assessed on their ability to decode a set of pseudowords. Participants were tested on their ability to read nonsense words using their understanding of decoding and spelling patterns. All words read promptly and accurately

without experimenter prompting were recorded as correct. Students who were able to identify 14 or fewer words scored one standard deviation below the mean for their grade level, and, therefore, were identified as poor readers. The split-half reliability coefficient reported in the manual is .91.

PPVT-4 - The Peabody Picture Vocabulary Test (Dunn & Dunn, 2007). This pretest was given to assess the possible influence of students' vocabulary knowledge on the effectiveness of the test accommodations provided. In this test, students were presented with a series of four pictures, and were asked to identify the picture that best illustrated the meaning of a word. This was particularly significant to explore since such a large proportion of the sample identified English as their second language. The question of interest was whether participants' knowledge of English vocabulary would influence their performance on the Gates-MacGinitie reading comprehension test in each test accommodation condition. The split-half reliability coefficient reported in the manual is .93.

Procedures for Selecting Average and Poor Decoders

Once pretests were completed, the researcher used students' performance to form the groups of poor decoders and average readers. Poor decoders were classified as those who showed grade level listening comprehension in response to directions or a story, but struggled with decoding or sight word reading. All of the students who scored lower than one standard deviation below the third-grade mean equivalent on either of the two CELF-4 subtests were removed from the study to eliminate those with possible comprehension difficulties. According to normative means reported in the CELF-4 administration booklet, the expected mean of the Concepts and Following Directions subtest for third grade was 41.9, with a standard deviation of 8.5. The expected mean of the Understanding Spoken Paragraphs subtest was 11.4, with a

standard deviation of 2.9. Therefore, all those students who either scored below a 33 on the Concepts and Following Directions subtest, or below a 9 on the Understanding Spoken Paragraphs subtests, were identified as poor comprehenders and were dropped from the study. All those who scored above those normative benchmarks were classified as average comprehenders, and were then tested on their word reading skills. Those who remained were considered to have at least average listening comprehension ability.

Scores on the WRMT-R word identification and word attack subtests were used to determine which of the remaining students qualified as poor decoders. Those who attained scores at or below one standard deviation lower than the third-grade-equivalent mean on either of the two WRMT-R subtests were classified as poor decoders. This was determined using normative means and standard deviations reported in W scale units, as shown in Table 5.2 of the WRMR-R User Guide. Since this third grade sample consisted of students from a variety of specific ages, it seemed most appropriate to use the overall normative means established for the grade as a reference point rather than differentiate each participant by age in years and months. According to the normative table, third graders scored a mean W of 495.2 on the word attack, with a standard deviation of 15.2. Therefore, one standard deviation below the W mean score would be a W score of 480. The raw score that corresponded to this W score was 16 for Form S. The mean W score on the Word Identification task was 482.3 for third graders, with a standard deviation of 23.3. Thus, those who fell one standard deviation below the mean would have a W score of 459, which, according to normative scores, corresponds to a raw score of 49. Therefore, students who had either a raw score of 15 or lower on the Word Attack task or 48 or lower on the Word Identification task, while also scoring above the benchmark scores in the CELF-4 measures, were classified as poor decoders. The remaining students were classified as

average readers. In sum, poor decoders showed average listening comprehension but poor decoding skill while average readers showed average decoding and listening comprehension skills, as identified through normative means established on each of the measures used.

Reading Comprehension Test

The Gates-MacGinitie Reading Comprehension Tests, 4th Edition, Reading Comprehension Subtest, Form S – Grade 3 (MacGinitie, MacGinitie, Maria & Dreyer, 2000).

This test was given to all participants in order to assess their reading comprehension of short passages. This test includes a series of narrative and informational passages for students to read, each passage followed by questions, 48 in total for the subtest. According to the test manual, students are given a total of 35 minutes to complete the subtest. Questions require students to demonstrate both literal recall and inferential thinking. Two alternate forms of the test, S and T, are available. One step of the partial accommodation condition required the examiner to show and pronounce proper nouns to the students. Since Form T included more proper nouns than Form S, the experimenter elected to use only Form S in order to limit the impact of this step as a cause of any differences between testing conditions on students' reading comprehension. Another reason for using Form S rather than Form T was that mean performance of average readers in the pilot study was lower on Form S than Form T and hence further from the maximum score possible. The Kuder-Richardson 20 reliability coefficient reported in the manual is .93.

After obtaining permission from the publishers to modify the test format and administration of the Gates-MacGinitie reading comprehension test, the principal investigator created an adapted presentation of the test to be used in the partial read-aloud with pacing

accommodation and pacing only conditions. This included printing each passage and each test item on a separate page. In doing so, the principal investigator sought to control students' progress through the test and prevent students from moving ahead or skipping parts of the test.

Experimental Treatments

The performance resulting from three different reading comprehension test conditions, the partial read-aloud with pacing accommodation, the pacing only accommodation, and standard condition were compared in poor readers and average readers. Under each of these conditions, students were tested in small groups ranging in size from 1 to 15. For the partial read-aloud and pacing conditions, poor readers and average readers were tested separately.

The Partial Read-Aloud with Pacing Condition (PRAP) - Under this condition, each group of average and poor decoders was guided through the test by the experimenter, who followed the standardized set of directions with some modifications (See Appendix C). An adapted format of each booklet was used. Prior to reading each passage, the experimenter showed students index cards displaying any proper nouns in the story. Each word was read aloud, and students were asked to repeat it. The experimenter indicated that the readers would see these words in the text. If there were no proper nouns within a particular passage, this step was eliminated. The experimenter prompted the children to begin reading each story and to give a signal when they were finished (placing their pencils at the edges of their desks). Participants were told not to continue until everyone was done. Then the experimenter read each question and all its multiple choice answers to the students. She prompted them to follow along, to find the correct answer, and to signal when they were ready for the next item. The test booklet included 11 passages, 5 of which included proper nouns for a total of 10 proper nouns introduced

during the exam. Adapted directions are included in the appendix, and identify each of the nouns that were introduced.

The Pacing Only Accommodation (PO) - Under this condition, the experimenter guided groups of average readers and poor readers through Form S of the Gates-MacGinitie Reading Comprehension Test in a way similar to that conducted in the partial read-aloud with pacing condition using a scripted set of directions (See Appendix D). However, the experimenter instructed students to follow timing prompts through pages of the booklet and she did not read any items aloud. The experimenter told the students when to begin each passage, to signal when they were done by placing a pencil on the edge of their desks, and when to turn to the next page. The experimenter also had students complete one test question at a time, to signal when they were done using the same pencil method, and when to turn the page to the next test item. Additionally, poor decoders were offered the accommodated version of the test booklet presented to all students in the partial read-aloud with pacing condition. This was done to help enhance student focus, as it was noted in the pilot study that poor decoders seemed distracted during the standard form of the reading comprehension test, while average readers were not. For these students, it seemed that an overwhelming amount of text on a single page had worsened their difficulty in focusing. In offering poor decoders an adapted booklet, the principal investigator sought to determine whether or not removing distracting factors, like multiple items on a page, could enable them to best respond to the pacing prompts provided.

Standard Condition. Under this condition, participants took the test in its standard format free of any accommodation procedures. All students read the stories and answered 48 items within a 35-minute time constraint.

Procedure

Before beginning any work with the children, the investigator obtained IRB approval, as well as written permission from all administrators, parents and teachers involved in the study. Once the schools were recruited, permission forms were distributed to all third grade students in each school or cohort. Students who returned permission slips also gave assent before they began the pretests (See Appendix E).

All tasks were administered from November of 2013 through January of 2015. Students from Cohorts 1 and Cohort 2 worked with the principal investigator from January 2014 through June 2014. Data was collected from the summer camp in July and August of 2014. The investigator worked with Cohort 4 from September 2014 through January of 2015. Students were given pretests individually at times chosen by their classroom teachers or camp counselors. To prevent students in Cohorts 1 and 4 from missing instructional time, the investigator met with these students during their lunch period. Students in Cohorts 2 and 3 met during recreational time in their after school program or summer camp. Additionally, the investigator collected reports from each student regarding the nature of the languages spoken at home. Students were asked whether or not they spoke a language other than English, and if they spoke their native language or English more frequently at home. The investigator obtained this information in order to assess the impact of linguistic background on the effectiveness of the accommodation.

