2024

IN SUPPORT OF A READ-ALOUD ACCOMMODATION: A META ANALYSIS STUDY

Sarah Elizabeth Ryan

IN SUPPORT OF A READ-ALOUD ACCOMMODATION: A META ANALYSIS STUDY

A dissertation submitted in partial fulfillment
of the requirements for the degree of

DOCTOR OF PHILOSOPHY

to the faculty of the

DEPARTMENT OF EDUCATION SPECIALTIES

of

THE SCHOOL OF EDUCATION

at

ST. JOHN'S UNIVERSITY

New York

by

Sarah Elizabeth Ryan

Date Submitted November 27, 2023 Date Approved January 31, 2024

Sarah Elizabeth Ryan	Dr. Michael Sampson

ABSTRACT

IN SUPPORT OF A READ-ALOUD ACCOMMODATION: A META-ANALYSIS STUDY

Sarah Elizabeth Ryan

This study examined the read-aloud accommodation benefits for students with disabilities (SWD) and students without disabilities (SWOD). In this study, researchers recommend ways in which to remove encumbrances faced by SWD in reading comprehension. In general, research has shown that both SWD and SWOD have met a level of success from the read-aloud accommodations. Further, this meta-analysis will attempt to reveal key factors that influence the effects of read-aloud accommodations. For example, how the accommodation effect size influences the subject area of reading versus that of math. This study will also show that the effect of read-aloud accommodations is significantly stronger when human proctors read the test than when read by video/audio players or computers. Student, parent, and teacher perceptions regarding the use of the read-aloud testing accommodation are examined, along with the relationship between student self-efficacy of testing accommodations and their disability status as well as their grade level. Research has found that most parents and teachers perceived testing accommodations to be fair and valid for SWD. This study will further investigate the perceived positive consequences of testing accommodations by providing further evidence of their continued use as well as their role in facilitating valid test scores.

Keywords: read aloud, accommodations, meta-analysis, disabilities

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1 INTRODUCTION

Background

What have decades of research revealed about the nature of comprehension and how to develop students' comprehension in schools? Scientific research has revealed individual instructional practices and combinations of practices that can be used to foster reading comprehension development. A review of the research literature seems to show that beyond the stages of reading acquisition, superior reading ability is not associated with a greater tendency to use the verbosity that is inherent in natural language to speed word recognition. Instead, general comprehension strategies and rapid context-free word recognition appear to be the processes that most clearly distinguish well from poor readers.

Over the last two decades, experimental psychologists have developed a renewed interest in phenomena that have a heavy cognitive part (i.e., mental imagery, psycholinguistics). Nevertheless, nowhere is this clearer than in the literature on the reading process (Venezky, 1977). Cognitive psychologists have recently applied their information processing perspective to components of reading performance (Massaro, 1975, 1978). For instance, there was a strong tendency among early cognitive theorists to depict information processing as a series of discrete stages, each performing a specific transformation on its input and passing on the new recoded representation as an input to a later stage (Stanovich, 1980).

Because the sequence of processing operations progresses from the incoming data to higher-level encoding, such conceptualizations have been termed bottom-up models. It is not surprising that because these models were so influential in the early development of

information processing theorizing, they were the first to be applied to reading. Therefore, several bottom-up, serial-stage models of reading and word recognition have been introduced into the literature. However, it is now reasonably well-established that such a move is inadequate because it does not account for important empirical results in the reading literature. Researchers Rumelhart and Danks discuss experimental phenomena (e.g., word, syntactic, and semantic context effects) that supply problems for bottom-up models (Stanovich, 1980).

However, there exists a class of models that conceptualize reading in a manner that is diametrically opposed to that embodied in serial-stage models. These have been termed top-down models because higher-level processes interact with, and direct the flow of information through, lower-level processes. Although top-down conceptualizations of reading exist, they all have in common a view of the fluent reader as being actively engaged in hypothesis-testing as he/she continues through the text. Since the reader is only sampling textual information to evaluate hypotheses, the reading process is viewed as being driven by higher-level conceptual processes rather than by low-level stimulus analysis.

In short, top-down analysis starts with hypotheses and then try to verify them by processing the stimuli, while bottom-up analysis starts by processing the stimuli. While the top-down hypothesis-testing models have often been attacked for excessive vagueness in their conceptualization, there are even more serious criticisms of the entire class of models. This argument against hypothesis-testing models is reinforced by research that shows that fluent readers do not use conscious expectations to help word recognition. Researchers have questioned the hypothesis-testing models because they need

implausible assumptions about the relative speeds of the processes involved. Specifically, the generation of hypotheses about a next word, or words, must take less time than is necessary to recognize the words based on purely visual information, otherwise the hypothesis generation is unnecessary. However, it seems unlikely that a hypothesis based on complex syntactic and semantic analysis can be formed in less than the few hundred milliseconds required for a fluent reader to recognize most words in the text (Stanovich, 1980).

Reading comprehension relies on the ability to read most or all the words in a text. Word reading is a necessary—although not sufficient—condition for reading comprehension. Students who have previously experienced repeated connecting of graphemes, phonemes, and meanings of words including orthographically complex words will be able to devote less attention to word reading and more to comprehension. As one might expect, given these examples, instruction aimed at improving students' word reading, including through phonemic awareness and phonics instruction, often has a positive impact on reading comprehension (e.g., Suggate, 2014). Also, important to reading comprehension are knowledge and skills that bridge word reading and comprehension. One such skill is graphing phonological semantic cognitive flexibility (GSF), a form of executive function (and a fancy label for a down-to-earth skill). GSF is the ability to simultaneously address, and flexibly switch between, the letters and sounds in words (graph phonological) and the meanings of words (semantic). Research has found a relation between readers' GSF and their reading comprehension; thus, researchers have examined whether there are ways to bolster students' GSF (Duke et al., 2021).

Morphological awareness—or attention to the smallest meaningful parts of words (e.g., roots, affixes, words in compound words)—also bridges decoding and comprehension. Instruction in morphological awareness affects many contributors to reading comprehension, and more intensive morphological awareness instruction positively affects reading comprehension directly (e.g., Goodwin & Ahn, 2013). Morphological awareness instruction is proper for younger and older children. In fact, one study on morphological awareness instruction in preschool showed positive impacts on reading comprehension in grade 6 (Lyster, Lervåg & Hulme, 2016). Reading fluency, reading with accuracy, automaticity, and prosody also serve as a bridge between decoding and reading comprehension.

Reading words accurately supports comprehension because it helps readers build a correct basic understanding of the text. Reading words with automaticity allows readers to devote more cognitive attention to comprehension. Reading with prosody (e.g., proper expression, phrasing) both reflects and supports comprehension. Instructional practices such as Readers Theater and choral, echo, and paired reading, which are aimed at improving prosody and other aspects of reading fluency, have been shown to positively affect reading comprehension in multiple studies (e.g., Kuhn, 2020; Turner, 2010).

As important as foundational word reading and bridging skills are for reading comprehension, research has shown that they are not sufficient for strong comprehension. For example, just because a reader can fluently read the words in a text does not mean that they will glean the characteristics of the text, as that understanding requires a broader range of knowledge, strategies, and dispositions. A key piece of evidence that foundational word reading and bridging skills alone are not sufficient for reading

comprehension is the existence of students who have poor reading comprehension despite strong decoding and reading fluency. Researchers have seen and investigated such cases for decades (e.g., Oakhill, 2020), and research has suggested that they are a large portion of students who are not proficient on state reading tests (e.g., Koon, Foorman & Galloway, 2020). Comprehension-focused interventions can foster reading comprehension in students who have specific difficulty with reading comprehension (Lee & Tsai, 2017).

Given the absolute necessity of foundational word reading skills, it is tempting to think that instruction should begin with a focus on developing those and later turn to comprehension. However, research has supported a simultaneous, rather than sequential, model of reading instruction. Along with the development of phonological awareness, print concepts, and alphabet knowledge, young learners in preschool and early elementary school benefit from efforts to develop oral language comprehension, including efforts to develop oral comprehension of written language (i.e., through reads-aloud, e.g., Cervetti, 2020; Swanson et al., 2011).

As young learners begin to read texts themselves, comprehension instruction, alongside phonics and other foundational skills instruction, has a prominent place. For example, comprehension monitoring supplies a form of feedback to readers whether they have read a word accurately. While reading the sentence "I can get a dog," a young reader might understandably pronounce get as jet (i.e., pronouncing *g-* as in *g -em* rather than *g-* as in *got*). It is comprehension monitoring that would alert readers that *jet* is not correct, so they may reread, try a different sound for *g*, and thus be left with a correct orthographic mapping for the word.

Some students monitor comprehension as they read without instructional support; however, others do not (Kinnunen, Vauras & Niemi, 1998), and researchers have long concluded that teaching comprehension monitoring is effective (National Institute of Child Health and Human Development, 2000). The relation between word reading instruction and reading comprehension instruction is more synergistic than competitive. Despite the wealth of research on how effective the read-aloud is as a literacy tool, researchers have found that fewer teachers seem to be attempting to read what they consider sophisticated stories and nonfiction books in preschool and kindergarten in favor of reading easier, predictable, and concept books (often in Big Book format), especially in classrooms with high percentages of at-risk children. Sophisticated picture books include, for example, stories in which readers must infer characters' motivations and thoughts and connect them to actions (i.e., causes and effects). Such books have a rich repertoire of vocabulary.

These books can be contrasted with predictable books in which readers do not need to infer character motivation, feelings, or thoughts to enjoy the repeated words and actions. Examples include Martin, B. (2011). *Brown Bear*, *Brown Bear*, *What Do You See?* or Wadsworth. (1986). *Over in the Meadow*.

While predictable books play a role in preschool and kindergarten literacy programs, sophisticated picture books play an added role of expanding vocabulary and enhancing oral comprehension. According to Serafini, (2020), reading aloud is the single most important activity for developing proficient readers. However, increasingly more children do not have access to books, are read to regularly, and do not have exposure to the written language. It is imperative for such students to learn to read and write to have

academic and life successes. Reading aloud builds many important foundational reading comprehension skills, introduces vocabulary, while also supplying a model of fluent and expressive reading.

Cochran-Smith's (1985) argument emphasizes that effective teachers model the role of an ideal reader as they read aloud. An ideal reader is one who intuitively and unconsciously makes proper inferences and predictions, and constantly rethinks current events in a story in relation to past events. Thus, effective teachers' model what ideal readers do by explicitly talking aloud as they read, making children aware that they are predicting, making an inference, or changing their ideas about what is happening in a story (e.g., Keene & Zimmerman, 1997; Miller, 2012).

The importance of reading aloud to students in school is well proven in the literature; controversy concerning the effectiveness of reading aloud as pedagogy suggests that a teacher's purpose and process for executing the read-aloud in instruction matters. However, reading aloud to students in various forms was a common and useful practice in the early years of literacy development. In fact, more than 40 years ago, McCormick (1977) discovered in her literature review a "direct relationship between reading aloud to children and reading performance, language development, and the development of reading interests." Research shows that the read-aloud helps building a community of readers and writers (Berkowitz, 2011), which contributes to positive academic benefits (Dreher, 2003) across the curriculum.

Statement of the Problem

A landmark study by the National Research Council (McDonnell, McLaughlin & Morison, 1997) explored the clear exclusion of SWD from many of the benefits of

educational reform. It confirmed the importance of the participation of SWD in large-scale assessments, particularly statewide assessments, if SWD were to receive help from standards-based educational reforms. This study also pointed to the importance of accommodations for students to have access to the general curriculum, and to be able to show their knowledge and skills on assessments. Perspectives on assessment accommodations have changed quite dramatically over time. Initially, accommodations were viewed to level the playing field as well as to improve the validity of assessment results and to supply a way for students to take assessments when they otherwise would not be able to take them. These goals do not necessarily mean that student performance will improve (Thurlow, Elliott & Ysseldyke, 1998).

Furthermore, accommodations were viewed as an important part of instruction for SWD. However, instructional accommodations covered changes such as changing the number of items in a classroom assignment or including only easier parts of an assignment—changes well beyond those considered for assessment (Elliott & Thurlow, 2006; Nolet & McLaughlin, 2005). Despite the recommendation that only accommodations used in instruction should be used for assessment, concerns grew about making changes that altered what a test was intended to measure. As policy and practice evolved, so did the language surrounding accommodations. With time, such changes that produced valid results were called accommodations, while those that altered what the test was intended to measure were called modifications. Almost exclusively, the early belief was that accommodations were reasonable for providing only to individuals with disabilities, and even then, there were many concerns about whether they provided an advantage to the SWD using them (Koretz & Barron, 1991; Koretz & Hamilton, 2001).

Shifts in these early perspectives occurred over time with the enactment of policy imperatives for SWD.

Perspective shifts were also included in assessments to the point that the discussion focused on the disadvantage imposed on students who did not receive needed accommodations. Accommodations are designed to decrease noise and maximize the strength of the inference based on a student's score, and without them, many students are at a disadvantage in demonstrating what they know and can do (Madaus, Russell & Higgins, 2009, p. 182) (Martha L. Thurlow, 2014).

Who Needs Accommodations

Three federal laws support the provision of accommodations to individuals with disabilities: (a) Section 504 of the Vocational Rehabilitation Act of 1973, (b) the Individuals with Disabilities Education Act (IDEA), (c) and the Americans with Disabilities Act (ADA). Section 504 states that an institution receiving federal funds must assure that admissions tests are selected and administered so as best to ensure that when a test is administered to an applicant who has a handicap that impairs sensory, manual, or speaking skills, the test results accurately reflect the applicant's aptitude or achievement level, or whatever other factor the test purports to measure, rather than reflecting the applicant's impaired sensory, manual, or speaking skills (except where such skills are the factors that the test purports to measure) (Section 84.42[b] [3]).

This was the first sign that it was considered discriminatory for a test to reflect an individual's disability rather than his or her knowledge and skills. Section 504 allows for the provision of accommodations to individuals with disabilities, including those of school age regardless of their eligibility for special education services. Students who

receive accommodations but are not receiving special education services are "504 students" who have 504 accommodation plans.

The reauthorization of IDEA in 1997 introduced, for the first time, the notion of access to state and district-wide assessments and referenced accommodations as an aspect of participation in assessments:

- In general – All children with disabilities are included in all general State and district-wide assessment programs, including assessments described under section 1111 of the Elementary and Secondary Education Act of 1965, with proper accommodations and alternate assessments where necessary, and as indicated in their respective individualized education programs.
- (B) Accommodation guidelines – The State (or, in the case of a district-wide assessment, the local educational agency) has developed guidelines for the provision of right accommodations (20 U.S.C. Section. 612 [17]). IDEA was reauthorized in 2004, and added a requirement that states report on the number of students receiving accommodations for state assessments.

ADA was reauthorized in 2008 and added clarifications of the definition of a "disability," which was defined in 1990 as an individual with a physical or mental impairment that substantially limits one or more major life activities, who has a record of such an impairment, or who is regarded as having an impairment. The clarifications included, for example, expanding the "illustrative" list of "major life activities" to include as examples, concentrating, thinking, and communicating.

The three federal laws (Section 504 of Rehabilitation Act, IDEA, and ADA) clearly view accommodations as being for individuals with disabilities. In addition, the Elementary and Secondary Education Act (ESEA), reauthorized as the No Child Left Behind Act in 2002, recognized the need for accommodation for English language learners. States decided who could and who could not receive accommodation in their assessments. It is noteworthy that some states showed that all students were eligible for accommodation during state testing. However, as universal design and accessibility concepts appeared along with the belief that assessments needed to be accessible for a wide range of students, the belief grew that changes previously viewed as "accommodations" needed by SWD or ELLs were appropriate for all individuals.

Characteristics That Lead to A Need for Accommodations

Although the focus on accommodations in K–12 education started with a focus on students with disabilities, policymakers soon recognized that the category of a student's disability (such as learning disability [LD], speech language impairments [SLI], or intellectual disability [ID]) was not a good basis for making decisions about the need for accommodations or which accommodations might be needed by an individual student (Thurlow, Elliott & Ysseldyke, 1998). Instead, the disability-related characteristics of a student (e.g., limited reading skills, poor memory, or being easily distracted) or the linguistic characteristics of a student (Koenig & Bachman, 2004) were identified as being the proper basis for making decisions about accommodations. Table 1 supplies a sample of characteristics, needs, and accommodations identified in some accommodation manuals as well as other tools available on the internet (Thurlow, Albus & Christensen, 2015) (Martha L. Thurlow, 2014).

Table 1

Student Characteristics, Needs, and Possible Assessment Accommodations

Student Characteristic	Need	Possible Accommodation
Weak manual dexterity; difficulty with pencil; difficulty typing on standard keyboard	Assistance with holding materials, using pencil, typing	Express response via scribe or pointing Use speech to text technology
Poor decoding skills	Assistance with reading when the skills measured is not decoding skills	Read AloudText to Speech Technology
Short attention span: easily distracted	Assistance with focusing on assessment or key parts of assessment	Highlighting of keywords Masking or template to mark placeLocation near teacher or proctor Read AloudText to speech technology

Note. Adapted from Tables C, E, and I in Thurlow, M. L., Albus, D., & Christensen, L. L. (2009). Minnesota manual of accommodations for students with disabilities in instruction and assessment: A guide to selecting, administering, and evaluating the use of accommodations. Roseville, MN: Minnesota Department of Education (Martha L. Thurlow, 2014).

Despite the increased research that is being conducted, and the fact that experimental and quasi-experimental studies are being done, there are still many issues with the research (Thurlow, Lazarus & Christensen, 2013). For instance, some studies have focused on students with one category of disability label (e.g., learning disabilities), although students with such label can vary considerably in their characteristics as well as in their needs for specific accommodations. One also knows from studying SWD and SWOD that there are overlapping characteristics that are likely to create a need for accommodation. When it is assumed that a student needs accommodation based solely on a category of disabilities, and that any student without a disability label does not need accommodation, there is a great likelihood that the research findings will be confounded (Thurlow et al., 2009).

Studies have clearly defined students with disabilities in terms of the characteristics that are indicative of a specific need that might be addressed by accommodation. For example, Laitusis (2007) further grouped students into those with reading-based learning disabilities and those without disabilities. With this clearer definition of the groups included in the study, she also used a repeated-measures approach, with each student taking a test with an audio presentation and without an audio presentation, in counterbalanced order. She found that students with reading-based learning disabilities benefited more than SWOD. Earlier research on this accommodation had produced conflicting findings, most likely due to the existence of the same characteristics and needs within the two groups being compared. Although the research by Laitusis focused on the audio presentation, students receiving the audio presentation also received extended time and recording answers in a test booklet rather than a bubble sheet as accommodations (Laitusis, 2007).

Unlike some earlier experimental research, Laitusis (2007) recognized that accommodations are rarely used in isolation. Rather, when students use accommodation, it is often used in conjunction with other accommodation (Elliott et al., 1999). The use of more than one accommodation was recognized by McKevitt and colleagues (2003) in their early work using single-subject designs, as well as by Fletcher and associates (2006, 2009) in their study of a bundled accommodation package that included reading of proper nouns, question stems, and answer choices as well as testing across two days. Crawford (2005) concluded in a report on the value and validity of testing accommodations by recommending that there be "an expansion of the research base related to 'bundled accommodation,' including the effect of bundled, or packaged, accommodations on the

construct being measured, on each other, and on individual students" (Hollenbeck & Crawford, 2005, p. 14).

In addition to research on the effectiveness of accommodations, studies have examined issues surrounding the implementation of accommodations (DeStefano, Shriner & Lloyd, 2001; Rhode Island Department of Education, 2003). Examination of the implementation of accommodations is important because research has found that there is a relationship between the provision of accommodations and participation in assessments (Anderson, Jenkins & Miller, 1998). Not long after the enactment of IDEA requirements for students to take part in assessments, with accommodations as needed, researchers began to examine the implementation of accommodations during instruction and during assessments (DeStefano, Shriner & Lloyd, 2001; Shriner & DeStefano, 2003).

The Read-aloud As Testing Accommodation for Students with And Without Disabilities

Earlier qualitative combinations of research on differences in test scores associated with read-aloud accommodations have offered no consensus on its effect. As part of a larger study, Elbaum (2007) computed the average effect size difference between SWD and SWOD. The results suggested that elementary school SWD experienced a greater score boost from the accommodation than did SWOD, but that at the secondary school level, SWOD benefited more from the accommodation than SWD. These mixed results reinforce the need to conduct a meta-analysis to combine evidence across studies quantitatively (Buzick & Stone, 2014).

The grade-level effect of the read aloud found by Elbaum (2007) is one example of a moderator variable that has been considered to have an impact on the strength or

direction of effect sizes. While results of individual research articles may or may not contribute support for the use of accommodations, meta-analysis can introduce greater statistical power to detect effects and supply context as well as a measure of consistency across studies. Conducting meta-analysis studies on the read aloud enables researchers to find conclusive evidence by quantitatively summarizing effect sizes for both SWD and SWOD.

Research Questions

The focus of this study will be on two research questions: (1) What are the effects of read-aloud accommodations for SWD and SWOD? (2) Which factors are likely to influence the effects of read-aloud accommodations?

Theoretical Framework

In recent years, the increasing participation of SWD in statewide assessments has stimulated considerable research and discussion concerning the proper assignment of testing accommodations, the impact of accommodations on test performance of SWD and SWOD as well as the validity of interpretations of test performance when students are awarded particular accommodations. Issues surrounding testing accommodations have important implications both for accountability systems and for individual students. Accountability systems must consider whether the test scores of students who are awarded accommodation can be considered equal with those of students who take the tests without accommodation (Elbaum et al., 2004a).

For individual students, the right assignment of accommodation for high-stakes tests could make the difference between passing to the next grade and retention, or between exiting school with or without a standard diploma. The urgency of this issue has

led to accumulating research literature on testing accommodations for SWD. To date, reviews of the testing accommodations research show that considered for groups of students, the effects of accommodations on test performance are generally quite small. A meta-analysis by Chiu and Pearson (1999, as cited in Tindal, 2003) found that studies using general education students as a comparison group yielded an overall weighted mean accommodation effect size for all target student population of 0.16, with a standard error of 0.02.

Chiu and Pearson's meta-analysis synthesis also revealed large and statistically significant variation in the effects associated with different accommodations, thus supporting the need to understand the effects associated with specific accommodations. Tindal and others have noted the complex nature of the effects of accommodation, underscoring the importance of investigating possible interactions between the effect of accommodation, student disability, student skill level in the area being tested, and characteristics not only of tests but also of specific test items (Elbaum et al., 2004b). A large percentage of SWD have serious difficulty with reading and are candidates for testing accommodations on tests of reading comprehension. Students with learning disabilities (LD) make up almost 50% of all SWD, and many of such students have individualized educational program goals in reading (Elbaum et al., 2004a).

Studies have investigated the impact of testing accommodation on the performance of SWD on tests of reading comprehension. For instance, Fuchs et al. (2000) administered a reading assessment to fourth- and fifth grade SWD and fourth grade SWOD under four different testing conditions: standard, large print, extended time, and student read-alouds. The students did not receive help from extended time or large prints.

Marquart (2000), as cited in Elliott (et al., 2002) similarly found no statistically reliable effect for extended time accommodation on reading tests. In contrast, SWD in the study by Fuchs et al. were shown to have benefited significantly more than the SWOD from reading the passages aloud. For the student read-aloud accommodation, there was a significant difference in the effect of accommodation for SWD (effect size [ES] = 0.06) and SWOD (ES = –0.12) (Elbaum et al., 2004a).

Consequently, of the various accommodations for reading tests that have been investigated, only the one allowing the student to read the passages aloud has been found to produce a differential gain in the performance of SWD, and the reasons to explain this pattern of results have varied. When extended time does not enhance the scores for SWD, it may be because their knowledge and skills in a particular area may not be commensurate with the difficulty of the test.

When extended time enhances the performance of all students, not only SWD, but the accommodation is also not considered valid. The theoretical implication is that if time affects all students equally, then all students should take the test under the same time conditions, or else, students taking the test in the accommodated condition would have an unfair advantage over other students. The explanation for the absence of impact with large prints is more straightforward. For students with no visual impairment, there is no benefit to having passages displayed in a larger font. This is equally true for SWD and SWOD (Elbaum et al., 2004a).

Regarding students reading aloud, Goldman, Hogaboam, Bell, and Perfetti (1980) studied elementary school students' recall of specific words read within a sentence and across a sentence boundary. They divided their sample of students into those of higher

and lower reading ability based on teacher reports, and had students read the stimulus material in one of two conditions: silently or aloud. They found that students of lower reading ability, particularly younger students (third vs. fourth graders), had greater recall for texts just processed when they read them aloud. Although this finding was somewhat incidental to their main investigation, it fit with the view of reading comprehension as being dependent on holding just-read texts in short-term memory until sufficient text (usually a clause) has been processed to encode a complete meaning unit. It would follow that if reading aloud aids less highly skilled readers to recall specific texts long enough to enhance comprehension, then allowing students to read the passages of a reading test aloud might constitute appropriate testing accommodation for students with LD (Goldman et al., 1980).

Based on the evidence provided by Goldman et al., this accommodation would not be likely to benefit more highly skilled readers and therefore would meet the criterion for valid testing accommodation. This is in fact what Fuchs et al. (2000) found to be the case. However, based on the findings by Goldman et al., it is unclear whether much older students (i.e., those in middle school and high school) would reap the same advantage (Goldman et al., 1980).

Several researchers have also investigated teachers' and/or students' beliefs of the impact of specific testing accommodations on student performance. For instance, Fuchs et al. (2000) found that teachers were not accurate in their assignment of testing accommodations. They awarded accommodations to students who did not receive help from them and did not award them to students who did benefit from them. Helwig and Tindal (2003) similarly found that teachers were not effective in predicting who would

benefit from accommodation. Furthermore, McKevitt and Elliott (2003) reported that eighth-grade students, responding to a questionnaire about test accommodations, thought they did better on tests when accommodations were provided. However, no analysis was conducted to verify whether the degree to which students perceived the accommodation to be effective was associated with the accommodation boost they experienced (McKevitt & Elliott, 2003).

Conclusion

Overall, the literature to date has shown weak effects on accommodation (Chiu & Pearson, 1999; Tindal, 2002). Nevertheless, this weak effect on a population of students (e.g., SWD) still does little to inform the selection of accommodation for a particular student. That is, in the absence of any other information, the mean accommodation boost associated with the assignment of accommodation to a group of students supplies the best estimate of the impact for any individual student. However, prediction will improve tremendously if information is available on the students' prior experience with a particular accommodation, especially if the student was afforded multiple assessment opportunities both in the classroom and in test situations using a variety of accommodations (cf. Helwig & Tindal, 2003; Tindal et al., 1998).

A caution in the implementation of test accommodations, especially on high-stakes tests, is that accommodations are not a remedy for low levels of skill in the construct that is being assessed. The large amount of attention being paid to providing SWD with right accommodations may suggest to some students and their families that the "right" combination of accommodations will result in students achieving an adequate level of performance in a test (Elbaum et al., 2004a).

Additionally, as pointed out by Elliott et al. (2002), accommodation may in fact remove a disability-related barrier for the student evaluated, but still not have a significant effect on scores. In conclusion, studies have added another piece to the experimental literature on testing accommodations for students with disabilities. The findings of research studies are consonant with those from previous research in suggesting that the challenge of assigning the most effective and appropriate testing accommodations for SWD (e.g., that of designing the most effective and appropriate instructional programs for these students) is unlikely to be successfully addressed by precepts affecting entire populations of students defined by their category of disability. Instead, much more attention will need to be paid to individual students' characteristics as well as to their responses to accommodations in relation to types of tests and testing situations (Elbaum et al., 2004a).

The study explored the possible benefits of reading aloud to students in the classroom. (Coiro, 2013). Discussions centered on how the interactive read-aloud, a powerful literacy tool, supplies so many instructional purposes to motivate, encourage, excite, build background, develop comprehension, help children in making connections, and serve as a model for fluent reading (Wadsworth, 2008).

Highlighting research by Lane and Wright (2007) supports the fact that reading aloud promotes a range of skills related to emergent literacy and can yield important academic benefits for children. A purposeful and interactive read-aloud, along with follow-up conversations, gives teachers an opportunity to scaffold learning for students lacking in background knowledge (Wadsworth, 2008). The read-aloud is used to motivate students to read, build topical knowledge, and to model discussions relating to text. The

read-aloud models expressive and enthusiastic reading, transmits the pleasure of reading, and invites listeners to be readers (Richardson, 2000, p. 3). Emphasis will be placed on how the read-aloud accommodation is one way that teachers can promote literacy at any age while helping to increase reading test scores for students with and without disabilities (Daisey, 1993).

Researchers have shown that the read-aloud is a productive way of exposing students to effective listening comprehension skills. Through this, they can also develop and increase their vocabulary, concepts of print, concepts of story as well as their background knowledge. Children develop their concepts of patterns and structures of the written language through listening to the read-aloud. Furthermore, children who will experience a read-aloud come to school better able to understand the components of structure and the function of narrative texts. Some researchers even note that the experience of reading aloud helps children better express themselves, connect with others, and to understand the world around them (Fisher et al., 2004).

CHAPTER 2 LITERATURE REVIEW

The topic of read-aloud accommodation is particularly relevant to students from diverse backgrounds. Many of such students are identified as struggling learners and are often among those identified as having a disability or limited English proficiency. The lessons learned from disability research and practice that have application to students other than those with identified disabilities or limited English ability are addressed through accommodations in the array of alternate assessments that federal policy has allowed for SWD (Elbaum et al., 2004a).

According to Thurlow, Lazarus, Thompson, and Morse (2005), there are five major accommodation categories: (a) timing alternative test schedules, (b) response to alternative ways of responding to the assessment, (c) setting changes to test surroundings, (d) equipment and materials the use of additional devices or references, and (e) presentation – alternative ways to present test materials. Read-aloud accommodations, the focus of this study, presents test materials in an alternative way. The effects of such accommodations are complicated by the involvement of different students, subject areas, accommodation delivery methods, and other factors (Thurlow, 2007).

Read-aloud accommodations are typically used for math tests, with the expectation that the accommodation will not change the construct being tested. Elbaum (2007) summarized four types of findings about the effects of read-aloud accommodations on math tests. The first group of studies reported a significantly positive result for SWD, with little or no effect for SWOD (e.g., Tindal, Heath, Hollenback, Almond & Harniss, 1998). The second group found significantly positive effects for all students, although the effects were stronger for SWD (e.g., Weston, 2003). The third

group showed significantly positive effects for all students, with no significant difference about the magnitude of the effects for SWD compared to those without disabilities (e.g., Meloy, Deville & Frisbie, 2002). The fourth group found no meaningful results for either group of students (e.g., Helwig & Tindal, 2003).

Compared to read-aloud accommodations for math tests, read-aloud accommodations are much more controversial in the context of reading tests. According to the simple view of reading, reading comprehension involves two components: decoding and linguistic comprehension (Hoover & Gough, 1990). Decoding refers to rapidly deriving a representation from printed input, while linguistic comprehension refers to taking lexical information and deriving sentence as well as discourse interpretations. Read-aloud accommodations make decoding words easier, which further helps reading comprehension. However, many researchers take the position that supplying read-aloud accommodation for a reading test changes the construct being measured and, therefore, should not be allowed.

This issue is still controversial. For example, as Crawford and Tindal (2004) have argued, although providing read-aloud accommodations in a reading test may change the skills being tested from reading comprehension to listening comprehension, listening and reading comprehension are so highly correlated that such an accommodated test still provides information about students' reading skills (Li, 2014).

Moreover, according to Laitusis (2010), when decoding skills are not considered to be a part of reading comprehension, reading a reading test aloud does not necessarily change the construct being evaluated. Studies have focused on using read-aloud accommodations for reading tests, and inconsistent results have been reported. For

instance, Meloy et al. (2002) and McKevitt and Elliott (2003) found similar gains for students with and without disabilities because of receiving read-aloud accommodations on reading tests. Crawford and Tindal (2004) and Laitusis (2010), however, found a differential boost from the read-aloud accommodation compared to the non-accommodation condition for SWD, compared to SWOD.

In summary, studies on the use of read-aloud accommodations in reading tests present mixed findings, and whether we should supply read-aloud accommodations in reading tests continues to be a controversial issue (Thurlow et al., 2010).

A few studies (e.g., Calhoon, Fuchs & Hamlett, 2000; Miranda, Russell & Hoffmann, 2004) have explored whether the effects of read-aloud accommodation differ depending on how it is delivered. Often, a human proctor—either a teacher or a test administrator—reads the test to students (e.g., Elbaum, 2007). The test is read to students by a video or audio player (e.g., Helwig & Tindal, 2003), and in others, the read-aloud accommodation is delivered via computers (e.g., Burch, 2002).