On the basis of the pretest scores, students were classified as poor decoders or average readers and were randomly assigned to one of three treatment groups: the partial read-aloud with pacing condition, the pacing only condition, or the standard condition. Six testing sessions took place within each cohort, since poor decoders and average readers were tested separately in the

two accommodation conditions. Because students' progress through the test was paced in both accommodation conditions, the sessions necessarily lasted longer than the standard test session.

Data Analysis

Once students completed all of the pretests and the reading comprehension test, descriptive statistics were calculated to compare performance of the treatment groups. Several ANOVAs were conducted with reading ability and accommodation condition as the independent variables and test performance as the dependent variables. Hypotheses were tested at an alpha level of $p < .05$ to assess whether main effects and interactions were statistically significant. ANOVAs on the pretests were intended to determine whether or not there were significant main effects of reader ability or treatment condition distinguishing the groups. It was expected that within each reader ability level, the three groups would perform similarly on the pretests. Further, an ANOVA was run to explore whether the linguistic background of students impacted outcomes. Finally, correlation coefficients were calculated between variables to examine the strength of relationships between pretest measures and the Gates-MacGinitie reading comprehension test.

Chapter 6 - Results

Characteristics of Participants

Demographics of each cohort are shown Table 4.

Table 4

Characteristics and Numbers of Participants Contributed by Each Cohort with Percentages in Parentheses Indicating the Preponderance of Characteristics within Each Cohort

Characteristics	Cohort 1	Cohort 2	Cohort 3	Cohort 4	Total
Gender					
Male	9	2	8	10	29
Female	23	4	3	23	53
Ethnicity					
Asian	19	5	3	23	50
Caucasian	10	1	2	7	20
African American	1	0	4	1	6
Hispanic	2	0	2	2	6
Reading Classification					
Average Readers	22	3	6	23	54
Poor Readers	10	3	4	11	28
Linguistic Background					
Bilingual (Native Lang.) ^a	4	4	0	0	8
Bilingual (English) ^b	3	0	4	12	19
Bilingual - (Both) ^c	18	0	0	14	32
Monolingual	7	2	7	7	23
% Reduced Price Lunch	22%	36%	33%	22%	

^a Bilingual students who are native language prominent at home.

^b Bilingual students who are English prominent at home.

^c Bilingual students who speak both their native language and English at home.

In Cohorts 1, 2, and 4, the majority of students were Asian American and female, while, in Cohort 3, the majority of students were African American and male. In each cohort, the majority of students were bilingual, although the reported nature of each student's language use at home varied. In Cohorts 1 and 4, the majority of students were bilingual, and reported speaking both their native language and English equally at home. In Cohort 2, the majority of students reported speaking their native language more than English at home. Cohort 3 was predominately monolingual. Specific demographics of each cohort can be seen in Table 4.

Table 5

Characteristics of Groups of Poor Readers and Average Readers

<u>Characteristics</u>	<u>Poor Readers</u>	<u>Average Readers</u>
Gender (M/F)	10/ 18	19 / 35
Ethnicity		
Asian	13	36
Caucasian	9	11
African American	3	2
Hispanic	3	5
Pretest [Raw Score Mean (<i>Standard Deviation</i>)]		
PPVT	139.14 (13.22)	153.22 (7.60)
<i>Age Equivalent</i>	8:9	10:2
Word Identification Form H	57.14 (10.52)	67.69 (5.18)
<i>Grade Equivalent</i>	3.40	4.40
Word Attack	12.75 (7.43)	24.87 (5.99)
<i>Grade Equivalent</i>	1.70	2.90
CELF Directions	48.64 (3.50)	50.20 (2.50)
<i>Age Equivalent</i>	>8:11	>8:11
CELF Listening Comprehension	13.07 (0.94)	13.96 (1.57)

Note: Students in the study had an average age of approximately 8 years, 9 months. In the PPVT, their estimated score would be 136. CELF Directions estimated score would be 42. Age Equivalents were not provided in the CELF Test Manual for the Understanding Spoken Paragraphs (Listening Comprehension) subtest.

Demographics and pretest means characterizing the average readers and poor decoders are displayed in Table 5. Of the 82 students, 54 were classified as average readers on the basis of pretests, while 28 were poor decoders. Of the 28 students, 26 were classified as poor decoders on the basis of their Word Attack subtest scores alone, while the remaining 2 were below the established benchmarks on both the Word Attack and Word Identification subtests. None of the poor decoders were identified as such on the basis of the Word Identification subtest alone, since most succeeded in identifying more than 48 words.

From Table 5, it is clear that mean performance between poor and average readers differed greatly on the PPVT, Word Reading and Word Attack pretests, while their scores on both of the CELF-4 measures were more similar. On the decoding tests, students generally

performed better on Word Identification. Students identified as poor decoders performed under at least one standard deviation below the mean on the Word Attack subtest.

Table 6

Analysis of Variance for Testing Condition and Reading Ability on the Pretests

Source	df	SE	F	p	Partial Eta
PPVT					
Read Ability	1	3641.15	38.98	.00**	.34
Test Condition	2	137.95	1.48	.24	.04
RA X TC	2	210.93	2.26	.11	.06
Error	76	93.40			
CELF-Directions					
Read Ability	1	45.51	5.33	.02*	.07
Test Condition	2	4.49	0.53	.59	.01
RA. X Con.	2	4.74	0.55	.58	.01
Error	76	8.531			
CELF-Listening Comprehension					
Read Ability	1	14.05	7.12	.01*	.09
Condition	2	1.64	0.83	.44	.02
RA. X Con.	2	0.02	0.01	.99	.00
Error	76	1.97			
Word Attack					
Read Ability	1	2,692.47	103.07	.00**	.53
Condition	2	9.98	0.38	.68	.01
RA X Con.	2	3.31	0.13	.13	.00
Error	76	26.11			
Word Reading					
Read Ability	1	45.51	5.33	.02*	.07
Condition	2	4.49	0.53	.59	.01
RA. X Con.	2	4.74	0.55	.58	.01
Error	76	8.531			

* $p < .05$. ** $p < .01$

Once students were classified as average or poor readers, they were randomly assigned to one of the three treatment groups – partial read-aloud with pacing, pacing only or standard.

Students were evenly distributed across the conditions, with 27 students (18 average, 9 poor decoders) in the standard accommodation, 27 students (17 average, 10 poor decoders) in the pacing condition, and 28 students (19 average, 9 poor decoders) in the partial read-aloud condition. Analyses of variance were conducted to ensure that the three treatment groups did not differ significantly on any of the pretests. From the ANOVAs in Table 6, it is apparent that no significant main effects or interactions involving treatment condition were detected in any of the analyses. Students performed comparably across the three treatment groups on the word reading, vocabulary, and listening comprehension pretests. Therefore it is safe to conclude that no group had any special advantage or disadvantage that would impact its performance on the reading comprehension test beyond that created by the differing treatment conditions.

Results of the ANOVAs reported in Table 6 reveal main effects of reader ability in all of the analyses indicating that average readers outperformed poor decoders on all of the pretests. Steps were taken to insure that poor decoders included in the present study possessed listening comprehension skill within the normal range. However, this did not create equivalent groups on the CELF listening comprehension measures. As evident in Table 5, mean scores of the poor decoders were a bit lower than those of average readers.

Effects of Reading Ability and the Accommodation Treatment Variable on Reading Comprehension. The central question of interest in the present study was whether and how the accommodation treatments would impact average readers' and poor decoders' success on a reading comprehension test. An ANOVA was conducted with reader ability and accommodation treatment as the independent variables and performance on the Gates-MacGinitie reading comprehension test as the dependent variable. Results are reported in Table 7.

A significant main effect was found for reading ability, with average readers scoring higher across all conditions than poor decoders. A significant main effect was also detected for treatment condition as well as a statistically significant interaction between treatment and reading ability. As shown in Table 8 and Figure 2, the mean score of poor readers was much higher under the partial read-aloud condition ($M = 36.33$) than under either the standard ($M = 28.22$) or pacing conditions ($M = 28.20$), whereas average readers did not show much of a difference between their performance on the partial read-aloud with pacing accommodation ($M = 42.11$), standard condition ($M = 42.17$), or pacing only condition ($M = 40.65$). These results show that the partial read-aloud with pacing accommodation did not favor everyone equally.

Table 7

Analysis of Variance for Testing Condition and Reading Ability on the Gates-MacGinitie Test

Source	df	SE	F	p	Partial Eta
Read Ability	1	2114.51	106.76	.00**	.58
Test Condition	2	163.06	8.23	.00**	.18
RA x TC	2	115.34	5.82	.00**	.13
Error	76	19.81			

** $p < .01$

Bonferoni tests showed that poor decoders comprehended text significantly better in the partial read-aloud with pacing condition than in both the pacing only and the standard conditions, but the pacing only and standard conditions did not differ. In contrast, Bonferoni tests confirmed that average readers comprehended text equally well across all three treatment conditions. These findings provide support for the main hypotheses, that a partial read-aloud with pacing accommodation would improve poor decoders' performance on a reading comprehension test but

would not make a difference on average readers’ performance compared to standard test conditions, and that simply pacing students through the test would not explain this benefit for poor decoders. Differences between means are also displayed on Figure 2.

Table 8

Mean Performance and Standard Deviations as a Function of Testing Condition and Reading Ability on the Gates-MacGinitie Reading Comprehension Test (Maximum score = 48 correct)

Reading Ability	PRAP ^a Mean (SD)	PO ^b Mean (SD)	Standard Mean (SD)	Bonf	<i>d</i> ^c	<i>d</i> ^d
Average Readers	42.11 (4.28)	40.65 (5.12)	42.17 (3.01)	RA=P=S	-.02	-.39
Grade Equivalent	6.2	5.4	6.2			
N	19	18	17			
Poor Decoders	36.33 (4.60)	28.20 (6.70)	28.22 (6.03)	RA>P=S	1.53	-.03
Grade Equivalent	4.4	3.2	3.2			
N	9	10	9			

^a PRAP = Partial Read Aloud with Pacing Condition

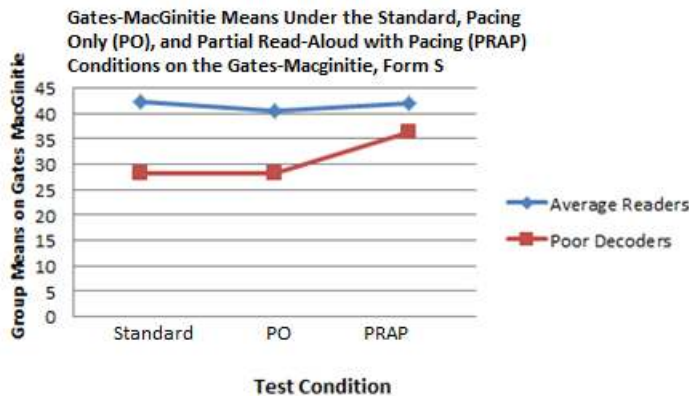
^b PO = Pacing Only Condition

^c Effect size = *M* of Partial Read-Aloud with Pacing Condition minus *M* of Standard Condition divided by pooled *SD*.

^d Effect size = *M* of Pacing Only Condition minus *M* of Standard Condition divided by pooled *SD*.

Figure 2

Group Means Under the Standard, Pacing and Read Aloud Conditions on the Gates-MacGinitie Reading Comprehension Subtest, Form S (Maximum Score = 48 Correct)



In addition, mean performances were examined to determine whether the linguistic backgrounds of the students exerted any impact on these outcomes. Four linguistic groups were

distinguished, a monolingual English group and three groups of bilinguals who differed in whether English was predominant, or their native language was predominant, or neither language predominated. Mean performance of the four groups is shown in Table 9 for each reader ability level.

Table 9

Mean Performance of Students from Different Linguistic Backgrounds on the Gates-MacGinitie Reading Comprehension Test

Test	<u>Average Readers</u> Mean (SD)	<u>Poor Decoders</u> Mean(SD)
Bilingual		
Native Language Prominent		
Standard Condition	44.00 (1.41)	26.00 (5.66)
Grade Equivalent	7.6	3.1
N	2	2
PO Condition ^a	38.50 (2.12)	27.00 (7.07)
Grade Equivalent	4.8	3.1
N	2	2
PRAP Condition ^b	45.00 (0.00)	34.00 (0.00)
Grade Equivalent	8.6	4.0
N	1	1
Bilingual		
English Prominent		
Standard Condition	40.50 (3.99)	33.33 (6.42)
N	6	2
PO Condition	40.50 (6.36)	28.00 (8.19)
Grade Equivalent	5.4	3.2
N	2	2
PRAP Condition	41.33 (3.06)	36.00 (6.68)
Grade Equivalent	5.8	4.4
N	3	4
Bilingual		
Both Languages Used Equally		
Standard Condition	42.57 (2.37)	25.50 (4.43)
Grade Equivalent	6.2	3.0
N	7	3
PO Condition	41.33 (3.00)	34.00 (0.00)
Grade Equivalent	5.8	4.0
N	8	1
PRAP Condition	41.33 (3.06)	38.50 (4.80)
Grade Equivalent	5.8	4.8
N	11	2

Table 9 (Continued)

Mean Performance of Students from Different Linguistic Backgrounds on the Gates-MacGinitie Reading Comprehension Test

Test	<u>Average Readers</u> Mean (SD)	<u>Poor Decoders</u> Mean(SD)
Monolingual		
Standard Condition	43.33 (2.31)	28.22 (6.03)
Grade Equivalent	6.8	3.2
N	3	4
PO Condition	41.33 (3.00)	27.60 (7.79)
Grade Equivalent	5.8	3.1
N	4	5
PRAP Condition	42.11 (4.28)	36.00 (2.83)
Grade Equivalent	6.2	4.4
N	5	2

^a Pacing Only

^b Partial Read-Aloud with Pacing

Inspection of mean performance in Table 9 shows that within each linguistic group, the same patterns of performance were evident. The average readers outperformed the poor decoders consistently, although the gap between the two was narrower in the partial read-aloud with pacing condition. Average readers differed little across the three testing conditions. However, poor decoders showed superior reading comprehension in the partial read-aloud with pacing condition than in the other two testing conditions. Thus, the pattern of results detected when all the groups' performance was combined also held for the separate linguistic groups. Although the majority of students in the sample were bilingual, their varied linguistic backgrounds did not impact the main pattern of outcomes reported previously.

Correlations between Tests

Correlations between the various tests were calculated. The correlation matrix appears in Table 10. Statistically significant correlations at the $p < 0.01$ level were found between the

Gates-MacGinitie reading comprehension scores and the PPVT (.35), the CELF-4 Listening Comprehension subtest (0.34), the WMRT-R Word Attack Subtest (0.63), and the WMRT-R Word Identification test (0.41). The strongest correlations were detected between the word reading tests and reading comprehension. This can be interpreted to support the importance of a partial read-aloud accommodation that lightens word decoding demands among poor decoders and better reveals their reading comprehension ability. Additionally, significant correlations existed at the $p < 0.01$ level between vocabulary and listening comprehension (.36), vocabulary and word reading (.43) and vocabulary and word attack (.49). Only one of the CELF measures was significantly correlated with reading comprehension, that involving the spoken paragraphs test, not the following directions test. This suggests that the paragraph comprehension task may be a better way to measure language comprehension as it explains variance in reading comprehension of text.