In an experimental study, Calhoon et al. (2000) did not find a significant difference between the effects of the read-aloud accommodation delivered by a human proctor and the accommodation delivered by a computer. However, 65% of the students in that study reported that they preferred receiving accommodation via computers due to the anonymity this method afforded them. Certainly, it would be interesting to decide whether the method of delivery has any bearing on the effects of read-aloud accommodations (Li, 2014).

Researchers have also found that grade level is related to the effect of read-aloud accommodations. For instance, in the study by Laitusis (2010), the differential boost was

greater in grade 4 than in grade 8, both for students with and without disabilities. In a meta-analysis of read-aloud accommodations in math tests for SWD, Elbaum (2007) found that the effect of accommodation was stronger for SWD than for SWOD at the elementary school level; however, the converse was true for secondary school students. Laitusis et al. (2012) also noted that read-aloud accommodation studies involving either middle school or high school students showed less effect compared to studies involving elementary school students.

Grade level, therefore, is a key factor in considering the effects of read-aloud accommodations. To supply read-aloud accommodation, extra time is sometimes allowed in the accommodated condition, not because this is a purposeful aspect of the accommodation design, but because the accommodation necessitates it (Olson & Dirir, 2010). For instance, extra time may be needed to turn a video player on and to change a tape. Therefore, when the read-aloud accommodation shows an effect, it is important to figure out whether extra time has addled the observed effect (Harker & Feldt, 1993).

Sireci et al. (2005) reviewed 59 studies on test accommodations for SWD, 23 of which used read-aloud accommodations. Despite the mixed results, they concluded that read-aloud accommodations in math tests appeared to lead to a more valid interpretation of the math achievement of SWD (Li, 2014).

Test accommodations for SWD found that the read-aloud accommodation for math tests appeared to be called for and that the accommodation effects were influenced by many factors. Read aloud accommodations could be used for English language arts (ELA) tests describing decoding as not a part of the construct being measured, and in

middle school if the read-aloud accommodation is given without significantly extending the testing time (Li, 2014).

Additionally, meta-analysis studies have been performed on test accommodations for SWD. For instance, Chiu and Pearson (1999) conducted a meta-analysis of diverse types of test accommodations for both English language learners and SWD. Among the 40 effect sizes for SWD, only five involved presentation formats (i.e., read-aloud accommodations). Chiu and Pearson (1999) found that on average, SWD had a score gain of .16 standard deviation units because of receiving test accommodations (Li, 2014).

Nevertheless, no conclusion has been drawn specifically about the use of read-aloud accommodations for SWD. Elbaum (2007) performed a meta-analysis to decide the effects of read-aloud accommodations on math tests for SWD. In total, 17 studies were included, published between 1998 and 2003. The effect sizes were examined across grade levels. For elementary school students, the effect sizes ranged from .10 to .82, while for secondary school students, the effect sizes ranged from -.07 to .30.

Recently, Vanchu-Orosco (2012) performed a meta-analysis of different types of test accommodations for SWD and based on 119 comparisons from 34 studies conducted and/or published from 1999 to 2011, she concluded that the effect size of test accommodations for SWD was .30, while the effect size for SWOD was .17. However, despite the large scale of this meta-analysis, Vanchu-Orosco did not specifically study read-aloud accommodations (Li, 2014).

As suggested by Zenisky and Sireci (2007), there is a need for more well-constructed meta-analysis of specific accommodations. Therefore, Li (2014) performed a meta-analysis to figure out the effects of read-aloud accommodations for SWD and

SWOD, and to investigate which factors are likely to influence these effects. The proposed meta-analysis differed from earlier meta-analyses in three major respects. First, Li's meta-analysis included a larger number of read-aloud accommodation studies than previous ones. For example, studies on both math and reading tests were included from both published and unpublished sources. Second, Li's meta-analysis focused exclusively on read-aloud accommodations, such that researchers were able to consider a larger number of variables to explain the accommodation effects such as subject area, accommodation delivery method, grade level, extra time, and research design. Third, unlike Chiu and Pearson (1999), Elbaum (2007), and Vanchu Orosco (2012), Li's meta-analysis used the variance known HLM approach (Li, 2014).

Historically, researchers have used fixed-effect models for meta-analyses, with the assumption that the effect size in each study is an estimate of a common effect size of the entire population of the studies (Hunter, Schmidt & Jackson, 1982). In contrast, random-effects models assume that the included studies are random samples drawn from a population of studies, such that the findings can be generalized beyond the studies included in the meta-analysis (DerSimonian & Laird, 1986). Raudenbush and Bryk (1985, 2002) proposed a two-level variance-known hierarchical linear modeling (HLM) approach to meta-analysis, which is regarded as the mixed-effects model (Fischer & Mansell, 2009). It goes beyond the random-effects approach by testing whether there is systematic variance that can be explained by study characteristics beyond simple random variation (Lipsey & Wilson, 2001).

Subjects are regarded as nested within the primary studies included in the meta-analysis. The level-1 model investigates how effect sizes vary across studies, while the

level-2 model explains the potential sources of this variation by examining multiple predictors of effect sizes simultaneously. Using a simulation study, Noortgate and Onghena (2003) have shown that the variance known HLM approach generally produces less biased estimates compared to the fixed effects approaches, unless the number of studies is small. Li (2014) was particularly interested in discovering factors that influenced the effects of read-aloud accommodations. Therefore, due to its flexibility (Hox, 2010), Li chose the variance known HLM approach for his meta-analysis study (Li, 2014).

CHAPTER 3 METHODS

Study 1

In the meta-analysis study, *The Effects of Read-Aloud Accommodations for Students with and Without Disabilities: A Meta-Analysis. March 2014. Educational Measurement Issues and Practice 33(3):3–16 DOI: 10.1111/emip.12027*, Researcher Hongli Li's focus is on the read-aloud accommodations that have been proposed as a way of helping remove barriers faced by SWD in reading comprehension. Li focused on empirical studies examining the effects of read-aloud accommodations; however, the results were mixed. These studies used a variance known HLM approach that was based on 114 effect sizes from 23 studies. An added meta-analysis was conducted to examine the effects of read-aloud accommodations for students with and without disabilities. In general, the studies concluded that SWD and SWOD received help from the read-aloud accommodations, and the accommodation effect size for SWD was significantly larger than the effect size for SWOD.

Further, the same meta-analysis study revealed the key factors that influence the effects of read-aloud accommodations. For example, the accommodation effect was significantly stronger when the subject area was reading than when the subject area was math. The effect of read-aloud accommodations was also significantly stronger when the test was read by human proctors than when it was read by video/audio players or computers (Li).

According to the National Center for Education Statistics (NCES), in 2020–21, the number of students ages 3–21 who received special education services under the Individuals with Disabilities Education Act (IDEA) was 7.2 million, or 15 percent of all

public-school students. Among students receiving special education services, the most common category of disability was specific learning disabilities (33 percent) (NCES).

Among students who received special education services under IDEA in school year 2020–21, the category of disabilities with the largest reported percentage of students was "specific learning disabilities." A specific learning disability is a disorder in one or more of the basic psychological processes involved in understanding or using spoken or written language that may manifest itself in an imperfect ability to listen, think, speak, read, write, spell, or do mathematical calculations. Students with autism, developmental delays, intellectual disabilities, and emotional disturbances each accounted for between 5 and 12 percent of the students served under IDEA. Students with multiple disabilities, hearing impairments, orthopedic impairments, visual impairments, traumatic brain injuries, and deaf blindness each accounted for 2 percent or less of those served under IDEA (NCES).

The aim of IDEA is to ensure that SWD receive help from standards-based reforms, and that they achieve high educational standards. However, a major concern is that historically, general large-scale assessments, which are intended for all students but those who take part in alternate assessments, were developed without consideration of SWD, and thus may constitute an additional challenge for SWD (Dolan, Hall, Banerjee, Chun & Strangman, 2005). To ensure that SWD are appropriately included in state testing programs, test accommodations have been proposed to level the playing field by removing construct-irrelevant variance caused by disabilities (Fuchs, Fuchs, Eaton, Hamlett & Karns, 2000; Lai & Berkeley, 2012).

Among existing test accommodation strategies, the read-aloud accommodation is one of the most used for SWD (Sireci, Scarpati & Li, 2005). With this accommodation, the test (or certain parts of it, such as directions, questions, or prompts) is read to students by a teacher or a device, in addition to the printed text (Thurlow, Moen, Lekwa & Scullin, 2010). The read-aloud accommodation is primarily provided to students with learning disabilities (Crawford & Tindal, 2004), and it is thought that students who struggle to decode written texts will receive help from this accommodation (Bolt & Roach, 2009).

Researchers Fuchs and Fuchs (1999) introduced the differential boost framework that is often used to evaluate the effects of read-aloud accommodations. In this framework, both SWD and SWOD are expected to receive help from the accommodation; however, SWD benefit differentially more than SWOD. A more strictly defined version of this framework is the interaction hypothesis (Sireci et al., 2005; Zuriff, 2000), according to which students who need the accommodation should benefit from it and students who do not need the accommodation should not benefit from it.

The interaction hypothesis is more stringent in that SWOD should not receive help from the accommodation. Empirical studies have examined the effects of read-aloud accommodations for SWD; however, the results are mixed (Elbaum, 2007). Further, it is not clear which factors influence the heterogeneous effects of read-aloud accommodations. Li's findings in the study suggest a quantitative synthesis of earlier studies that are of particular importance in regard to providing solid information about read-aloud accommodations to educators and policymakers (Li, 2014).

Methods

Search and Inclusion Criteria. Li selected studies for inclusion in the meta-analysis based on the following criteria. First, only studies in which a read-aloud accommodation featured as the single test accommodation strategy for SWD and/or SWOD were eligible for inclusion. Second, to address the issue that studies reporting significant effects are more likely to be published (Glass, 1977), Li considered both published and unpublished studies. Third, only studies with an experimental or quasi-experimental design and that present sufficient information for calculating effect sizes were included. Finally, due to the small number of read-aloud accommodation studies in the context of science tests, only studies involving math or reading tests were considered.

Li used the following procedures to search for eligible studies. First, using various combinations of key words and phrases such as "read-aloud," "oral," and "test accommodation," we searched several well-known online databases, including ERIC, JSTOR, ProQuest, and PsycINFO. Second, reviews on test accommodations for SWD and SWOD were used (e.g., Chiu & Pearson, 1999; Elbaum, 2007; Laitusis et al., 2012; Rogers et al., 2012; Sireci et al., 2005; Vanchu-Orosco, 2012), as well as major journals for relevant articles. Finally, references cited in the studies that had already been determined to be eligible were added to those that had not already been found through other sources. After an initial search, 94 studies were retrieved, among which 71 were located, studied and consequently excluded because they did not meet the study inclusion criteria. All of the 94 studies initially retrieved were published or released before 1990.

It was discovered that most of the eligible studies involved more than one comparison (or effect size). For instance, Laitusis (2010) included two groups of SWOD

and two groups of SWD, such that this one article generated four comparisons (or effect sizes). Multiple strategies have been used to address this issue (Marsh, Bornmann, Mutz, Daniel & O'Mara, 2009). First, the multiple effect sizes within each study can be averaged, or one effect size can be selected from each study. As a result, the number of effect sizes is drastically reduced. Second, this dependence can be modeled by adding a third level to the variance known HLM analysis. This approach, however, is constrained by the number of studies included in the meta-analysis. Third, the dependence can be ignored when it is right to do so. For instance, as recommended by Borenstein, Hedges, Higgins, and Rothstein (2009, p. 223), when each of the subgroups in a single study contributes independent information, the "independent subgroups are no different than independent studies." Although ignoring the dependence slightly biases standard errors downward, Marsh et al. (2009) did not find much difference between results of the method that models dependence as a third level and results of the method that ignores the dependence.

Li noted that the researcher Vanchu-Orosco (2012) also obtained comparable results, whether she averaged the multiple effect sizes from a single study or ignored the dependence. Because the samples used to calculate the four effect sizes were mutually exclusive in Laitusis (2010), Li treated each single comparison from the Laitusis study as the unit of analysis in the meta-analysis. Similarly, after a thorough search and screening, Li found that 114 comparisons from 23 studies were eligible for inclusion in the present meta-analysis (See Appendix A for a list of the studies included and their characteristics). Nevertheless, a single study contributed different numbers of effect sizes to this meta-analysis, ranging from 1 (e.g., Johnson, 2000) to 16 (e.g., Helwig & Tindal, 2003). A

sensitivity analysis was performed to evaluate whether excluding a 13-particular study with many effect sizes would substantially change the results of the meta-analysis.

Based on the literature, the study named five variables—disability status, subject area, delivery method, grade level, and extra time—which were all identified as closely related to the effects of read-aloud accommodations. These variables were later used as potential predictors to account for variations in the effect sizes among the studies. The sample size, mean test score, and standard deviation for both the experimental groups and the control groups were also extracted to calculate effect size statistics. The author coded the studies according to a coding scheme, and a measure of inter-rater reliability (percentage agreement) was used to calculate each coded variable.

Discussion

What Are the Effects of Read-Aloud Accommodations for Students with and Without Disabilities? According to the differential boost framework, because of receiving accommodations, SWD are expected to obtain a larger increase in their scores compared to SWOD. The results of Li's meta-analysis support the requirement of this framework. In both Models 1 and 7, disability status was a statistically significant predictor, showing that the effect of read-aloud accommodations for SWD was significantly stronger than that for SWOD, whether or not other predictors were controlled. Specifically, in Model 1, the accommodation effect size was .14 for SWOD and .27 for SWD. These results do not differ substantially from those from previous meta-analyses. For instance, Vanchu-Orosco (2012) reported an effect size of .30 for SWD and .17 for SWOD for multiple types of test accommodations. In Elbaum (2007), the mean effect size was .37 for elementary school SWD and .10 for secondary school SWD.

The interaction hypothesis says that students who need accommodation should benefit from it and that those who do not need accommodation should not benefit from it. Here, the estimated effect sizes in Model 7 are referred to when grade level, extra time, and research design were controlled for. As shown in Figure 2, when the subject area was reading, regardless of the accommodation delivery method, both SWD and SWOD received help from receiving read-aloud accommodations, with effect sizes ranging from .26 to .61. When the subject area was math, both SWD and SWOD received help from read-aloud accommodations provided by human proctors, with effect sizes of .35 and .22, respectively. However, for math tests, when the accommodation was provided by a computer or a video/audio player, the effect sizes for SWD were exceedingly small and the effect sizes for SWOD were zero or almost zero. Therefore, the read-aloud accommodations did not always meet the criteria of the interaction hypothesis.

Summary of the Study

Except when read-aloud accommodations were provided in math tests via a computer or a video/audio player, both SWD and SWOD received help from the accommodation, although the effect size was generally greater for SWD. The fact that SWOD may also receive help from read-aloud accommodations, however, raises a fairness and validity issue (Li & Suen, 2012; Phillips, 1994). If read-aloud accommodations are only provided to SWD, SWOD may be at a disadvantage because they could have received help from the accommodations as well. In other words, the accommodation may even offer SWD an unfair advantage over SWOD.

Although studies have addressed the effects of read-aloud accommodations, more research is needed to fully understand the fairness and validity of test accommodations (Li, p. 23).

Study 2

In their study *Effects of a Student-Reads-Aloud Accommodation on the Performance of Students with and Without Learning Disabilities on a Test of Reading Comprehension*, Batya Elbaum, Maria Elena Arguelles, Yvonne Campbell, and Maya Bardawil Saleh (2010) investigated the impact of a student reads-aloud accommodation on the performance of middle school and high school students with and without learning disabilities (LD) on a test of reading comprehension. Data for the analyses came from 311 students (n = 230 with LD) who took alternate forms of a reading test in a standard and an accommodated condition. In the accommodated condition, students were instructed to read each passage aloud at their own pace, and then to read each comprehension question and the response choices aloud before marking their answer. As a group, students' test performance did not differ in the two conditions, and students with LD did not benefit more from the accommodation than students without LD. However, students with LD showed greater variability in their response to the accommodation such that they were almost twice as likely as students without LD to show a substantive change in test performance in either the positive or negative direction. The findings from this study underscore the need to go beyond the interpretation of group mean differences in deciding the validity of testing accommodations.

The increasing participation of SWD within the past several years in statewide assessments has stimulated considerable research and discussion concerning the proper

assignment of testing accommodations, the impact of accommodations on test performance of students with and without disabilities, and the validity of interpretations of test performance when students are awarded accommodations. Issues surrounding testing accommodations have important implications both for accountability systems as well as for individual students (Elliott & Roach, 2002; Individuals with Disabilities Education Act, 1997; No Child Left Behind Act, 2001; Tindal, 2002). Accountability systems must consider whether test scores of students who are awarded accommodations can be considered equal with scores of students who take the tests without accommodations. For individual students, the right assignment of accommodations for high-stakes tests could make the difference between passing to the next grade and retention or between exiting school with or without a standard diploma.

The urgency of the issue has led to a burgeoning research literature on testing accommodations for SWD. To date, reviews of the testing accommodations research (e.g., Chiu & Pearson, 1999; Thurlow & Bolt, 2001; Tindal, 2002) show that considered for groups of students, the effects of accommodations on test performance are generally quite small. A meta-analysis by Chiu and Pearson (1999, as cited in Tindal, 2002) found that studies using general education students as a comparison group yielded an overall weighted mean accommodation effect size for all target population students of 0.16, with a standard error of 0.02. The synthesis also revealed large and statistically significant variation in the effects associated with different accommodations, thus supporting the need to understand the effects associated with specific accommodations. Tindal and others have noted the complex nature of accommodation effects, underscoring the importance of investigating possible interactions between the accommodation effect,

student disability, student skill level in the area being tested, and characteristics not only of tests but also of specific test items. A significant percentage of SWD have severe difficulty in reading and are candidates for testing accommodations on tests of reading comprehension. SWD make up almost 50% of all SWD, and these students have individualized educational program goals in reading. This is the population of SWD we investigated in this study.

Studies have investigated the impact of testing accommodation on the performance of students with LD on tests of reading comprehension. For instance, Fuchs et al. (2000) administered a reading assessment to fourth- and fifth grade SWD as well as fourth graders SWOD under four different testing conditions: standard, large print, extended time, and student reads aloud. Students did not receive help from extended time or large print. Marquart (2000, as cited in Elliott et al., 2002) similarly found no statistically reliable effect for an extended time accommodation on reading tests. In contrast, SWD in the study by Fuchs et al. benefited significantly more than SWOD from reading the passages aloud. For the student reads-aloud accommodation, there was a significant difference in the accommodation effect for SWD (effect size [ES] = 0.06) and SWOD (ES = –0.12).

Thus, of the various accommodations for reading tests that have been investigated, only the one which allows the student to read the passages aloud has been found to produce a differential gain in the performance of SWD. The reasons to explain this pattern of results have varied. When extended time does not enhance the scores of SWS, it may be because their knowledge and skills in a particular area may not be commensurate with the difficulty of the test. When extended time enhances the

performance of all students, and not only those with LD, but the accommodation is also not considered valid. The theoretical implication of this is that if time affects all students equally, then all students should take the test under the same time conditions, or else students taking the test in the accommodated condition would have an unfair advantage over other students.

The explanation for the absence of the impact of large print is more straightforward. For students with no visual impairment, there is no benefit to having passages displayed in a larger font. This is equally true of students with and without reading disabilities. Regarding students reading aloud, Goldman, Hogaboam, Bell, and Perfetti (1980) studied elementary school students' recall of specific words read within a sentence and across a sentence boundary. They divided their sample of students into those of higher and lower reading ability based on teacher reports and had students read the stimulus material in one of two conditions: silently or aloud. They found that students of lower reading ability, particularly younger students (third vs. fourth graders), had greater recall for text just processed when they read it aloud. Although this finding was somewhat incidental to their main investigation, it fit with the view of reading comprehension as being dependent on holding just-read text in short-term memory until sufficient text (usually a clause) has been processed to encode a complete meaning unit. It would follow that if reading aloud helps less highly skilled readers to recall specific text long enough to enhance comprehension, then allowing students to read the passages of a reading test aloud might constitute an appropriate testing accommodation for SWD. Based on the evidence provided by Goldman et al., this accommodation would not be likely to benefit more highly skilled readers, and thus would meet the criterion for a valid

testing accommodation. This is in fact what Fuchs et al. (2000) found to be the case. However, based on the findings by Goldman et al., it is unclear whether older students (i.e., those in middle school and high school) would reap the same advantage.

Teachers' and/or students' beliefs of the impact of specific testing accommodations on student performance has also been investigated. Fuchs et al. (2000) found that teachers were not accurate in their assignment of testing accommodations. They awarded accommodations to students who did not receive help from them and did not award them to students who did receive help from them. Helwig and Tindal (2003) similarly found that teachers were not effective in predicting who would receive help from an accommodation. McKevitt and Elliott (2003) reported that in responding to a questionnaire about test accommodations, eighth-grade students thought they did better on tests when accommodations were provided. However, no analyses were conducted to verify whether the degree to which students perceived the accommodation to be effective was associated with the accommodation boost they experienced.

Consequently, the study by Elbaum, Arguelles, Campbell, and Saleh were designed to accomplish two primary goals. First, extending the work of Fuchs et al. (2000) to an older group of students, the researchers in this study wished to examine the impact of the student reads-aloud accommodation on the reading test performance of SWD and SWOD in middle and high schools. Second, the researchers of the study also wished to investigate the accuracy of students' feelings of the impact of this accommodation on their test performance in reading.

Method

Participants. Participants in the study were 456 students (283 with LD; 276 male) in grades 6 through 10. The students were recruited from six schools (three middle schools and three high schools) in a large urban school district in the southeastern United States. The school population in the district is highly diverse in terms of race, ethnicity, and socioeconomic status.

Measures. The reading tests used in this study were constructed using third- to fifth-grade level reading passages and accompanying comprehension questions designed for use as test preparation exercises in language arts classes. Researchers Elbaum, Arguelles, Campbell, and Saleh found out in advance that the specific passages being used had not been included in any practice activities in the participating schools. On-grade-level passages were administered to a sample of students with LD in grades 7 and 9 attending schools that were comparable to schools taking part in the study. The passages and test questions were similar in content, presentation format, and response format to those on the statewide reading assessment. The purpose of the pilot test was to discern whether the performance of the target students in a standard administration condition was adequate for the study. That is, test passages that produced a floor or ceiling effect would not yield correct information on the potential benefits of an accommodation. The distribution of students' scores on a set of grade-level passages was in fact highly positively skewed, with students unable to answer more test items correctly. Given this outcome, the researchers assessed the students on a choice of third- through fifth-grade level passages.

Discussion. This study examined the effects of a student reads-aloud accommodation on the performance of middle and high school SWD and SWOD on a test of reading comprehension. Overall, the test scores that students achieved in the accommodated condition were not statistically significantly different from those obtained in the standard condition. For 17% of the SWD, the accommodation raised their performance; for 20%, the accommodation compromised their performance. Considering SWOD, 10% showed an accommodation benefit, while 11% showed an accommodation detriment. Although an ANOVA revealed no statistically reliable difference in the accommodation boost for students with and without LD, the analysis of residualized gain scores suggested that at approximately equal levels of performance in reading comprehension, SWOD may, as a group, benefit more from this accommodation than SWD.

This study cited recommendations in the literature concerning the experimental investigation of testing accommodations. Tindal, Heath, Hollenbeck, Almond, and Harniss (1998) urged that "to provide the most convincing empirical support for an accommodation, students with a specific need have to be compared to others without such a need, who are otherwise comparable in achievement" (p. 442). In this study, a comparison group of general education students was chosen that was very close to the reading performance level of the SWD. Also, the students took the test in both conditions, thereupon acting as their own controls. The finding that as a group, SWD did not have higher scores in either the accommodated or in the standard test condition calls into question the effectualness of the accommodation; the fact that they did not benefit more from the accommodation than SWOD calls into question its validity. Regarding

efficacy, it may be the case that as suggested by Goldman et al. (1998), older students with low reading skills are less likely than younger ones to benefit from producing an overt phonological representation of the text. Although reading aloud may enhance retention of exchange in working memory, it also slows down reading fluency. For older readers, the trade-off of increased retention versus slower processing of the text—particularly for longer passages—may not be sufficiently promising to result in overall gains in comprehension. Regarding validity, the findings from the study raised the question of whether the scores achieved in the accommodated condition can be interpreted in the same way as those obtained in the unaccommodated condition. Tindal (2002) described perspectives on the validity of accommodations. For example, Phillips (1994) specified five conditions for an appropriate definition of accommodations, including that the meaning of scores should be the same, regardless of any changes being made in the way the test is given or administered, and that the accommodation should not have the potential for benefit for SWOD (for further discussion, see Elliott et al., 2002).

The insinuation is that a necessary condition for test validity under accommodated conditions is that SWOD do not receive help from the accommodation (cf. Tindal et al., 1998). In the case of Fuchs et al. (2000), the student reads-aloud accommodation was concluded to be valid because although the gain for students with LD in the student reads-aloud condition was very small (ES = 0.06), SWOD suffered a detriment in performance (ES = –0.12), which resulted in a significant differential accommodation boost.

Moving beyond mean group differences in accommodation boost, alternative evidence for the impact of accommodation in this study was investigated. The lower

correlation of scores across testing accommodations was interpreted and compared to the correlation across alternate test forms in the standard condition to show that the accommodation did have an impact. Of potential interest is the fact that SWOD showed greater separation of the accommodation boost than general education students, such that the SWD were twice as likely to be substantially impacted by the accommodation.

In effect, the performance of SWD was more likely to be apprehensive in one direction or the other because of the accommodation. This is in line with evidence regarding the considerable heterogeneity of SWD (Morris et al., 1998). This finding is also consistent with those of McKevitt and Elliot (2003, as cited in Elliot et al., 2002). McKevitt and Elliott studied the impact of teacher-recommended accommodations on the performance of students with and without disabilities on a reading test. They found considerable unpredictability in the accommodation effects. Accommodations positively affected the scores for half of all SWD and 38% of all SWOD. There is no report of any negative impact on students' performance due to the accommodations. Hence, it is important to note that the initial discussion of validity is predicated on the assumption that accommodations have the potential to alter the construct being tested and that it must be demonstrated that they do not do so. Regarding the accommodation investigated in this study—having the student read reading passages aloud—it could be asserted that the construct of reading is indifferent to whether a text is read silently or aloud. Although a full discussion on the controversy surrounding models of skilled reading would go beyond the purpose of this study, comments are relevant to an understanding of the accommodation being investigated. Proponents of a strong phonological theory of visual word recognition (e.g., Frost, 1998) present persuasive arguments against the argument

that skilled performance in reading involves bypassing the mechanisms that convert orthographic structures into phonological structures. A more penurious explanation of skilled reading is that with practice, the reader's efficiency in computing a prelexical phonological representation increases; that the reader also acquires greater efficiency in accessing the lexicon with impecunious phonological information. In this view, beginner readers or older readers who have not gotten a high level of efficiency in reading must undertake a detailed phonological analysis of the printed word before lexical access, and hence, comprehension is achieved. Poor readers, or at least those whose primary difficulty lies in phonological decoding, may receive help from reading aloud because it helps them arrive at the more complete phonological analysis, they need to achieve lexical access. If the construct of reading comprehension is not yielding to acts of silent reading, then we must reevaluate both the supposition of the accommodation and the experimental findings.

The benefit of the accommodation—as in the study by Goldman et al. (1980)—may extend to any lower skilled reader whether the lower level of skill is due to a specific disability. Indeed, from this perspective, the accommodation is only an accommodation to conventional test administration practices, which, for the sake of efficiency, typically involve the testing of large groups of students in the same room at the same time. In this regard, there is no theoretical reason for which all students should not have the option of reading a reading test aloud. Certainly, some students are allowed to read to themselves if they do not do so (very) loudly. The finding that SWOD responded similarly to SWD, although less extremely, can be viewed as supporting the use of the accommodation with all students who might receive help from it, with no regard of disability status.

The study findings underscore the importance of taking an individual perspective on testing accommodations and of necessitating that accommodation decisions be based on trials adopted by each student. In most test accommodation studies, students' scores either increased because of the accommodation or remained unchanged. In this study, however, students suffered a potential harm (impaired test performance) because of the accommodation. The potential for harm makes it essential that great caution be applied in assigning this accommodation, and that it be assigned only based on prior evidence of benefit to the individual student.

Regarding students' feelings, previous research indicated that students are generally well responsive to testing accommodations. For example, McKevitt and Elliot (2003) found that students in their study had positive views of a tester-reads-aloud accommodation, although they expressed concern that having the test items read aloud made them difficult to follow. Elliott et al. (2002) reported on a dissertation study by Marquart (2000), which explored extended time on a mathematics test for eighth-grade students. The students in this study were surveyed concerning their feelings of the accommodation, and most reported that they felt more comfortable, were more motivated, thought they had performed better, and preferred taking the test with the extended time accommodation. Interestingly, neither the effect size for SWOD (ES = 0.34) nor that for SWD (ES = 0.26) was statistically different from zero.

Implications for Practice. Overall, the literature to date has shown weak effects for accommodations (Chiu & Pearson, 1999; Tindal, 2002). Nevertheless, a weak effect for a population of students (e.g., SWD) still does little to inform the selection of an accommodation for a particular student. That is, in the absence of any other information,

the mean accommodation boost associated with the assignment of an accommodation to a group of students supplies our best estimate of the impact for any individual student. However, prediction will improve tremendously if information is available on the students' prior experience with a particular accommodation, especially if the student was afforded multiple assessment opportunities both in the classroom and in test situations using a variety of accommodations (cf. Helwig & Tindal, 2003; Tindal et al., 1998). A caution in the implementation of test accommodations, especially on high-stakes tests, is that accommodations are not a remedy for low levels of skill on the construct that is being assessed. The large amount of attention being paid to providing SWD with proper accommodations may suggest to some students and their families that the "right" combination of accommodations will result in students achieving an adequate level of performance on a test. As pointed out by Elliott et al. (2002), an accommodation may in fact remove a disability-related barrier for the student assessed and still not have a significant effect on scores.

In conclusion, this study adds another piece to the experimental literature on testing accommodations for SWD. The findings from the study are congruous with those from previous research in suggesting that the challenge of assigning the most effective and appropriate testing accommodations for SWD (like that of designing the most effective and appropriate instructional programs for these students) is unlikely to be successfully addressed by doctrines affecting entire populations of students defined by their category of disability. Instead, much more attention will need to be paid to individual students' characteristics and responses to accommodations in relation to types of tests and testing situations.

Study 3

In the study *Consequences of Using Testing Accommodations: Student, Teacher, and Parent Perceptions of and Reactions to Testing Accommodations; Sylvia C. Lang, Patrick J. Kumke, Corey E. Ray, Erin L. Cowell, Stephen N. Elliott, Thomas R. Kratochwill & Daniel M. Bolt Wisconsin Center for Education Research and Department of Educational Psychology University of Wisconsin — Madison*, the researchers examined student, parent, and teacher perceptions of the use of testing accommodations as well as the relationship between student perceptions of testing accommodations and their disability status and grade level. Students with and without disabilities completed math and reading achievement tests with and without accommodations. The students, parents, and teachers completed a questionnaire to share their views on the use of testing accommodations.

Many SWD require testing accommodations to fully take part in large-scale achievement tests. Although a small percentage of children and adolescents with disabilities may be eligible for alternate assessments, most students who receive special education services can take large-scale achievement tests with accommodations. Testing accommodations are designed to increase the validity of the inference made about a test score and are intended to impact the skills needed to access a test, not the skills targeted for measurement by the test (Elliott, Braden & White, 2001). Accommodations can take different forms, including changes to the setting, test presentation, response format, and timing.

Both researchers and practitioners are interested in understanding the effects of accommodations on students' test scores. Such information is valuable not only for

making testing decisions for students in special education but also for understanding the consequences of testing accommodations for these students. The purpose of this study was to name several consequences of testing accommodations for students by examining student, teacher, and parent reactions to testing accommodations on large-scale achievement tests for students with and without disabilities.

Reactions to Testing Accommodations

Student and teacher reactions to the use of testing accommodations has been a topic of research over the years. For example, Elliott and Marquardt (2004) documented the positive reactions of eighth-grade students to an accommodation of extra time on a math test, and students at risk academically, SWD, SWOD, as groups, all preferred the accommodated test condition. Most students denoted that the testing accommodation allowed them to feel more comfortable, interested, and motivated—and less frustrated—while working on the test. They also reported that the test seemed easier, and that they performed better when extra time was provided.