Table 10
Correlations Between Pretests and Gates-MacGinitie Reading Comprehension Test

	1	2	3	4	5	6
1. Gates-MacGinitie	———					
2. Word Attack	.628**	———				
3. Word Reading	.414**	.525**	———			
4. CELF CFD ^a	.195	.267*	.070	———		
5. CELF USP ^b	.340**	.295**	.185	.162	———	
6. PPVT	.354**	.487**	.425**	.210	.362**	———
Mean	37.96	20.73	64.09	49.67	13.66	148.41
Standard Deviation	7.12	7.64	8.92	2.96	1.44	11.89

* $p < .05$, ** $p < .001$

^a CELF = Clinical Evaluation of Language Functions, Fourth Edition CFD = Concepts and Spoken Directions Subtest

^b CELF = Clinical Evaluation of Language Functions, Fourth Edition USP = Understanding Spoken Paragraphs Subtest

Anecdotal Observations

The researcher collected observations and feedback from participants. It was common for students in the pacing only condition to be unresponsive and sometimes visibly frustrated. At least two students in the pacing only condition in each cohort had to be told to stop moving ahead and wait for the signal to turn the page. Several average readers questioned why they had to wait for prompting to move on to the next passage or question. During the field study, it was suggested that poor decoders be tested separately to avoid this problem of faster readers having to wait for others to finish before moving on. However, the same problem arose among the average readers in the pacing only condition in the current study. In addition, some students had to be reminded several times to give the signal that they had finished, by placing their pencils on the desk. To address this problem, the investigator double checked that everyone had finished by repeating, “If you’ve answered the question, please show me the signal.”

Additionally, different behaviors were observed in poor decoders across conditions. Poor decoders in the standard and pacing only conditions seemed fatigued and unfocused. While these students were given an accommodated booklet for the pacing only condition in order to help maintain focus on one question at a time, this did not seem to enhance their ability to concentrate on the task at hand. Some got easily distracted, and turned to neighbors to talk in the midst of the assessment. Likewise, average readers were similarly unfocused in this condition. Meanwhile, far fewer distractions were observed in the partial read-aloud accommodation. All students in this condition seemed to be less frustrated and more focused on the task at hand.

Chapter 7 - Discussion

The data resulting from this study provide support and insight regarding the research questions and hypotheses posed. The results carry strong implications for student reading ability classification and the importance of listening comprehension. Additionally, the results support the use of the partial read-aloud with pacing condition as a valid accommodation for students with reading disabilities, particularly those who struggle in decoding.

Student Ability Classification and Listening Comprehension

As mentioned earlier in the study, a method of classifying students as poor decoders or average readers based upon The Simple View of Reading by Gough and Tunmer (1985) was employed. Past research provided support for the claims of the Simple View by showing that students classified as reading disabled vary in their particular areas of struggle. Results of the pretests in the current study further supported this claim. Students classified as reading disabled performed one standard deviation below the mean on the Word Identification or Word Attack subtests based on standardized norms. However, the same students performed similarly to their peers on the Concepts and Following Directions and Understanding Spoken Paragraphs subtests. This is consistent with the findings of Aaron, Joshi and Williams (1999), who also showed that those who were identified as poor readers varied in the nature of their reading disabilities. Most students in the current study were below grade level norms in word reading, but within these norms on listening comprehension. Additionally, there were students who scored within one standard deviation of the mean on the Word Attack or Word Reading subtests, but scored at least one standard deviation below the mean in Concepts and Following Directions or Understanding

Spoken Paragraphs, indicating adequate decoding but poor comprehension skill. These students were eliminated from the study.

The Simple View of Gough and Tunmer (1985) has led other researchers to obtain support for the claim that listening comprehension is a significant indicator of overall reading comprehension. For this reason, two measures of listening comprehension were included as pretests and used to classify students as poor decoders. The correlation matrix shown in Table 10 illustrates that there was a significant correlation between scores on the Understanding Spoken Paragraphs subtest and the Gates-MacGinitie Reading Comprehension test. This is consistent with previous studies. Dreyer and Katz (1992) and Tilstra et al. (2009) each found that listening comprehension was significantly correlated with reading comprehension, and that this correlation increased over time. The correlation between listening comprehension, as measured by Understanding Spoken Paragraphs, and The Gates-MacGinitie Reading Comprehension test was only moderate because the students were third graders. Further, the Concepts and Following Directions subtest was only weakly correlated with reading comprehension. This may be attributed to the fact that following a set of directions in isolation is a different skill than listening to a short text and responding to comprehension questions. This is further illustrated by the nonsignificant correlation between the Concepts and Following Directions and Understanding Spoken Paragraphs subtests. However, this may also be explained by the findings of both Dreyer and Katz (1992) and Tilstra et al. (2009), which showed that the relationship between listening comprehension and reading comprehension increases over time. Thus, in future studies, it may be interesting to try each of these subtests with an older population of students and see if a stronger relationship between listening comprehension and reading comprehension is evident.

Despite the low correlation between one of the listening comprehension subtests and the Gates-MacGinitie test, it is clear that the data from this study supports the claim of The Simple View of Reading. Results showed that there are students classified as reading disabled who struggle specifically in decoding, while others have difficulty in comprehension. Thus, it is important to identify the specific nature of a reading disability when considering the effectiveness of an accommodation.

Research Question 1: Does the partial read-aloud with pacing condition improve performance of a reading comprehension test more than pacing only and the standard conditions? At the onset of the study, it was hypothesized that the partial read-aloud with pacing accommodation would improve scores on a reading comprehension test more than pacing only or the standard condition. Results showed that this was the case, but only for students who were classified as poor decoders. Poor decoders in the partial read-aloud with pacing accommodation performed better collectively than those in either the pacing condition or the standard condition. This is consistent with some past research, particularly studies exploring the read-aloud accommodation in tests of mathematics. Schulte et al. (2001), Helwig et al. (2002), and Helwig and Tindal (2003) all showed that learning disabled students benefitted under a read-aloud accommodation. However, Schulte et al. (2002) found this to be problematic, because there was no differential impact for students with disabilities since average students also improved under this condition. Similar results were found in most research conducted on the read-aloud accommodation on tests of reading comprehension. Erbaum et al. (2004), McKevitt & Elliot (2002) and Melov (2002) all found that LD students improved under the read-aloud accommodation, as did their non-disabled peers. This possibility that all students would benefit from the accommodation was addressed in the second research question of the current study.

Research Question 2: Do poor decoders show significantly greater improvement in reading comprehension with the partial read-aloud with pacing condition than with pacing only and standard conditions, whereas average readers show little improvement in reading comprehension with the partial read-aloud with pacing condition compared to the other conditions? It was hypothesized that poor readers would benefit more from the partial read-aloud with pacing condition than average readers, who would not benefit significantly. Results of the current study supported this hypothesis. There was a significant interaction between student classification and test condition. Students who were identified as poor decoders scored higher on the Gates-MacGinitie Reading Comprehension Subtest under the partial read-aloud with pacing accommodation than they did under either the standard or pacing only conditions, whereas average readers did not vary significantly in their scores under any of the three conditions.

These findings differ from those of previous studies, which have questioned the validity of a read-aloud accommodation on tests of reading comprehension. McKeivitt and Elliot (2003) and Melov (2002) gave reading comprehension tests to students in elementary and middle school and found that all students improved under an accommodated condition. Likewise, a study by Bolt (2006) called into question the validity of a read-aloud accommodation, claiming that it yielded significant measurement error and did not seem to benefit students with learning disabilities. These studies, therefore, claimed that a read-aloud accommodation was not valid since it did not specifically impact the learning disabled students for whom it was intended.