In a similarly designed study, McKevitt and Elliott (2003) examined the effects of read-aloud and teacher-recommended accommodations on eighth-grade students. The SWD preferred taking the test under the accommodated conditions, while the SWOD thought that the accommodated and non-accommodated conditions were the same. However, SWOD were more likely than SWD to report that they thought they did better on the test with accommodations. Conversely, SWD reported that they believed it was easier to show what they knew under the accommodated test condition and that they felt more comfortable taking the test with accommodations.

McKevitt and Elliott (2003) also examined teacher feelings of testing accommodations and found overall mixed attitudes. For instance, the teachers reported that testing accommodations were fair for SWD but felt that standardized reading tests, with or without accommodations, did a poor job of supplying an opportunity for SWD to demonstrate their knowledge. Thus, teachers characterized testing accommodations as only "somewhat helpful" in helping SWD to demonstrate their knowledge on standardized reading tests. Nonetheless, the teachers most strongly endorsed the belief that accommodations helped SWD feel more positive about taking tests.

In a national study among 401 general education teachers, Jayanthi, Epstein, Polloway, and Bursuck (1996) found that teachers, either alone or with a special educator, oversaw making decisions about testing accommodations for their students. Although teachers oversaw identifying and applying testing accommodations, they reported that many of the most useful testing accommodations were not practical for implementation in the classroom. Elementary school teachers were more likely than middle and high school teachers to report that accommodations were relatively easy to implement. Most teachers (67%) in this study reported that providing testing accommodations only for SWD was unfair, citing two primary reasons: First, most teachers (78%) who thought it was unfair to provide testing accommodations only for SWD reported that all students need some accommodations, regardless of whether they receive special education services. Second, some teachers showed that if SWD are included in the general education class, they must adhere to the standards of that class and not be given accommodations that are not part of the general education curriculum (Jayanthi et al., 1996).

In another study on teacher beliefs and use of testing accommodations, Gairia, Salend, and Hemrick (1994) surveyed middle and high school general education teachers. The results showed that although the teachers were familiar with many diverse types of testing accommodations, they were most likely to use those that could be provided to all students and that were not tailored to individual needs. Approximately one-third of the testing accommodations were rated significantly higher for perceived effectiveness than for use. Based on their findings, Gajria et al. suggested that the use of testing accommodations is influenced not only by presumed effectiveness but also by the feasibility of implementation. Similarly, in a study examining teacher perceptions of the desirability and feasibility of 30 classroom-based testing accommodations, Schumm and Vaughn (1991) found that testing accommodations requiring little individualization were rated as more desirable and feasible, and that accommodations related to curriculum use or evaluation procedures were rated as least viable. Finally, classroom-based accommodations related to increasing students' social and emotional adjustment and not requiring any environmental or curricular adaptations were rated as the most desirable.

Consequential Validity of Testing Accommodations

Student, teacher, and parent feelings of testing accommodations as well as the effects of testing accommodations on these groups are the primary components of the consequential validity of the use of testing accommodations. Substantial validity refers to the influence or effects that an assessment or intervention has on its consumers. It is as much a matter of social policy as it is a major part of the overall validity of an assessment system (American Educational Research Association, 1999). As Messick (1996) emphasized, the meaningful aspect of validity should not be viewed as a separate type of

validity in that the values served in the intended and unintended outcomes of test analysis use both to interpret from and contribute to the meaning of the test scores. This aspect of validity affirms the need to collect evidence about the intended benefits and potential unintended negative consequences of supplying testing accommodations to SWD. Currently, the consequential validity of the use of testing accommodations remains unclear.

Past research in this area has primarily examined teachers' (rather than students' or parents') feelings of the use of testing accommodations. For several reasons, gaining a broader understanding of student perspectives is important. First, students are the consumers of accommodations provided within the classroom or on large-scale assessments. Therefore, it is important to consider how such accommodations affect their motivation. Some students—especially those in middle and high school—may be motivated to perform when accommodations are provided. Second, it is important to consider the acceptability of diverse types of accommodations to students. Whereas some accommodations may be viewed as helpful and welcomed, others may be regarded negatively. In a study among middle and high school students, Vaughn, Schumm, Niarbos, and Daugherty (1993) found that students appreciated teachers who supplied accommodations and leaned toward certain types of accommodations to others. Students preferred classroom accommodations such as adjusted instruction by teachers who were sensitive to individual needs and diverse learning styles (i.e., teachers who changed presentation formats, met with students individually, and changed workgroups). They reported less positive discernment of accommodations on tests, homework, materials, or

textbooks. Surprisingly, high-achieving students were more likely to prefer accommodations than were low-achieving students.

Vaughn, Schumm, Niarbos, and Gordon (1993) found that elementary students also showed a preference for teachers who provided instructional accommodations, except for accommodations provided on tests, homework, and textbooks. As in the Vaughn, Schumm, Niarhos, and Daugherty (1993) study, high-achieving students were more likely to prefer accommodations than were low-achieving students. Extended interviews supplied insight into these group differences. For example, responses from students in the high-achieving group revealed an awareness of different learning needs and the requisite of accommodations for students. In contrast, students in the low-achieving group expressed a desire to "fit in" and can discuss assignments with classmates and discover what their classmates were being taught. Based on these findings, Vaughn, Schumm, Niarhos, and Gordon (1993) suggested that accommodations might draw unwanted attention to students who are low achieving or otherwise having difficulty in the classroom. Furthermore, accommodations may transfer more of the responsibility for learning to the student, a responsibility that students may not want to take upon themselves.

Vaughn and Schumm (1993) also examined the perspectives of students with learning disabilities on accommodations. As expected, students with learning disabilities showed an overall preference for a teacher who supplied accommodations. Being that elementary and high school students with learning disabilities showed no significant preference for or against accommodations on tests, homework, or textbooks, middle school students reported a preference for accommodations on homework and textbooks,

but not on tests. Again, accommodations in teacher instructional style were preferred by all students. Vaughn and Schumm had predicted that the preferences of students with learning disabilities would be more like those of the low-achieving students than the high-achieving students. Instead, their results uncovered that high- and low achieving students across all grade levels were more like each other in their preferences for accommodations than were students with learning disabilities to students in either group without disabilities. The authors attributed these findings to the earlier accommodation experiences of students with learning disabilities. These students may have been familiarized with receiving accommodations on assignments or tests; therefore, they viewed such accommodations—although not preferred in the general education classroom—more positively.

The results from these studies highlight important themes. First, teachers appear to be primarily responsible for identifying, developing, and implementing accommodations for SWD in their classrooms. Second, teachers may perceive accommodations as beneficial or desirable, but not workable for implementation, given realistic constraints. Research suggests that these constraints often lead to questionable practices in school-wide testing of SWD. Third, students' preferences for certain accommodations differ across ability groups and grade levels. Fourth, student perspectives on acceptable accommodations may not align with teacher perspectives of viable accommodations. These themes underscore the importance of gathering more information about the alignment of teacher and student perspectives on accommodations recommended for use during large-scale assessments.

Discussion

Purpose of the Study. This study examined student, parent, and teacher perspectives of testing accommodations for standardized tests. Student beliefs of testing accommodations were examined in relation to (a) their comfort level while taking the test, (b) their ability to demonstrate knowledge under both accommodated and non-accommodated conditions, (c) the ease of the test, and (d) their preference for the accommodated or non-accommodated test condition. Parent and teacher beliefs were examined based on (a) the fairness of granting accommodations solely to SWD on state and districtwide tests and (b) the comparability of scores achieved by SWD who received accommodations and SWOD who received no accommodations. Finally, the differences in student feelings were inspected based on groupings of disability status and grade level. The research did not involve a high-stakes testing situation, although the IEP-based accommodations used in the study were like those used in statewide assessment in most states. A clear sense of educational or personal consequences for poor test performance is, therefore, missing from the study. It was hypothesized that this high-stakes variable would influence students', teachers', and parents' reactions to accommodations. It was determined that future investigations would be needed to better understand the influence of high stakes on student and consumer reactions to tests and testing accommodations.

Student Perceptions of Accommodations. The findings from this study indicate that a significant majority of students with and without disabilities felt more comfortable taking tests under the accommodated condition. However, there was no significant difference between the proportion of SWD who felt more comfortable taking tests with accommodations and the proportion of SWD who either felt more comfortable taking

tests without accommodations or who felt that the testing conditions were the same. In fact, the percentage of SWD who felt more comfortable taking the tests under the accommodated condition (46%) was comparable to the percentage of SWD who felt that the accommodated and non-accommodated conditions were about the same (44.4%). However, there was a significant difference between the proportion of SWD who felt more comfortable with accommodations and the proportion of SWD who felt more comfortable without accommodations. A slight majority of students with and without disabilities viewed the tests to be easier when accommodations were provided; however, more than a quarter of the students with and without disabilities claimed that the test conditions seemed the same. The proportion of SWD who felt that the tests were easier with accommodations was significantly greater than the proportion of SWD who either felt the tests were easier without accommodations or felt the conditions were the same. The SWD also showed an overall preference for the accommodated testing condition compared to the non-accommodated condition. However, when the SWD who indicated that the testing conditions seemed about the same were grouped with those who indicated a preference for the non-accommodated condition, the difference between this combined group and the SWD who preferred the accommodated condition was not significant.

Students disliked the accommodations because they felt they were unnecessary and slowed down the testing process. Yet other students reported a liking for one of the testing conditions based on how the accommodations affected their self-perception or innate test taking abilities. Specifically, the accommodated test condition made students feel more capable and on level academically with their peers. Such findings about the effects accommodations can have on students' self-perception are important to consider

when making decisions about supplying testing accommodations. Thus, based on the responses of students in this study, it appears that testing accommodations can be helpful or harmful for a student's sense of capability.

Parent and Teacher Perceptions of Accommodations. Both parents and teachers regarded testing accommodations to be fair for SWD, although teacher ratings of fairness were slightly higher. Most parent participants were parents of SWD, while the teacher participants included both general and special educators. With the inclusion of SWD in general education classrooms, it can be expected that general education teachers have had experiences educating or working with SWD. Hence, we predicted that parents and teachers would view testing accommodations as fair given their experiences with children who have diverse learning needs. Parents and teachers agreed that the score of a student with disabilities who received accommodations would be somewhat identical to the score of a student without a disability who received no accommodations on the same test. On the assumption that SWD necessitates testing accommodations to disclose the opportunity to communicate their knowledge, it was expected that the scores of SWD who received accommodations would be resemblant to the scores of SWOD who did not receive accommodations. This study proves that parents and teachers viewed the scores from accommodated tests for SWD as somewhat valid, as reflected by their ratings of these test scores being somewhat comparable to those of a test without accommodations for SWOD. Unfortunately, parent and teacher questionnaires did not query about the desirability or feasibility of identifying, designing, or implementing testing accommodations for SWD, an investigation that would have been useful given previous research (Schumm & Vaughn, 1991).

Group Differences. The exploration of group differences revealed mixed findings. As predicted, a significantly greater proportion of SWD than SWOD felt the adapted condition was easier and preferred that condition on the math and reading tests. However, contrary to the study's prediction that students in fourth grade would prefer testing accommodations more than students in eighth grade, there was no significant association found between grade level and partiality for either test condition. These findings were unexpected given that researchers found that SWD reported less preference for testing accommodations at higher grade levels than at lower ones (Elliott & Marquardt, 2004; McKevitt & Elliott, 2003).

Conclusion

This study found that students, parents, and teachers have multifarious perceptions of the requirement of testing accommodations for SWD. Most students indicated a slight preference for testing accommodations, and most parents and teachers reported feeling that testing accommodations were fair and valid for SWD. An examination of open-ended responses on the student questionnaire revealed different ways in which testing accommodations can sway a student's self-efficacy. Whereas some students viewed the testing accommodations as contributing to their self-assuredness during testing, others reacted negatively to the accommodations, signaling that they made them feel less than academically prepared as their peers. These findings attest to further research on the effects that testing accommodations can have on students' self-efficacy, particularly since many SWD are likely to be exposed to testing accommodations during mandated tests. Taken together, this research on student, teacher, and parent reactions to

the use of testing accommodations provides evidence for the consequential validity of the resulting scores for all students.

The three studies selected for this meta-analysis will provide insight into the validity of providing the read aloud as an accommodation for SWD, and SWOD. By combining the results of these studies, I can construct associations between theoretically important variables such as, in what types of testing conditions does the read aloud increase scores for both SWD, and SWOD? How is the accommodation viewed by students, parents, and teachers, and how does the accommodation impact student self-efficacy?

Both SWD, and SWOD have been shown to receive help from the read-aloud accommodations with the accommodation effect size for SWD being significantly larger than the effect size for SWOD. Li (2014) revealed in his study the important factors that influence the effects of read-aloud accommodations. For example, the accommodation effect was significantly stronger when the subject area was reading than when the subject area was math. The study also shed light on the effect of read-aloud accommodations being significantly stronger when the test was read by human proctors than when read by video/audio players or computers. As with all the studies being referenced, the implications, limitations, and directions for future research were discussed.

CHAPTER 4 RESULTS

As noted, the three studies selected for this meta-analysis intend to yield insight into the validity of providing the read aloud as an accommodation for SWD and SWOD. By joining the results of these studies, I can build upon the associations between theoretically important variables such as the types of testing conditions and if they do in fact increase scores for both SWD and SWOD.

Li's (2014, online first) *The effects of the read-aloud accommodations for students with and without disabilities: A meta-analysis. Educational Measurement: Issues and Practices* is a meta-analysis study that proposed the read aloud accommodation to help remove obstacles faced by SWD in reading comprehension. Li (2014) conducted his meta-analysis to examine the effects of read-aloud accommodations for students with and without disabilities. Typically, both SWD and SWOD improved their test scores from the read-aloud accommodations, albeit the accommodation effect size for SWD was significantly larger than the effect size for SWOD.

This meta-analysis revealed key factors that influence the effects of read-aloud accommodations. For example, the accommodation effect was significantly stronger when the subject area was reading than when the subject area was math. The effect of read-aloud accommodations was also significantly more predominant when the test was read by human proctors than when it was read by video/audio players or computers. Ultimately, the implications, limitations, and directions for future research were discussed throughout the study (Li, 2014).

Data Analysis

Coding Procedures

First, disability status was coded as "with disabilities" or "without disabilities" based on information provided in the studies. Specifically, in all the studies included in this meta-analysis, SWD mostly were identified as having learning disabilities. Second, subject areas were coded as either math or reading. Third, the methods whereby the read-aloud accommodations were delivered were coded according to three categories: when the test was read by a teacher or a test administrator, it was coded as "read by human proctors"; when read by a video tape player (Crawford & Tindal, 2004) or a cassette player (Harker & Feldt, 1993), it was coded as "read by video/audio players"; and when read via a computer, it was coded as "read by computers."

Third, the coding of extra time was less straightforward than the coding for the other variables. Studies in which extra time was deliberately combined with the read-aloud accommodation so that a package of accommodations were provided (e.g., Schulte, Elliott & Kratochwill, 2001) were excluded from this meta-analysis. Li included only studies in which the read-aloud was offered as a single accommodation strategy and extra time was inevitably allowed due to practical reasons relating to delivering the read-aloud accommodation (Olson & Dirir, 2010). When it was specifically said that the read-aloud accommodated condition allowed more time than the standard condition, this variable was coded as "yes"; otherwise, Li coded it as "no." In a few instances, Li contacted the authors to collect sufficient information to code this variable.

In addition, Li coded the research design of each study. In a study with an independent group design, typically students were randomly assigned to either the

accommodation or the control condition. Effect size was calculated as the standardized mean difference between the two groups using the pooled standard deviation of the two groups (raw-score effect size) (Morris & DeShon, 2002). However, many of the read-aloud accommodation studies that Li included used a repeated measure design (i.e., each student took the test under both conditions, with read aloud accommodations and without). Typically, the design was counterbalanced to minimize the order effects. For a study using the repeated-measure design, the correlation between the pre-test and post-test scores was needed to calculate the effect size (change-score effect size) (Morris & DeShon, 2002). However, many of the studies using the repeated measure 15 design did not report this correlation, so Li was unable to calculate the change-score effect size. Li, therefore, calculated the raw-score effect size for both the independent group design and the repeated measure design studies. This practice was also adopted in other test accommodation meta-analysis studies, such as Gregg and Nelson (2012) and Kieffer, Rivera, and Francis (2012). To adjust this artifact due to different research designs, Li further coded research design as a dichotomous variable (repeated measure design or independent group design) and used this variable as one of the level-2 predictors in the subsequent analysis (Briggs, Ruiz-Primo, Furtak, Shepard & Yin, 2012; Hox, 2012).

Individual Effect Sizes

The meta-analysis was performed following the variance known HLM approach with the HLM 6.08 software (Raudenbush, Bryk & Congdon, 2004). As demonstrated in Raudenbush and Bryk (2002), dj, the effect size estimate for comparison j, is the standardized mean difference between an experimental group and a control group. It is

defined as dj = (Ej - Cj) /Sj [1] where Ej is the mean outcome for the experimental group; Cj is the mean outcome for the control group; and Sj is the pooled within-group standard deviation. 16 According to Hedges (1981), dj can be viewed as a statistic to estimate the corresponding population effect size. It is approximately unbiased and normally distributed with variance Vj = (nEj + nCj) / (nEjnCj) + δj 2 / [2(nEj + nCj)] [2] where δj is the corresponding population effect size; nEj is the experimental-group sample size; and nCj is the control-group sample size. The observed dj is used to substitute for δj in equation 2, and Vj is assumed to be known. When there are at least 20 (Hedges & Olkin, 1985) or 30 (Raudenbush & Bryk, 2002) cases per study, it is reasonable to assume that the variance Vj can be estimated with sufficient accuracy (Hox, 2010). Hedges (1981) also presented a correction for bias in the calculation of effect sizes when the sample size of the experimental group, nEj or that of the control group, nCj , is very small: Adjusted dj = dj (1 – [3] With the variance-known HLM approach, the level-1 outcome variable in the meta-analysis is the effect size reported for each comparison (Li, 2014).

Table 2

Variables and Frequencies

Variable	Coding and Notation	Frequency
Disability status	Students without disabilities (Reference group)	60
	Students with disabilities (W_1)	54
Subject area	Reading (Reference group)	26
	Math (W_2)	88
Delivery method	Read by human proctors (Reference group)	41
	Read by computers (W_3)	21
	Read by video/audio players (W_4)	52
Grade level	Elementary school (< 6^{th} grade) (Reference group)	51
	Middle school (6^{th} to 8^{th}) (W_5)	46
	High school (9^{th} to 12^{th}) (W_6)	17
Extra time	Accommodated condition does not allow more time than the non-accommodated condition (Reference group)	94
	Accommodation condition allows more time than the non-accommodated condition (W_7)	20
Research design	Repeated measure design (Reference group)	93
	Independent group design (W_8)	21

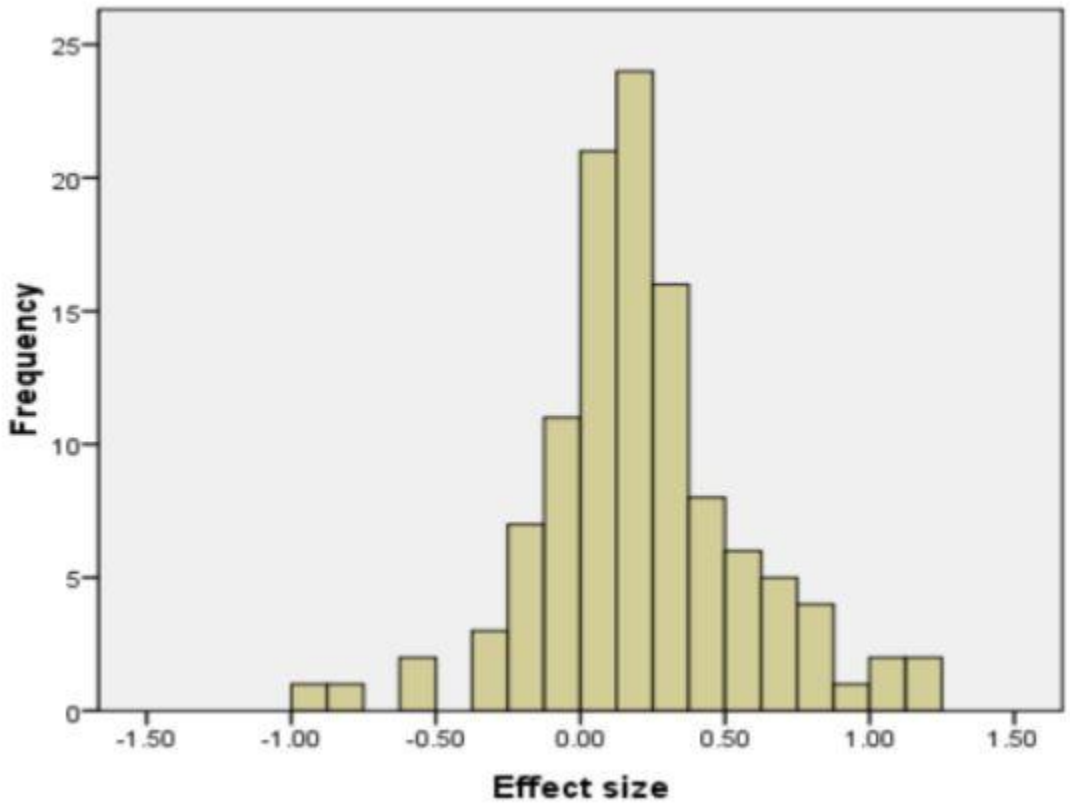
As described by Raudenbush and Bryk (2002), the level-1 model (often referred to as the unconditional model) is $dj = \delta j + ej$ [4] where δj is the true overall effect size across comparisons; and ej is the sampling error associated with dj as an estimate of δj . Here, we assume that $ej \sim N(0, Vj)$. 17 In the level-2 model, the true population effect size, δj , depends on comparison characteristics and a level-2 random error term: $\delta j = \gamma 0 + \gamma 1 W1j + \gamma 2 W2j + \ldots + \gamma 6 W6j + \gamma 7 W7j + \gamma 8 W8j + \mu j$ [5] where W1j …W8j are the comparison characteristics predicting δj (see Table 1 for the list of variables and the corresponding frequencies); $\gamma 0$ is the expected overall effect size when each Wij is zero; $\gamma 1 \ldots \gamma 8$ are regression coefficients associated with the comparison characteristics W1 to W8 ; and μj is a level-2 random error (Li, 2014).

Results

Based on the procedure described in the methods section, 114 effect sizes from 23 studies were included in this meta-analysis. The distribution of the effect sizes is illustrated in Figure 1. More of the effect sizes were positive than negative, and no outliers were detected. The effect sizes ranged from -.95 to 1.20, with a mean of .20 and a standard deviation of .36. The effect sizes were approximately normally distributed with a skewness of .23 and a kurtosis of 1.51 (Li, 2014).

Figure 1

Histogram of Effect Sizes



The predictors were entered into the model by category separately and then in a combined way. Table 2 summarizes the estimated regression coefficients, the 95% confidence intervals, and the random components of a series of models. Due to space

limitations, confidence intervals were not referred to in the later sections. Model 0 shows the results when no 18 predictors were included. The intercept (i.e., the estimated grand-mean effect size) was .20, which was statistically different from zero ($t\ [113] = 6.70$, $p < .001$). This result shows that on average students who received read-aloud accommodations scored about .20 standard deviation units higher than their non-accommodated peers. Furthermore, the estimated variance of the effect size was .06, which was significantly different from zero. This suggests that variability existed in the true effect sizes across comparisons. Therefore, the results show that analysis should proceed to a level-2 conditional model in order to determine which characteristics explain this variability (Li, 2014).

Table 3

Results of the Models

	Model 0	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7
Fixed effect								
Intercept	.20*** (.14, .26)[a]	.14*** (.07, .22)	.41*** (.30, .52)	.34*** (.25, .43)	.24*** (.15, .32)	.17*** (.10, .23)	.17*** (.10, .23)	.48*** (.33, .64)
Disability		.13* (.01, .24)						.13* (.03, .23)
Math			-.27*** (-.40, -.15)					-.26*** (-.38, -.14)
Computer				-.23** (-.38, -.07)				-.22** (-.36, -.07)
Video/audio player				-.22** (-.35, -.10)				-.20** (-.33, -.06)
Middle school					-.10 (-.23, .03)			-.15* (-.26, -.03)
High school					.03 (-.14, .21)			-.02 (-.18, .14)
Extra time						.21* (.05, .36)		.13 (-.01, .27)
Independent group design							.19* (.04, .34)	.08 (-.09, .24)
Random effect								
Standard deviation	.25387	.24696	.23165	.23206	.25247	.25008	.24635	.18431
Variance Component	.06445[b]	.06099	.05366	.05385	.06374	.06254	.06069	.03397
Degree of freedom	113	112	112	111	111	112	112	105
Chi-square value	431.17	392.30	393.62	385.11	420.00	470.22	415.16	260.77
Proportion of variance explained	N/A	5.37%	16.74%	16.45%	1.10%	2.96%	5.83%	47.29%

Note. * p < .05, ** p < .01, *** p < .001; a Numbers in the parentheses are the 95% confidence interval; b The variance component is significant at .001 level across all models.

Results of the Models

In Model 1, disability status was statistically significant ($\gamma = .13$, $t[112] = 2.13$, $p < .05$). SWOD who received read-aloud accommodations scored about .14 standard deviation units higher than their non-accommodated peers, whereas SWD who received read-aloud accommodations scored about .27 (i.e., .14 +.13) standard deviation units higher than their non-accommodated peers. In Model 2, the accommodation effect size for math tests was significantly smaller than that for reading tests ($\gamma = -.27$, $t[112] = -4.18$, $p < .001$). Specifically, students who received a read-aloud accommodation on reading tests scored about .41 standard deviation units higher than their non-

accommodated peers; however, the increase was only .14 (i.e., 41 - .27) standard deviation units for math tests. In Model 3, both two variables related to accommodation delivery methods were statistically significant. When a human proctor read the test, students who received read-aloud accommodations scored about .34 standard deviation units higher than their non-accommodated peers, whereas the increase was .11 (i.e., .34 - .23) standard deviation units when the read-aloud was delivered by a computer and .12 19 (i.e., .34 - .22) standard deviation units when the read-aloud was delivered by a video/audio player. In Model 4, for middle school students and for high school students, the effect of read aloud accommodations was not significantly different from that for elementary school students (Li, 2014).

In Model 5, compared to the effect size when extra time was not evidently provided in the accommodated condition, the effect size when extra time was provided was significantly larger by .21 standard deviation units. In Model 6, compared to the effect size in studies with a repeated measure design, the effect size in studies with an independent group design was significantly larger by .19 standard deviation units. In Model 7, all the predictors at one time were entered to investigate their effects simultaneously. As shown in Table 2, the regression coefficient related to disability status was significantly positive in both Models 1 and 7. This result indicates that whether we controlled for other predictors, the effect size of receiving read-aloud accommodations was larger for SWD than for those without disabilities. The effects of subject areas and accommodation delivery methods were also consistent whether other predictors were included in the model. There were, however, minor variations regarding the other predictors across models. The difference between middle schools and elementary schools

became statistically significant in Model 7. Extra time and research design, however, became statistically non-significant in Model 7. These minor variations indicate a potential interaction among the predictors, although this interaction is considered slight (Li, 2014).

In addition to the regression coefficient, the proportion of variance explained was also calculated with Model 0 as the baseline model (Raudenbush & Bryk, 2002). As shown in the last row of Table 2, subject area and accommodation delivery method each explained over 16% of the variance in effect sizes, followed by research design (5.8%), disability status (5.4%), extra time (3.0%), and grade level (1.1%). In Model 7, when all the predictors were included, the proportion of variance explained was 47.3%. Still, the estimated variance of the effect sizes in this model was .034, which was significantly different from zero ($\chi 2 = 260.77$, $df = 105$, $p < .001$). This indicates that unknown sources of variability still exist among the effect sizes beyond what have been accounted for in this meta-analysis (Li, 2014).

Figure 2 stands for the estimated effect sizes in Model 7 when variables were controlled for grade level, extra time, and research design. For example, when the subject area was reading, for SWOD, the estimated effect sizes were as follows: .48 when the test was read by a human proctor, .26 (i.e., 48 - .22) when read by a computer, and .28 (i.e., .48 - .20) when read by a video/audio player. The estimated effect size for math was calculated in an equivalent way. As shown in Figure 2, the estimated accommodation effects varied substantially across combinations of disability status, subject area, and accommodation delivery method (Li, 2014).

Few studies contributed effect sizes to this meta-analysis. The mean of the eight effect sizes in Calhoon et al. (2000) was .24, and the mean of the 16 effect sizes in Olson and Dirir (2010) was .18, both of which were close to the overall mean of .20. Also, both studies involved multiple accommodation delivery methods and or multiple subject areas. However, the mean of the 16 effect sizes in Helwig and Tindal (2003) was only .03, and the mean of the 14 effect sizes in Helwig et al. (2012) was .00. These two studies focused on read aloud.

Figure 2

Estimated Effect Sizes

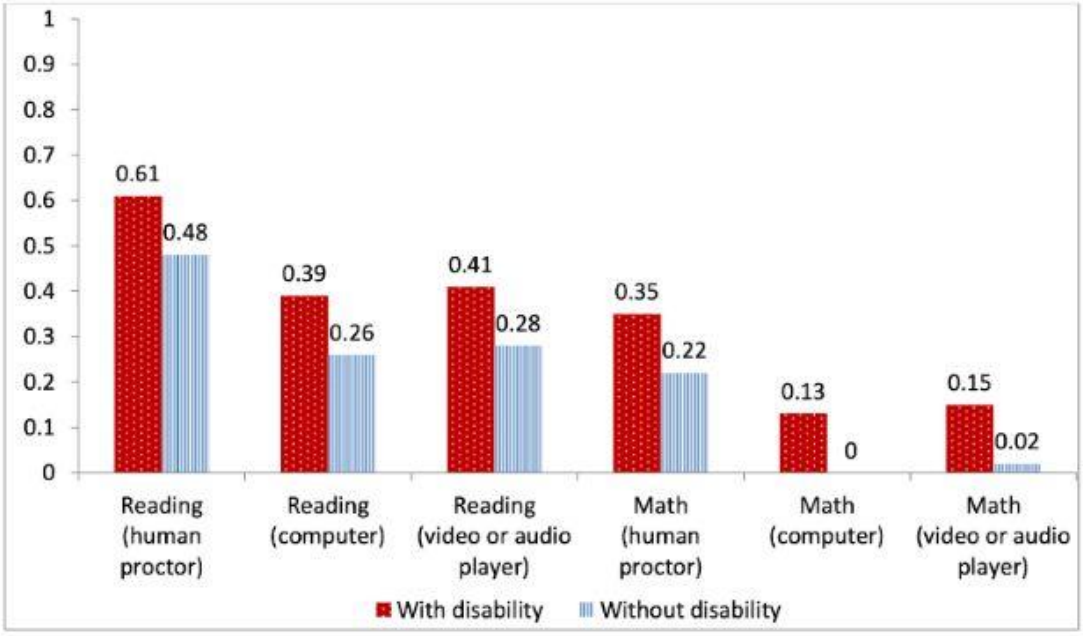


Figure 2. Estimated effect sizes.

Note. The effect sizes are based on Model 7, where reference groups are elementary school, no extra time, and repeated-measure design.

Few studies contributed many effect sizes to this meta-analysis. The mean of the eight effect sizes in Calhoon et al. (2000) was .24, and the mean of the 16 effect sizes in Olson and Dirir (2010) was .18, both of which were close to the overall mean of .20.

Also, both studies involved multiple accommodation delivery methods and or multiple subject areas. However, the mean of the 16 effect sizes in Helwig and Tindal (2003) was only .03, and the mean of the 14 effect sizes in Helwig et al. (2012) was .00. These two studies focused on read aloud accommodations for math tests delivered via video/audio players. In addition, a sensitivity analysis was performed by removing the effect sizes produced in each of the four studies one at a time. The only changes observed were as follows: (1) extra time became significant in Model 7 ($\gamma = .14$, $t(89) = 2.04$, $p < .05$) when removed Helwig and Tindal (2003), and (2) middle school became nonsignificant in Model 7 ($\gamma = -.10$, $t(89) = -1.38$, $p > .05$) when removed Olson and Dirir (2010). Admittedly, one study contributing a large number of effect sizes may create a dependence issue and bias the standard errors. However, the sensitivity analysis shows that conclusions from the study did not change much as a result of including those studies in this meta-analysis (Li, 2014).

Table 4

Distribution of All Participating Students by Grade Grouping, and Gender

	Students With LD		*Students Without LD*	
	Male	*Female*	*Male*	*Female*
Middle school (Grades 6–8)	114	59	45	51
High school (Grades 9–10)	68	42	49	28
Total	182	101	94	79

Note. LD = learning disabilities.