However, the current study differed from previous research in several ways. Present students were given an accommodation specifically intended for poor decoders, and this classification was determined by pretest performance. In contrast, the previous studies used

teacher recommendations or participant IEPs to identify students as learning disabled. This has been noted as being potentially problematic by Spear-Swerling and Sternberg (1996), who have stressed that teacher or IEP recommended accommodations are not always ideal for students with learning disabilities. Additionally, a study by Helwig et al. (2002) found that teacher recommendations were ineffective in establishing a beneficial test accommodation, and students did not tend to perform well under the conditions established by the classroom teacher. Therefore, it may be that when a more specific assessment of a reading disability is used to identify a child's disability, this yields more promising results in identifying the most helpful accommodation.

Further, the nature of the read-aloud accommodation in the current study differed from that of previous studies. In past research, the read-aloud accommodation involved reading an entire test aloud to students. This was done with a proctor (McKevitt & Elliot, 2003; Melov, 2002), video monitor (Helwig et al., 2002; Crawford and Tindal, 2004), or by the students themselves (Erlbaum et al., 2004). In the current study, only directions, proper nouns, and test items were read-aloud to students. Further, the investigator read from a standardized sheet of directions. A meta-analysis by Li (2014) showed that the use of test proctors yielded more significant results than alternative methods of test administration, a point further supported by the current study.

Additionally, the current study included an adapted test booklet that helped to facilitate the nature of the partial read-aloud with pacing and pacing only conditions. It was predicted that students might look ahead at future test items rather than read along with the proctor as each item was encountered. Thus, the publishers of the Gates-MacGinitie Test granted the investigator permission to adapt the test so that there would be the passage and a single item would be

presented at a time on each page. None of the previous studies referenced the use of an adapted booklet to correspond with the pacing of the proctor.

Meanwhile, unlike much of the literature, this study is consistent with findings by Crawford and Tindal (2004). The work of Crawford and Tindal, like the current study, found that there was a significant interaction between classification and administration, in which students who were classified as learning disabled scored higher on a reading comprehension test when presented with a video format than when presented with standard conditions. Unlike the Crawford and Tindal study, however, the current study showed a main effect of reader ability. This may have resulted from administration of the test, which was delivered by the principal investigator reading the items aloud to participants. This also may have been due to the altered nature of the read-aloud accommodation. While Crawford and Tindal read the entire test aloud on the video screen, the investigator in the current study only read aloud test items. Since both results were significant, it may be interesting to further explore if one would be a better format for students using the partial read-aloud with pacing accommodation.

Finally, the current study expanded upon the findings of the field study. The field study showed a significant interaction between the partial read-aloud with pacing accommodation and reading ability. However, the investigator wondered if it was the task of reading items aloud alone that led to those results. Since students were guided through the test, it was considered that the act of pacing students could have improved their scores. However, results of the pacing only accommodation in the current study showed that the act of helping students work through a test one item at a time did not have a significant effect on student scores. On the contrary, the investigator observed that this condition confused and sometimes annoyed some of the students.

Thus, it can be concluded that it was not the pacing of students through the exam that benefitted poor decoders, but, rather, the reading aloud of the test items and choices.

Limitations

While the data of this study supported the hypotheses, there are some limitations to consider in its design. One such limitation involves its sample of participants. In the field study, it was noted that its participants were generally higher level students. Many participants who were classified as average readers scored above grade level norms on the Gates-MacGinitie Test. It was the intention of the principal investigator to seek out for the current study a population of students who would score closer to grade level on the Gates-MacGinitie test. However, the sample selected for this study yielded similar results; many of the students tested above grade level. There are alternative explanations for why this may have happened. Primarily, grade level norms on the Gates-MacGinitie were based on a standardized format in a large-scale study, while the conditions set forth in the current study included accommodated conditions in small groups. High scores could also be due to the age and background of the students. This sample population came from middle-class, high performing schools. Further, some of the students were tested at the end of their third grade year, which could have also improved scores. In future research, it may be beneficial to work with a population of students from more diverse backgrounds to see if there is a difference in their performance.

Additionally, the nature of the pacing only accommodation presents some possible limitations. Pacing only was quite frustrating to all students involved, but the partial read-aloud with pacing condition was not. This could be due to test taking strategies set in place by the participants. Perhaps students typically employ a set of strategies, including previewing

questions and underlining key words, and may have become frustrated with this imposed break in their routine. However, particularly in the case of average readers, this could also be attributed to the design of the study. In order to prevent poor decoders from becoming overwhelmed with text, they were presented with an adapted test booklet in the pacing only condition. Since average readers were not similarly distracted with the text in the pilot study, they were not presented with the adapted booklet in the pacing only condition. While this did not dramatically hurt the performance of the average readers, the nature of the pacing only accommodation was distracting to them, and it may have been worthwhile to see if they would also benefit from an adapted booklet. However, poor decoders also showed limited focus in the pacing only condition regardless of the presence of an adapted booklet, suggesting that the use of the booklet alone did not enable students to improve or show enhanced concentration. Further, all readers were given the same adapted booklet in the partial read-aloud condition, which did not seem to be of any particular benefit to average readers in comparison to the pacing only or standard accommodations. The difference in focus may not be due to the adapted booklet, but, rather, the nature of the accommodation. The issue of having to wait for other students seemed to be most frustrating to those in the pacing only condition. It is possible that, if this test were taken individually, students would show less frustration with the need to wait eliminated. Future studies exploring test accommodations with pacing may help to clarify these issues.

Another point of interest involved performance on pretests. Although participants scored high in reading comprehension, they collectively scored low on the Word Attack task. Students overall did not seem to struggle with Word Identification, but identifying pseudowords and utilizing an understanding of spelling patterns seemed to be a particular area of struggle, even for

average readers. Why this sample was far above grade level in reading comprehension but below grade level in the act of decoding new words remains to be studied.

Further, the sample population was predominately Asian American and bilingual. This may not be representative of a wider population of readers, particularly those who are reported as having reading disabilities. It was taken into consideration that students who are bilingual may react differently to the accommodated conditions than would monolingual students. Comparison of performance patterns showed no differential effects among subgroups of poor decoders. All benefited from the partial read-aloud condition. However, future studies may explore whether this finding is consistent across different linguistic groups, particularly since the majority of participants identified themselves as Asian American. This classification encompasses a variety of languages and alphabetic systems. It may be informative to consider the writing system associated with the native languages of ELL participants, and how this may impact their decoding skills. Additional research would more clearly explore whether any struggle in decoding experienced by ELL students could be attributed to their decoding skill or their limited experience with the English alphabetic system. Further, it would be beneficial to see how this accommodation benefits a larger sample of students from low-income backgrounds, since the majority of this sample included children from a middle class community.

There were some also limitations in the measures and their usage. Primarily, the benchmark scores established for the Word Attack and Word Identification subtests were gathered from the mean W score reported for third grade. This benchmark was used to locate corresponding raw scores one standard deviation below the third grade mean. However, several different age groups fit into this category. Some participants were just about to begin third grade, while others were at the end of third grade. A year of education would likely make a

difference in a child's anticipated score and the benchmark established for ability classification. The use of expected levels of performance to classify individual students may be beneficial in further studies, as a single cut-off point may not be precise enough to reflect the true abilities of participants across different age groups.

Further, there were pros and cons in the use of Form S rather than Form T of the Gates-MacGinitie to test reading comprehension. Form S was originally chosen because it contained fewer proper nouns than Form T. However, after further consideration, it may have been more appropriate to use Form T. In utilizing the measurement tool with more proper nouns, the investigator could have determined how well the partial read-aloud accommodation benefited students when their accommodation included a higher number of low frequency proper nouns. However, Form T had a downside. In the pilot study the mean scores of average readers on Form T were higher than the mean scores on Form S (i.e., $M = 43$ vs. 39 correct). Choice of Form S thus reduced chances of a ceiling effect among average readers in the current study. It turned out that the mean score of average readers on Form S in the dissertation study was somewhat higher than that on Form S in the pilot study ($M = 42$ vs. 39 correct, respectively). However, the mean was still substantially below the maximum score of 48 correct. This strengthens confidence that the absence of differences among average readers across testing conditions was not likely a result of ceiling effects produced by a test that was too easy.