Table 5

Test Performance of Students With and Without Learning Disabilities in Each Test Condition

	Students With LD[a]		*Students Without LD*[b]	
Condition	*M*	*SD*	*M*	*SD*
Standard	9.19	3.36	13.20	2.30
Student reads aloud	9.21	3.50	12.62	2.17

Note. LD = learning disabilities.
[a]$n = 283$. [b]$n = 173$.

Table 6

Distribution of Students in Main Analyses by Grade Grouping, and Gender

	Students With LD		*Students Without LD*	
	Male	*Female*	*Male*	*Female*
Middle school (Grades 6–8)	91	47	16	24
High school (Grades 9–10)	54	38	23	18
Total	145	85	39	42

Note. LD = learning disabilities.

Table 7

Test Performance of Students With and Without Learning Disabilities in Each Test Condition

	Students With LD[a]		*Students Without LD*[b]	
Condition	*M*	*SD*	*M*	*SD*
Standard	9.27	2.45	9.16	3.25
Student reads aloud	11.52	1.70	11.86	2.39

Note. Students whose data are reported in this table were those who achieved scores between 5 and 13 (out of a maximum score of 16) in the standard test condition. LD = learning disabilities.
[a]$n = 230$. [b]$n = 81$.

Table 8

Cross-Tabulation of Accommodation Benefits by Disability Status

	Accommodation Benefit			
Disability Status	*Benefit*	*No Difference*	*Detriment*	*Total*
Students with LD				
Count	40	144	46	230
% within disability status	17.4	62.6	20.0	100.0
% within accommodation benefit	83.3	69.2	83.6	74.0
% of total	12.9	46.3	14.8	74.0
Students without LD				
Count	8	64	9	81
% within disability status	9.9	79.0	11.1	100.0
% within accommodation benefit	16.7	30.8	16.4	26.0
% of total	2.6	20.6	2.9	26.0
All students				
Count	48	208	55	311
% within disability status	15.4	66.9	17.7	100.0
% within accommodation benefit	100.0	100.0	100.0	100.0
% of total	15.4	66.9	17.7	100.0

Note. LD = learning disabilities.

Table 9

Percentages and Numbers of Students With Disabilities and Students without Disabilities and Preferred Test Conditions

Questionnaire item	Student responses: Preferred accommodations % (n)	Preferred no accommodations % (n)	Felt conditions were same % (n)
Comfort level			
SWD*	46% (58)	9.5% (12)	44.4% (56)
SWOD**	43.4% (72)	17.5% (29)	39.2% (65)
Seemed easier			
SWD	62.2% (79)	9.4% (12)	28.3% (36)
SWOD	50% (83)	16.9% (28)	33.1% (55)
Overall preference			
SWD	46% (59)	17% (22)	36% (46)
SWOD	39% (64)	21% (35)	40% (66)
Better able to show knowledge in math			
SWD	36.2% (46)	27.6% (35)	36.2% (46)
SWOD	25.5% (42)	40% (66)	34.5% (57)
Better able to show knowledge in reading			
SWD	42.5% (54)	21.3% (27)	36.2% (46)
SWOD	30.7% (51)	35.5% (59)	33.7% (56)

*The total number of SWD was 127. **The total number of SWOD was 166.

CHAPTER 5 DISCUSSION AND CONCLUSION

What Are the Effects of Read-Aloud Accommodations for Students with and Without Disabilities?

According to the differential boost framework, because of receiving accommodations, SWD are expected to obtain a larger increase in their scores compared to SWOD. The results of the Li (2014) meta-analysis study support the requirement of this framework. Disability status was shown as being a statistically significant predictor, signifying that the effect of read-aloud accommodations for SWD was significantly stronger than the effect for SWOD whether other predictors were controlled. The significant predictors of the accommodation effect sizes were .14 for SWOD and .27 for SWD. Results of the study did not differ primarily from previous metanalysis findings. For instance, Vanchu-Orosco (2012) reported an effect size of .30 for SWD and .17 for SWOD for multiple types of test accommodations. In Elbaum (2007), the mean effect size was .37 for elementary school SWD and .10 for secondary school SWD. At present, in the test accommodation literature for SWD, the categories of small, medium, and large effects are not clearly defined (Vanchu-Orosco, 2012). If a general scheme is used, as implied by Cohen (1992), a differential boost of .13 standard deviation units as found in the present meta-analysis is regarded as small in practical terms.

The interaction hypothesis says that students who require the accommodation should benefit from it and that students who do not need the accommodation should not benefit from it. Here, Li (2014) referred to the estimated effect sizes, when grade level, extra time, and research design were collected for. When the subject area was reading, regardless of the accommodation delivery method, SWD and SWOD both benefited from

receiving read-aloud accommodations, with effect sizes ranging from .26 to .61. When the subject area was math, SWD and SWOD both received help from read-aloud accommodations provided by human proctors, with effect sizes of .35 and .22, respectively. Nevertheless, for math tests, when the accommodation was provided by a computer or a video/audio player, the effect sizes for SWD were small and the effect sizes for SWOD were zero or almost zero. Therefore, the read-aloud accommodations did not always meet the criteria of the interaction hypothesis (Li, 2014).

Elbaum (2010), and colleagues examined the effects of a student-reads-aloud accommodation on the performance of middle school and high school students with and without LD on a test of reading comprehension. Test scores that students achieved in the accommodated condition were not statistically significantly different from scores obtained in the standard condition. For 17% of SWD, the accommodation boosted performance; for 20%, the accommodation impaired performance. Whereas 10% of SWOD showed an accommodation benefit, while 11% showed an accommodation detriment. Although an ANOVA revealed no statistically reliable difference in the accommodation boost for SWD and SWOD, the analysis of residualized gain scores suggested that at approximately equal levels of performance in reading comprehension, SWOD may, as a group, benefit more from this accommodation than SWD (Elbaum et al., 2004).

Elbaum's (Elbaum, 2010) brought recommendations in the literature to the forefront concerning the experimental investigation of testing accommodations. Tindal, Heath, Hollenbeck, Almond, and Harniss (1998) urged that "to provide the most convincing empirical support for an accommodation, students with a specific need have

to be compared to others without such a need who are otherwise comparable in achievement" (p. 442). In this study, a comparison group of general education students was chosen that was remarkably close to the reading performance level of SWD. As well, students took the test in both conditions, therefore acting as their own controls. The finding that as a group, SWD did not have higher scores in the accommodations than in the standard test condition calls into question the effectiveness of the accommodation; the fact that they did not benefit more from the accommodation than SWOD calls into question its credibility. Regarding accommodation effectiveness, it may be the case that as suggested by Goldman et al. (1980), older students with low reading skills are less likely than younger ones to benefit from producing a distinct phonological representation of the text. Although reading aloud may enhance retention of discourse in working memory, it also slows down reading fluency. For older readers, the trade-off of increased retention versus slower processing of the text, particularly for longer passages, may not be sufficiently beneficial to result in overall gains in comprehension. With regard to validity, the findings of the study raise the question of whether scores achieved in the accommodated condition can be explicated in the same way as scores obtained in the unaccommodated condition. Tindal (2002) described several points of view on the validity of accommodations. For example, Phillips (1994) specified five conditions for a suitable definition of accommodations including that the meaning of scores should be the same regardless of any changes being made in the fashion in which the test is given or taken and that the accommodation should not have the likelihood of benefit for SWOD. The insinuation is that a necessary condition for test validity under accommodated conditions is that SWOD do not have an unfair advantage from the accommodation (cf.

Tindal et al., 1998). In the case of Fuchs et al. (2000), the student-reads-aloud accommodation was concluded to be valid because although the gain for SWD in the student-reads-aloud condition was very small (ES = 0.06), SWOD suffered a decline in performance (ES = –0.12), resulting in a significant differential accommodation boost. (Li, 2014)

Elbaum's (20210) findings also emphasize the importance of taking an individual perspective on testing accommodations and of requiring that accommodation decisions be based on trials undertaken by each student. In most studies conducted on test accommodations, students' scores either increased because of the accommodation or remained unchanged. In this study, in disparateness, students suffered a potential detriment (impaired test performance) because of the accommodation. The potential for the detriment makes it essential that great vigilance be applied in assigning this accommodation and that it be assigned only because of prior evidence of benefit to the individual student. Regarding students' feelings, previous research indicated that students are generally well responsive to testing accommodations. For example, McKevitt and Elliott (2003) found that students in their study had positive views of a tester-reads-aloud accommodation, although they expressed concern that having the test items read aloud made them difficult to follow. Elliott et al. (2002) reported on a dissertation study by Marquart (2000) investigating extended time on a mathematics test for eighth-grade students. Students in this study were surveyed concerning their understanding of the accommodation. Most reported that they felt more comfortable, were more motivated, thought they had performed better, and preferred taking the test with the extended time accommodation. Interestingly, neither the effect size for SWOD (ES = 0.34) or that for

SWD (ES = 0.26) was statistically different from zero. Students were not very accurate in their perceptions of whether their test performance was enhanced as a result of the accommodation. In contrast to earlier research investigating teachers and students' predictions of benefit, this study examined students' post dictions of benefit. That is, students in this study had the potential advantage of having experienced the accommodated condition just prior to being questioned about their feelings. Still, the reflections of students in this study after being given the assessment, although statistically slightly better than chance, were far from correct. Translated into practical application, the experience of the accommodation in and of itself did not provide students with an accurate basis for determining whether they would be appropriate candidates for this accommodation. Conceivably, with repeated experience and feedback on results, student accuracy would improve. The incapability to accurately assess the impact of the accommodation characterized SWOD as well as SWD. Therefore, students appear to be no more accurate than teachers (cf. Fuchs et al., 2000; Helwig & Tindal, 2003) in their perceptions of the actual or potential impact of an accommodation on their test performance. (Elbaum,2010).

Lang's study (2005) extends the information available about student, parent, and teacher perceptions of accommodations for SWD on large-scale achievement tests. Prior research findings concerning students' preferences for accommodations on tests were diverse. Consistent with studies by Elliott and colleagues, Lang's study found that both fourth- and eighth grade SWD, and SWOD reported an overall preference for testing under the accommodated condition. Many of the connections between different groups of students and partiality for either testing condition were not significant, however.

Students' open-ended responses showed an inconsistent effect of testing accommodations on students' self-perception and self-efficacy. While some students showed feeling more confident and relaxed under the accommodated test condition, other Students reported feeling less secure in their abilities. When considering the potential consequences of testing accommodations for SWD, the self-beliefs of individual students should be taken into consideration. Although this study did not directly examine students' self-efficacy in relation to testing accommodations, the open-ended responses from the student questionnaire did supply an indication that testing accommodations can have either a positive or a negative effect on a student's reported sense of self efficacy. It is probable that this effect, in turn, could significantly influence a student's behavior during the testing. It would be unfortunate if testing accommodations perceived by teachers as beneficial had the opposite effect on some students, resulting in a tapering off in their motivation to work or a sign of their success to the testing accommodations rather than themselves. Therefore, it is important to have a clear understanding of student and teacher beliefs of testing accommodations to ensure that when testing accommodations are provided to students, the students' and teachers' beliefs of the accommodations are aligned (Lang et al., 2005).

Understanding the consequences of testing accommodations for students is paramount for setting up the consequential validity of such accommodations on large-scale achievement tests. The intended consequence of testing accommodations for students is to provide them with an opportunity to show their knowledge on tests without the interference of their disability. So, students' positive perception of testing accommodations provides support for the consequential validity of testing

accommodations on large-scale achievement tests. Moving forward, teachers' and parents' positive feelings of testing accommodations for SWD offer evidence that the consequences of the accommodations are indeed positive. Therefore, this study supplies further support for the continued use of testing accommodations through evidence affirming the consequential validity of their use.

In summary, except when read-aloud accommodations were provided in math tests via a computer or a video/audio player, both SWD and SWOD profited from the accommodation, though the effect size was generally greater for SWD. The fact that SWOD may also receive help from read-aloud accommodations, still raises a fairness and validity issue when providing the accommodation (Li & Suen, 2012; Phillips, 1994). Therefore, if read-aloud accommodations are only provided to SWD, SWOD may be at a detriment because they could have received help from the accommodations as well. In other words, the accommodation may even offer SWD an unfair advantage over SWOD. Many studies have contended with the effects of read-aloud accommodations, and it is apparent that more research is needed to fully understand the fairness and validity of test accommodations.

Future research would need to address the issues surrounding the use of accommodations with high stakes testing and the effect of testing accommodations on students' self-efficacy. Researchers would benefit from examining students' perceptions of specific types of testing accommodations or accommodation packages. It may be that SWD are more accepting and in favor of specific types of testing accommodations given their previous experiences with them. If a relationship was found between students' positive or negative ratings and certain types or packages of accommodations, it would be

advantageous to examine whether that relationship was mediated by disability type. Parents' feelings of testing accommodations also deserve further research since parents are key customers of testing accommodations through their children and imminently can have a large influence on the provision of accommodations for their children. It would also be helpful for future research to examine whether parents' perceptions of testing accommodations differ according to whether their children have disabilities.

Findings from these studies warrant further research on the effects that testing accommodations can have on students' self-efficacy, particularly since many SWD are likely to be exposed to testing accommodations during mandated tests. Taken into consideration, this research on student, teacher, and parent reactions to the use of testing accommodations provides evidence for the consequential validity of the resulting scores for all students.

The studies also add another piece to the experimental literature on testing accommodations for SWD. The findings of the studies are congruent with previous research in suggesting that the challenge of assigning the most effective and appropriate testing accommodations for SWD, like that of designing the most effective and appropriate instructional programs for these students, is unlikely to be successfully addressed by principles affecting entire populations of students defined by their category of disability. Rather, much more attention will need to be paid to individual students' characteristics and responses to accommodations in relation to types of tests and testing situations.

REFERENCES

Accommodations toolkit [Human Read Aloud: Research]. (n.d.). National Center on Educational Outcomes. https://publications.ici.umn.edu/nceo/accommodations-toolkit/human-read-aloud-research

Albright, L. K. & Ariail, M. (2005). Every child, every day. *Educational Leadership, 69*(6), 10–15.

American Association of State Colleges and Universities Teacher Preparation Task Force. (2017). *Preparing teachers in today's challenging context: Key issues, policy directions, and implications for leaders of AASCU universities*. https://www.aascu. org/AcademicAffairs/TeacherEdReport.pdf.

Anderson Koenig, J. & Bachman, L. F. (2004). *Keeping score for all*. National Academies Press. https://doi.org/10.17226/11029

Ariail, M. & Albright, L. K. (2006). A survey of teachers' read-aloud practices in middle schools. *Reading Research and Instruction*, *45*(2), 69–89. https://doi.org/10.1080/19388070609558443

Barrera, M., Liu, K., Thurlow, M. & Chamberlain, S. (2006). Use of chunking and questioning aloud to improve the reading comprehension of English language learners with disabilities. ells with disabilities report 17. *National Center on Educational Outcomes, University of Minnesota.*

Barton, K. E. W. (2001). *Stability of constructs across groups of students with different disabilities on a reading assessment under standard and accommodated administrations*. National Center on Educational Outcomes. https://nceo.info/references/thesis/10831

Berkowitz, D. (2011). Oral Storytelling: Building community through dialogue, engagement, and problem solving. *YC Young Children*, *66*(2), 36–40.

Berryhill, J., Linney, J. & Fromewick, J. (2009). The effects of educational accountability on teachers: Are policies too stress provoking for their own good? *International Journal of Education Policy and Leadership*, *4*(5). https://doi.org/10.22230/ijepl.2009v4n5a99

Bintz, W. P. (2020). Analyzing the frequency and content of read aloud articles published in selected journals across the K-8 curriculum. *Australian International Academic Centre*.

Bolt, S. E. & Roach, A. T. (2009). Inclusive assessment and accountability: A guide to accommodations for students with diverse needs. *New York, NY: Guilford Press*.

Bolt, S. E. & Thurlow, M. L. (2007). Item-level effects of the read-aloud accommodation for students with reading disabilities. *Assessment for Effective Intervention*, *33*(1), 15–28. https://doi.org/10.1177/15345084070330010301

Bornstein, M., Hedges, L. V., Higgins, J. P. T. & Rothstein, H. R. (2009). Introduction to meta-analysis. *Chichester, UK: John Wiley & Sons*. https://doi.org/10.1002/9780470743386

Braun, P. (2010). Taking the time to read aloud. *Science Scope, 34*(2), 45–49.