Educational Implications

The findings of this study yield several important educational implications. Primarily, the correlation matrix and abilities of the participants further support the Simple View of Reading by Gough and Tunmer (1985). Moreover, they illustrate the necessity in specifically

identifying what is meant by a reading disability, particularly when a student classified as having such a disability is mandated for test accommodations. Much of the literature on test accommodations explores their impact on students with learning disabilities, and does not clearly identify the nature of that disability (Spear-Swerling & Sternberg, 1996; Schulte, Elliot, Kratochwill, 2001; Li, 2014). As a result, the studies that have previously explored the read-aloud accommodation have not taken into consideration that the learning disabled participants within the study may not struggle in decoding, but, rather, in comprehension. The participants in the current study, who were identified as poor decoders on the basis of pretests, benefitted from the partial read-aloud with pacing accommodation.

This finding also illustrates the benefit in presenting the partial read-aloud accommodation with pacing as opposed to extended time, which is most commonly used. Poor decoders who took the standard form of the Gates-MacGinitie test, which was given with limited time, became unfocused and frustrated. As an educator, the principal investigator has observed struggling readers who have taken standardized tests with extended time reporting a prolonged feeling of frustration and inability to focus. The current study illustrated how the partial read-aloud accommodation gave poor decoders the opportunity to take a standardized test in a way that enhanced their focus, controlled their movement through the text, reduced the load on decoding, and hence improved their performance.

Despite this benefit, read-aloud accommodations have not been adopted due to the results of previous studies, in which entire tests were read aloud to students. This did not have a differential impact on learning disabled participants. This could be due largely to the fact that the entire test was read aloud, which facilitated the process of answering test questions for all test takers regardless of ability (McKevitt & Elliot, 2003; Melov, 2002; Li, 2014). However, in the

current study, when test items, proper nouns, and multiple choice answers were the only portions of the test read aloud, students identified as poor decoders performed much better under the accommodated condition, while average readers did not. This stands to reason; in tests of mathematics, in which the read-aloud accommodation has worked, the proctor in question simply reads the questions. Perhaps enabling students to show their understanding of a passage being read with the barrier of decoding questions removed helps to level the playing field for poor decoders without necessarily overestimating their abilities.

The effectiveness of the partial read-aloud with pacing condition can be explored further if the nature of the words presented to students at the onset of each passage was adapted. In the current study, words shown and read aloud in the partial read-aloud accommodation were limited to proper nouns. However, standardized reading comprehension tests often present multisyllabic words that are an area of struggle for poor decoders. Additional studies may seek to explore the impact of the partial read-aloud with pacing condition if it were further adapted to include these more challenging words in the set of items introduced to students at the onset of the passage.

While this study focused specifically on students classified as poor decoders, further research could also explore how other reading disabled students may respond to the partial read-aloud accommodation. Poor comprehenders were eliminated from this study, but, in future studies, it may be interesting to see if they also show enhanced focus and improved performance while having portions of reading comprehension tests read aloud to them. Since some students struggle in both decoding and comprehension, it is expected that these students could benefit somewhat from the partial read-aloud accommodation, although perhaps not as dramatically due to their difficulty in comprehending text.

Based on the results of the current research, the partial accommodation with pacing might even be considered as a general testing procedure to replace extended time for all students. No harm was done to average readers, and poor decoders benefitted. The nature of this condition enabled all students to be focused and perhaps more relaxed or less anxious. Given these ideal circumstances, future research may seek to explore the benefit of offering the partial read-aloud accommodation to all students regardless of ability.

The current study shows the merit in the partial read-aloud with pacing accommodation on tests of reading comprehension, while past research has questioned the validity of such an accommodation. Despite contrary findings in previous studies, the design and positive results of the current study clearly illustrate that this is a topic worth exploring further. Such an accommodation may be highly beneficial to students who classify as poor decoders.

Appendix A

Information Form to Administrators

The Effect of the Read Aloud Accommodation on the Reading Comprehension Scores of Good and Poor Third Grade Readers
Administrator Information Letter

Thank you for considering your school's participation in a research project, "The Effectiveness of a Partial Read Aloud Accommodation To Assess Reading Comprehension of Students With A Reading Disability". The primary research investigator is Mrs. Michelle Giusto, a third grade teacher at P.S. 221 and graduate student in the PhD program in Educational Psychology at the CUNY Graduate Center. The co-investigator is Dr. Linnea Ehri, a professor in Educational Psychology at The Graduate Center.

Mrs. Giusto will be studying conditions for optimizing performance on tests of reading comprehension. This project will examine whether students of various academic abilities, when taking tests in reading comprehension, will benefit from an accommodation in which the examiner reads aloud all test directions and questions with multiple choice answers.

For children whose parents provide written consent, Mrs. Giusto will give initial reading tests to determine how well these students perform on skills involving word decoding and listening comprehension. This is to select a sample of students who demonstrate grade level comprehension while struggling in decoding. These pre-tests include the Word Identification and Word Attack subtests of the Woodcock Reading Mastery Test Revised (WRMT-R) and two subtests of the Clinical Evaluation of Language Functions 4th Edition Test (CELF-4). The Word Identification subtest asks students to read aloud a series of words in order to determine how well they are identified and decoded by participants. The Word Attack subtest asks students to do the same with a series of nonsense words to ascertain how well students approach them with the decoding skills they have already acquired. The CELF-4 subtests include the Concepts and Directions and Understanding Spoken Paragraphs subtests. The Concepts and Directions Subtest asks students to interpret spoken directions of increasing length and complexity by responding to a verbal request. Students are shown a line of pictures and are directed as to which picture they should point, often asked to interpret ordinal words ("point to the third car") or differentiation in direction or size ("to the right of" "the little car"). The Understanding Spoken Paragraphs subtest asks students to listen to short passages and answer comprehension questions. Pre-tests will be given one-on-one by me at a time arranged between myself and the child's teacher. I will take a group of 3-4 students at a time and work with each one individually. I will ensure that students are not taken out while they are scheduled for mandated services or at any other conflicting time.

Once students have been given all of their pre-tests, they will be given a reading comprehension test. Specifically, I will be administering Forms S and Form T of the grade 3 Gates MacGinitie reading test. This is a test of reading comprehension using a series of short reading passages accompanied by multiple choice questions, similar to the format of the state ELA exam. Students will take one form of this test under one of three possible conditions - standard administration, a pacing condition, or with a partial read aloud accommodation. Under the standard administration, all passages and questions will be read and answered independently by the students. In the pacing condition, I will guide students through the test and advise them when to begin each passage and question. Under the partial read-aloud accommodation, I will show students index cards with any proper nouns in the story and model how these nouns are pronounced, ask students to read the passages independently and read each question and choice to the students before they arrive at an answer. All tests will be given to a larger group at once. Again, I will work with the classroom teacher to arrange the most convenient time to do so.

There are no known risks for children involved in the study. They will be given tasks

Appendix A (Continued)

Information Form to Administrators

quite similar to the E-CLAS and ELA practice exams given to them in school. Children can benefit by receiving individual attention while reviewing important skills and getting additional practice in standardized testing before their state exams in May. All information gathered will be strictly confidential and will be stored in a locked filing cabinet. The identities of all individuals will remain anonymous. We may publish the results of the study, but the names of students, or any identifying characteristics, will not be used in any of the publications. However, any educational information that could help your child's teachers best meet his or her academic needs will be communicated with them, so that they can use the results of these tests to help your child in school.

In order for all children to participate in the study, Mrs. Giusto must obtain written permission from each parent. Parents will be made aware of the fact that participation in the study is voluntary. Parents are free to withdraw their children from the study at any time and for any reason without consequence. In addition, children may choose to withdraw from the study at any time and for any reason without consequence.

If you have any further questions, you can contact me, Michelle Giusto, @ Mmontefinise@gc.cuny.edu, Dr. Linnea Ehri, at 212-817-8294 or Lehn@gc.cuny.edu. If you have any questions about the rights of the participants in this study, you can contact Kay Powell, IRB administrator, The Graduate Center City University of New York, (212) 817-7525, or kpowell@gc.cuny.edu. Thank you again for your participation.

Appendix B
Parent Permission Form

1

CITY UNIVERSITY OF NEW YORK

The Graduate Center

Department of Educational Psychology

**PARENTAL/LEGAL GUARDIAN PERMISSION FORM FOR
CHILD'S PARTICIPATION IN RESEARCH**

Project Title: *Effectiveness of a Partial Read-Aloud Test Accommodation to Assess Reading Comprehension in Poor Decoders*

Principal Investigator:

*Michelle Grusto
Graduate Student
The Graduate Center, CUNY
365 Fifth Avenue
New York, NY 10016
718-496-5147*

Faculty Advisor:

*Dr. Linnea Ehri
Distinguished Professor
The Graduate Center, CUNY
365 Fifth Avenue
New York, NY 10016
212-827-8294*

Site where study is to be conducted:

- P.S. 221, 5740 Marathon Parkway, Little Neck, NY 11362*
- P.S. 115, 80-51 261st Street, Glen Oaks, NY 11004*
- Samuel Field Y/P.S. 178 Summer Program, Queens High School of Teaching, 74-20 Commonwealth Boulevard, Queens, NY 11426*

Introduction/Purpose: Your child is invited to participate in a research study. The study is conducted under the direction of Michelle Grusto. The purpose of the research study is to determine the effect of a partial read-aloud accommodation on the reading comprehension scores of third graders. The results of this study may determine that some students will benefit from a test condition in which test questions, multiple choice items and proper nouns are read aloud to them, while all passages are read independently.

Procedures: Approximately 90 children are expected to participate in this study. I will give initial reading tests to determine how well these students perform on skills involving word reading and listening comprehension. In addition, students will be given a reading comprehension test, including short reading passages with corresponding multiple choice questions. The test will be given once to each student, under one of three possible conditions. Some students will take the exam under its standard format. A second

Appendix B (Continued) Parent Permission Form

group will be given the partial read-aloud accommodation, in which all test questions, multiple choice items and proper nouns will be read aloud. A third group will be given a pacing condition, which will involve a test administrator telling students when to begin and end each test passage and item. All word reading and listening comprehension tasks will be given one-on-one, while the reading comprehension tests will be given to students in a small group. All testing will be done during regular school hours, during which students will work with the primary researcher for two 40-60 minute sessions. I will arrange a schedule with your child's teacher to make sure that testing is done during an ideal time and will not conflict with mandated services, special programs or important projects. The time commitment of each participant is expected to be a span of approximately two hours across two different time periods. One hour will involve tests of listening and reading individual words. The second hour will be spent taking the reading comprehension test under one of the aforementioned conditions.

Possible Discomforts and Risks: There are no known risks for children involved in the study. All information gathered will be kept strictly confidential, and will be stored in a locked filing cabinet, to which only my advisor and I will have access. We may publish the results of the study, but the names of the students, or any identifying characteristics, will not be used in any of the publications. If you would like a copy of the study, please provide me with your address and I will send you one in the future.

Benefits: Children can benefit by receiving individual attention while reviewing important skills and getting additional practice in standardized testing before state exams. This study can also benefit students who have difficulty showing their understanding of their reading under standard test conditions, and can show the merits of having these students work under the partial read-aloud accommodation.

Voluntary Participation: Your child's participation in this study is voluntary, and you may decide to withdraw your child from participation without prejudice, penalty, or loss of benefits to which you are otherwise entitled. If you decide to remove your child from the study, please contact the principal investigator Michelle Giusto to inform her of your decision. Children will also be reminded at the onset of the study that they can stop working with the principal investigator at any time without any consequence. If students seem uncomfortable at any point, the principal investigator will ask them if they would like to continue, and will honor any requests made by children to leave the study.

Confidentiality: The information obtained from your child will be collected in the form of test booklets. The collected records will be accessible only to the primary investigator. The researcher will protect your child's confidentiality by coding the data and storing it in a locked filing cabinet.

Contact Questions/Persons: If you have questions about your rights as a research participant, or you have comments or concerns that you would like to discuss with someone other than the researchers, please call the CUNY Research Compliance Administrator at 646-664-8918. Alternately, you can write to:

CUNY Office of the Vice Chancellor for Research
Attn: Research Compliance Administrator
205 East 42nd Street
New York, NY 10017

Appendix B (Continued) Parent Permission Form

Statement of Consent:

"I have read the above description of this research and I understand it. I have been informed of the risks and benefits involved, and all my questions have been answered to my satisfaction. Furthermore, I have been assured that any future questions that I may have will also be answered by the principal investigator of the research study. I voluntarily agree to allow my child to participate in this study.

By signing this form I have ~~not~~ waived any of my legal rights to which my child would otherwise be entitled.

I will be given a copy of this statement."

Printed Name of
Subject's Legal
Guardian

Signature of Subject's Legal Guardian

Date Signed

Printed Name of the Child-subject

Printed Name of
Person Explaining
Form

Signature of Person Explaining Form

Date Signed

Printed Name of
Investigator

Signature of Investigator

Date Signed

Approved By: _____
Date: _____

Appendix C
Directions for Partial Read-Aloud Condition of Gates-MacGinitie, Form S

ADAPTATIONS TO THE GATES-MACGINITIE TEST FOR THE PARTIAL READ-ALOUD
 CONDITION

After page 38 in the instruction manual, read the paragraph below.

“You will be following along with me as you take the test. When I tell you to read each passage, you will read it silently to yourself. Once you complete the passage, you will stop and wait for the rest of the class before we move on to the questions together.. To signal to me that you have finished, you will place a pencil on the edge of your desk. (Model). Can we try that? Show me what you will do when you finish. (Students show the signal). Good. After we have finished each passage, I will read you each question. After I have read the question and all of the choices, you will choose the best answer based upon what you have read. When you have finished answering your question, you give me the signal to show that you are done. Do not move on to the next question until I tell you to do so. We will follow the same set of rules for each passage and each question; nobody is to move on until they have been told to do so.”

AT THE ONSET OF EACH PASSAGE

**There are ___ special words in this passage. [Here they are. (Show them). Can you say _____? (For each word). Good!].

Read this passage independently. When you are finished, give me the signal. Do not move on until we are all ready to do so.

AT THE ONSET OF EACH QUESTION

Look at question number _____. Read the question, and select the choice that you think best answers it. When you are finished, stop and give me the signal.

**This will vary from story to story. Follow the key below to fill in the blanks appropriately. For each, say the words listed here and show them on corresponding index cards. If there are no proper nouns in the story, just say “No” and omit the portion in the brackets.

Page of the Test Booklet	Words
8	Frog, Toad
9	None – Skip

Appendix C (Continued)
Directions for Partial Read-Aloud Condition of Gates-MacGinitie, Form S

Page of the Test Booklet	Words
10	David
12	None - Skip
13	Katie, Main Street, Children's Day, Tuesday
14	Gogo
16	None
17	Old Red
18	Inuit
20	None
21	None

Appendix D
Directions for Pacing Condition of Gates-MacGinitie, Form S

ADAPTATIONS TO THE GATES-MACGINITIE TEST FOR THE PACING CONDITION

**After page 38 in the instruction manual.

“You will be following along with me as you take the test. You will read each passage and answer each question independently, but I will ask you to remain at the same point of the test at the same time. When I tell you to read each passage, you will read it silently to yourself. Once you complete the passage, you will stop and wait for the rest of the class until answering each question. To signal to me that you have finished, you will place a pencil on the edge of your desk. (Model). Can we try that? Show me what you will do when you finish. (Students show the signal). Good. After we have finished each passage, you will answer each question one at a time. Do not begin the question until I tell you to do so. You will read each question independently. Then, you will answer it based upon what you read. When you have finished answering your question, you give me the signal to show that you are done. Do not move on to the next question until I tell you to do so. We will follow the same set of rules for each passage and each question; nobody is to move on until they have been told to do so.”

AT THE ONSET OF EACH PASSAGE

Read this passage independently. When you are finished, give me the signal. Do not move on until we are all ready to do so.

AT THE ONSET OF EACH QUESTION

Look at question number _____. Read the question, and select the choice that you think best answers it. When you are finished, stop and give me the signal.

Appendix E Student Assent Form

SCRIPT FOR OBTAINING STUDENT'S ORAL ASSENT PRIOR TO THE STUDY

"Today, if you would like to help me out, I will ask you to show me how well you can read words and listen to directions. Before I tell you a little more about what you may be doing with me today, I want to let you know that you do not have to participate if you do not wish to do so, today or any other day; that is your choice and you can stop at any time. If you would like to participate, this will be your chance to show off. I will ask you to read a list of words to me, follow directions and point to some pictures and answer questions about short stories that I will read to you. It will take about fifteen minutes.

"After that, in a few days, if you would like to, you will take a practice test in reading just like the ones you have taken in your teacher's classroom. You might take one on your own like you usually do, or you might take it another way, which will be a little different. Instead of taking the test all by yourself, I might tell you when to read each story and answer each question, or I might read the questions to you. You will only take one of these tests. When you do, you can show me how much you're learning about reading in school and what a great reader you are.

"Now that you know what will be asked of you, you can continue to work with me if you want to. You are free to tell me you would like to stop at any time; during our tests today if you wish, or later on when we take the practice tests. If you stop, it will be okay. About 30 other students will be taking part in my project too.

"Do you agree to work with me and take part in the project?" (Obtain child's assent).

Child's willingness to participate in the research study "Effectiveness of a Partial Read-Aloud Test Accommodation to Assess Reading Comprehension in Poor Decoders":

Child agrees _____ Child says no _____

References

- Aaron, P.G., Joshi, M., Williams, K.A. (1999). Not all reading disabilities are alike. *Journal of Learning Disabilities*, 32(2), 120-137
- Bolt, S.E., Ysseldyke, J.E. (2006). Comparing DIF across math and reading/language arts tests for students receiving a read-aloud accommodation. In C. Hulme & R.M. Joshi (Eds.) *Applied Measurement in Education* 19(4), 329-355.
- Cook, L., Eignor, D., Sawaki, Y., Steinberg, J., Cline, F. (2010). Using factor analysis to investigate accommodations used by students with disabilities on an English-Language Arts Assessment. *Applied Measurement in Education* 23, 187-208.
- Crawford, L., & Tindal, G. (2004) Effects of a read-aloud modification on a standardized reading test. *Exceptionality*, 12, 89-106.
- Dunn, L.M., Dunn, D.M. (2007). PPVT-4: The Peabody Picture Vocabulary Test. Minneapolis, MN: Pearson.
- Dreyer, L., & Katz, L. (1992). An examination of the simple view of reading. iHaskins Laboratories Status Report on Speech Research, SR-III, 161-166
- Educational Records Bureau Comprehensive Testing Program II* (1987). Princeton, NJ. Educational Testing Service.
- Elbaum, B., Arguelles, M.E., Campbell, Y., & Saleh, M.B. (2004). Effects of a student-read-aloud accommodation on the performance of students with and without learning disabilities on a test of reading comprehension. *Exceptionality*, 12, 71-87.
- Elliot, S.N. & Marquart, A.M. (2004) Extended time as a testing accommodation: its effects and perceived consequences. *Exceptional Children*, 70 (3), 349-367.
- Fletcher, J.M., Frances, D.J., Boudousquie, A., Copeland, K., Young, V., Kalinowski, S., & Vaughn, S. (2006). Effects of accommodations on high-stakes testing for students with reading disabilities. *Exceptional Children*, 72(2), 136-150.
- Fuchs, L.S., Fuchs, D., & Capizzi, A.M. (2005). Identifying appropriate test accommodations for students with learning disabilities. *Focus on Exceptional Children*, 37, 1-8
- Gough, P.B., & Tunmer, W.E. (1986). Decoding, reading and reading disability. *Remedial and Special Education*, 7, 6-10.
- Gregg, N. & Nelson, J.M. (2012) Meta analysis on the effectiveness of extra time as a test accommodation for transitioning adolescents with learning disabilities: More questions than answers. *Journal of Learning Disabilities*, 45(2), 128-138.

- Hollenback, K. (2002). Determining when test alterations are valid accommodations or modifications for large-scale assessment. In G. Tindal & T.M. Haladyna (Eds.) *Large-scale assessment programs for all students* (pp 395-426).
- Li, H. (2014). The effects of read aloud accommodations for students with and without disabilities: A meta-analysis. *Educational Measurement Issues and Practice*, 33(3), 13-16.
- Joshi, R.M., Williams, K.A., & Wood, J.R. (1998). Predicting reading comprehension from listening comprehension: Is this the answer to the IQ debate? In C. Hulme & R.M. Joshi (Eds.) *Reading and spelling: Development and disorders* (pp. 319-325).
- Kim, J.S. (2012) The effect of “read-aloud” as a test accommodation for students with visual impairments in South Korea. *Journal of Visual Impairment and Blindness* (pp. 319-325).
- Helwig, R. & Tindal, G (2003). An experimental analysis of accommodation decisions on large-scale mathematics tests. *Exceptional Children*, 69 (2), 211-225.
- Helwig, R., Rozeck-Tedesco, M.A., & Tindal, G. (2002). An oral versus standard administration of a large-scale mathematics test. *Journal of Special Education*, 36, 39-47.
- MacGinitie, W.H., MacGinitie, R.K., Maria, K., & Dreyer, L.G., & Hughes, K. (2000). *Gates-MacGinitie Reading Tests, fourth edition*. Itasca, IL: Riverside.
- McKevitt, B.C., & Elliott, S.N. (2003). Effects and perceived consequences of using read-aloud and teacher-recommended testing accommodations on reading achievement. *School Psychology Review*, 32, 583-600.
- Melov, L. L. (2002). The effect of a read aloud accommodation on test scores of students with and without a learning disability in reading. *Remedial and Special Education*, 23, 248-55.
- Nation, K., & Snowling, M. (1997). Assessing reading difficulties: the validity and utility of current measures of reading skill. *British Journal of Educational Psychology*, 67, 359-370.
- Schulte, A.L., Elliot, S.N., & Kratochwill, T.R. (2001). Effects of testing accommodations on standardized mathematics test scores: an experimental analysis of the performance of students with and without disabilities. *School Psychology Review*, 30, 527-547.
- Semel, E., Wiig, W.H., & Secord, W.A. (2003). *Clinical evaluation of language fundamentals – fourth edition..* San Antonio, TX: Pearson.
- Schrank, F.A., McGrew, K.S., & Woodcock, R.W. (2001). *Woodcock Johnson III tests of achievement*. Itaska, NY: Riverside Publishing

- Spear-Swerling, L. & Sternberg, R.J. (1996). *Off-track – When Poor Readers Become “Learning Disabled”*. Boulder, CO.: Westview Press.
- Stanovich, K. (1991). Discrepancy definitions of reading disability: has intelligence led us astray? *Reading Research Quarterly*, 26 (1), 7-29
- Thurlow, M.L., Lazarus, S.S., Thompson, S.J., Morse, A.B. (2005). State policies on assessment participation and accommodations for students with disabilities. *The Journal of Special Education*, 38(4), 232-240.
- Tilstra, J., McMaster, K., Van den Broek, P., Kendeou, P., Piapp, D. (2009). Simple but complex: components of the simple view in reading across grade levels. *Journal of Research in Reading*, 32(4), 383-401.
- Weschler, D. (2005). *Weschler Individual Achievement Test, Second Edition*. London: The Psychological Corp.
- Woodcock, R.W., McGrew, K.S., Mather, N.. (2001). *Woodcock Johnson III Tests of Achievement*.. Rolling Meadows, Illinois: Riverside Publishing
- Woodcock, R.W. (1987). *Woodcock Reading Mastery Tests – Revised*. New York, NY; Pearson.