Briggs, D. C., Ruiz-Primo, M. A., Furtak, E., Shepard, L. & Yin, Y. (2012). Meta-analytic methodology and inferences about the efficacy of formative assessment. *Educational Measurement: Issues and Practice*, *31*(4), 13–17. https://doi.org/10.1111/j.1745-3992.2012.00251.x

Brumfield, G. A. (2014). *The effectiveness of reading accommodations for high school students with reading disabilities*. National Center on Educational Outcomes. https://nceo.info/references/thesis/12461

Buzick, H. & Stone, E. (2014a). A meta-analysis of research on the read aloud accommodation. *Educational Measurement: Issues and Practice*, *33*(3), 17–30. https://doi.org/10.1111/emip.12040

Buzick, H. & Stone, E. (2014b). A meta-analysis of research on the read aloud accommodation. *Educational Measurement: Issues and Practice*, *33*(3), 17–30. https://doi.org/10.1111/emip.12040

Calhoon, M., Fuchs, L. S. & Hamlett, C. L. (2000). Effects of computer-based test accommodations on mathematics performance assessments for secondary students with learning disabilities. *Learning Disability Quarterly*, *23*(4), 271–282. https://doi.org/10.2307/1511349

Campbell, D. & Kenny, D. (1999). A primer on regression artifacts. *New York, NY: Guilford*.

Cawthon, S. W. (2007). Accommodations use for statewide standardized assessments: Prevalence and recommendations for students who are deaf or hard of hearing. *Journal of Deaf Studies and Deaf Education*, *13*(1), 55–96. https://doi.org/10.1093/deafed/enm029

Center for the Study of Readers, Center for the Study of Readers, Anderson, R. C., Hiebert, E. H., Scott, J. & Wilkinson, I. (1985). *Becoming a nation of readers: The report of the commission on reading* (Highlighting Underlining ed.). Becoming a Nation of Readers.

Cervetti, G. N., Pearson, P., Palincsar, A. S., Afflerbach, P., Kendeou, P., Biancarosa, G., Higgs, J., Fitzgerald, M. S. & Berman, A. I. (2020). How the reading for understanding initiative's research complicates the simple view of reading invoked in the science of reading. *Reading Research Quarterly*, *55*(S1). https://doi.org/10.1002/rrq.343

Cervetti, G. N., Wright, T. S. & Hwang, H. (2016). Conceptual coherence, comprehension, and vocabulary acquisition: A knowledge effect? *Reading and Writing*, *29*(4), 761–779. https://doi.org/10.1007/s11145-016-9628-x

Chiu, C. W. & Pearson, P. D. (1999). Synthesizing the effects of test accommodations for special education and limited English proficient students. *American Institutes for Research in the Behavioral Sciences*.

Cochran-Smith, M. (1984). *The making of a reader: (language and learning for human service professions)*. Praeger.

Cochran-Smith, M. (1985). Looking like readers, talking like readers. *Theory Into Practice*, *24*(1), 22–31. https://doi.org/10.1080/00405848509543142

Cohen, J. (1992). A power primer. *Psychological Bulletin*, *112(1)*, 155–159. https://doi.org/10.1037/0033-2909.112.1.155

Collins & Shaeffer. (1997). *Vol. 52, no. 5, July 1997 of young children on jstor*. https://www.jstor.org/stable/i40102659

Creswell, J. W. & Creswell, J. D. (2018). *Research design: Qualitative, quantitative, and mixed methods approaches* (5th ed.). SAGE Publications, Inc.

Daisey, P. (1993). Three ways to promote the values and uses of literacy at any age. *Journal of Reading*, *36*(6), 436–440.

Delo, L. (2008). Reading aloud: Integrating science and literature for all students. *The Science Teacher, 75*(1), 33–37.

DerSimonian, R. & Laird, N. (1986). Meta-analysis in clinical trials. *Controlled Clinical Trials, 7*(3), 177–188. https://doi.org/10.1016/0197-2456(86)90046-2

Destefano, L., Shriner, J. G. & Lloyd, C. A. (2001). Teacher decision making in participation of SWD in large-scale assessment. *Exceptional Children*, *68*(1), 7–22. https://doi.org/10.1177/001440290106800101

Dolan, R. P., Hall, T. E., Banerjee, M., Chun, E. & Strangman, N. (2005). Applying principles of universal design to test delivery: The effect of computer-based read aloud on test performance of high school students with learning disabilities. *The Journal of Technology, Learning and Assessment*, *3*(7), 4–32.

Douglas, K. M. & Albro, E. R. (2014). The progress and promise of the reading for understanding research initiative. *Educational Psychology Review*, *26*(3), 341–355. https://doi.org/10.1007/s10648-014-9278-y

Dreher, S. (2003a). A novel idea: Reading aloud in a high school English classroom. *The English Journal*, *93*(1), 50–53.

Dreher, S. (2003b). A novel idea: Reading aloud in a high school English classroom. *The English Journal*, *93*(1), 50. https://doi.org/10.2307/3650570

du Plessis, A., Cullinan, M., Gramotnev, G., Gramotnev, D. K., Hoang, N. T., Mertens, L., Roy, K. & Schmidt, A. (2020). The multilayered effects of initial teacher education programs on the beginning teacher workforce and workplace: Perceptions of beginning teachers and their school leaders. *International Journal of Educational Research*, *99*, 101488. https://doi.org/10.1016/j.ijer.2019.09.010

Duke, N. K., Ward, A. E. & Pearson, P. (2021a). The science of reading comprehension instruction. *The Reading Teacher*, *74*(6), 663–672. https://doi.org/10.1002/trtr.1993

Duke, N. K., Ward, A. E. & Pearson, P. (2021b). The science of reading comprehension instruction. *The Reading Teacher*, *74*(6), 663–672. https://doi.org/10.1002/trtr.1993

Duncan, S. (2010). Instilling a lifelong love of reading. *Kappa Delta Pi Record*, *46*(2), 90–93. https://doi.org/10.1080/00228958.2010.10516701

Education & Chicago, U. O. (1992). *The elementary school journal*. University of Chicago Press.

Elbaum, B. (2007). Effects of an oral testing accommodation on the mathematics performance of secondary students with and without learning disabilities. *The Journal of Special Education*, *40*(4), 218–229. https://doi.org/10.1177/00224669070400040301

Elbaum, B., Arguelles, M., Campbell, Y. & Saleh, M. (2004). Effects of a student-reads-aloud accommodation on the performance of students with and without learning disabilities on a test of reading comprehension. *Exceptionality*, *12*(2), 71–87. https://doi.org/10.1207/s15327035ex1202_2

Fernandes, F., de La Higuera Amato, C., Cardoso, C., Navas, A. & Molini-Avejonas, D. (2016). Reading in autism spectrum disorders: A literature review. *Folia Phoniatrica et Logopaedica*, *67*(4), 169–177. https://doi.org/10.1159/000442086

Fien, H., Santoro, L., Baker, S. K., Park, Y., Chard, D. J., Williams, S. & Haria, P. (2011). Enhancing teacher reads-aloud with small-group vocabulary instruction

for students with low vocabulary in first-grade classrooms. *School Psychology Review*, *40*(2), 307–318. https://doi.org/10.1080/02796015.2011.12087720

Fincher, M. (n.d.). *Investigating the effects of a read-aloud alteration on the third-grade reading criterion-referenced competency test (crct) for SWD*. ScholarWorks @ Georgia State University. https://doi.org/10.57709/4083395

Fischer, R. & Mansell, A. (2009). Commitment across cultures: A meta-analysis approach. *Journal of International Business Studies*, *40(8)*, 1339–1358. https://doi.org/10.1057/jibs.2009.14

Fisher, D., Flood, J., Lapp, D. & Frey, N. (2004). Interactive reads-aloud: Is there a common set of implementation practices? *The Reading Teacher*, *58*(1), 8–17. https://doi.org/10.1598/rt.58.1.1

Fletcher, J. M., Francis, 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. https://doi.org/10.1177/001440290607200201

Fletcher, J. M., Francis, D. J., O'Malley, K., Copeland, K., Mehta, P., Caldwell, C. J., Kalinowski, S., Young, V. & Vaughn, S. (2009). Effects of a bundled accommodations package on high-stakes testing for middle school students with reading disabilities. *Exceptional Children*, *75*(4), 447–463. https://doi.org/10.1177/001440290907500404

Flowers, C., Kim, D.-H., Lewis, P. & Davis, V. (2011a). A comparison of computer-based testing and pencil-and-paper testing for students with a read-aloud

accommodation. *Journal of Special Education Technology*, *26*(1), 1–12. https://doi.org/10.1177/016264341102600102

Flowers, C., Kim, D. H., Lewis, P. & Davis, V. (2011b). A comparison of computer-based testing and pencil-and-paper testing for students with a read-aloud accommodation. *Journal of Special Education Technology*, *26*(1), 1–12. https://doi.org/10.1177/016264341102600102

Fox, M. (2001). Reading magic. *San Diego, Harcourt, Inc.*

Fox, M. (2013). What next in the read-aloud battle? win or lose? *The Reading Teacher*.

Fuchs, D. & Fuchs, L. S. (2006). Introduction to response to intervention: What, why, and how valid is it? *Reading Research Quarterly*, *41*(1), 93–99. https://doi.org/10.1598/rrq.41.1.4

Fuchs, L. S. & Fuchs, D. (1999). Helping teachers formulate sound test accommodation decisions for students with learning disabilities. *Learning Disabilities Research & Practice*, *16(3)*, 174–181. https://doi.org/10.1111/0938-8982.00018

Fuchs, L. S., Fuchs, D. & Capizzi, A. M. (2017). Identifying appropriate test accommodations for students with learning disabilities. *Focus on Exceptional Children*, *37*(6). https://doi.org/10.17161/fec.v37i6.6812

Fuchs, L. S., Fuchs, D., Eaton, S. B., Hamlett, C. L. & Karns, K. M. (2000). Supplementing teacher judgments of mathematics test accommodations with objective data sources. *School Psychology Review*, *29(1)*, 65–85. https://doi.org/10.1080/02796015.2000.12085998

Giusto, M. & Ehri, L. C. (2018). Effectiveness of a partial read-aloud test accommodation to assess reading comprehension in students with a reading

disability. *Journal of Learning Disabilities*, *52*(3), 259–270. https://doi.org/10.1177/0022219418789377

Glass, G. V. (1977). Integrating findings: The meta-analysis of research. *Review of Research in Education*, *5*(1), 351–379.

Goldman, S. R., Hogaboam, T. W., Bell, L. C. & Perfetti, C. A. (1980). Short-term retention of discourse during reading. *Journal of Educational Psychology*, *72*(5), 647–655. https://doi.org/10.1037/0022-0663.72.5.647

Goldstein, D. (2022, May 22). *In the fight over how to teach reading, this guru makes a major retreat*. https://www.nytimes.com/2022/05/22/us/reading-teaching-curriculum-phonics.html.

Goodwin, A. P. & Ahn, S. (2013a). A meta-analysis of morphological interventions in English: Effects on literacy outcomes for school-age children. *Scientific Studies of Reading*, *17*(4), 257–285. https://doi.org/10.1080/10888438.2012.689791

Goodwin, A. P. & Ahn, S. (2013b). A meta-analysis of morphological interventions in English: Effects on literacy outcomes for school-age children. *Scientific Studies of Reading*, *17*(4), 257–285. https://doi.org/10.1080/10888438.2012.689791

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. https://doi.org/10.1177/0022219409355484

Hall, M. (1971). Literature experiences provided by cooperating teachers. *The Reading Teacher*, *24*(5).

Hamman, D., Button, K., Olivarez, A., Lesley, M., Chan, Y., Griffith, R. & Woods, K. (2007). Relation between the reading instruction of cooperating and student teachers. *Literacy Teaching and Learning*, *10*(2), 77–94.

Harker, J. K. & Feldt, L. S. (1993). A comparison of achievement test performance of nondisabled students under silent reading and reading plus listening modes of administration. *Applied Measurement in Education*, *6*(4), 307–320. https://doi.org/10.1207/s15324818ame0604_4

Harlin, R. P. (1999). Developing future professionals: Influences of literacy coursework and field experiences. *Reading Research and Instruction*, *38*(4), 351–370. https://doi.org/10.1080/19388079909558301

Harris, L. S. (2008). *Comparison of student performance between teacher read and CD-ROM delivered modes of test administration of English language arts tests*. National Center on Educational Outcomes. https://scholarworks.wmich.edu/cgi/viewcontent.cgi?article=4212&context=dissertations.

Harvey, S. & Goudvis, A. (2007). Strategies that work: teaching comprehension for understanding, engagement, and building knowledge. *Portland, ME: Stenhouse Publishers.*

Hays, W. L. (1973). Statistics for the social sciences. *Holt, Rinehart, and Winston*, (2nd ed.).

Helwig, R., Rozek-Tedesco, M. A. & Tindal, G. (2002). An oral versus a standard administration of a large-scale mathematics test. *The Journal of Special Education*, *36*(1), 39–47. https://doi.org/10.1177/00224669020360010401

Helwig, R. & Tindal, G. (2003). An experimental analysis of accommodation decisions on large-scale mathematics tests. *Exceptional Children*, *69*(2), 211–225. Retrieved 2023, from https://doi.org/10.1177/001440290306900206

Hoffman, J. L., Collins, M. F. & Schickedanz, J. A. (2015). Instructional challenges in developing young children's science concepts: using informational text reads-aloud. *Reading Teacher, 68*(5), *363–372.* https://doi.org/10.1002/trtr.1325

Hoffman, J. V., Roser, N. L. & Battle, J. (1993). Reading aloud in classrooms: From the modal toward a "model". *The Reading Teacher*, *46*(6), 496–503.

Hollenbeck, K. & Crawford, L. (2005). Validity issues and decisions about testing accommodations. *Assessment for Effective Intervention*, *31*(1), 7–17. https://doi.org/10.1177/073724770503100102

Hoover, W. A. & Gough, P. B. (1990). The simple view of reading. *Reading and Writing*, *2*(2), 127–160. https://doi.org/10.1007/bf00401799

Horn, J. & Wilburn, D. (2013). *The Mismeasure of Education*. Information Age Publishing.

Hox, J. J. (2004). *Multi-level analysis: Techniques and applications* (2nd ed.). New York, NY: Routledge.

Hox, J. J. (2010). *Multilevel analysis: Techniques and applications* (2nd ed.). New York, NY: Routledge. https://doi.org/10.4324/9780203852279

Hunter, J. E., Schmidt, F. L. & Jackson, G. B. (1982). *Meta-analysis: Cumulating findings across research*. Beverly Hills, CA: Sage.

Huynh, H., Meyer, J. & Gallant, D. J. (2004). Comparability of student performance between regular and oral administrations for a high-stakes mathematics test.

Applied Measurement in Education, *17*(1), 39–57. https://doi.org/10.1207/s15324818ame1701_3

International Literacy Association. (n.d.). *Standards for the preparation of literacy professionals 2017*. ILA.

Jacobs, S. J. & Morriso, T. G. (2000). Reading aloud to students: A national probability study of classroom reading practices of elementary school teachers. *Reading Psychology*, *21*(3), 171–193. https://doi.org/10.1080/02702710050144331

Juhasz, B. J. & Berkowitz, R. N. (2011). Effects of morphological families on English compound word recognition: A multitask investigation. *Language and Cognitive Processes*, *26*(4-6), 653–682. https://doi.org/10.1080/01690965.2010.498668

Kappel, A. T. (2002). *The effects of testing accommodations on subtypes of students with learning disabilities*. National Center on Educational Outcomes. https://nceo.info/references/thesis/10856

Keene, E. O. & Zimmermann, S. (1997a). *Mosaic of thought: Teaching comprehension in a reader's workshop* (Teachers ed.). Heinemann.

Keene, E. & Zimmermann, S. (1997b). *Mosaic of thought: Teaching comprehension in a reader's workshop* (Teachers ed.). Heinemann.

Kieffer, M. J., Rivera, M. & Francis, D. J. (2012 update). *Practical guidelines for the education of English language learners: Research based recommendations for the use of accommodations in large-scale assessments*. Portsmouth, NH: RNC Research Corporation, Center on Instruction.

Kiili, C., Mäkinen, M. & Coiro, J. (2013). Rethinking academic literacies. *Journal of Adolescent & Adult Literacy*, *57*(3), 223–232. https://doi.org/10.1002/jaal.223

Kim, D. H., Schneider, C. & Siskind, T. (2009). Examining the underlying factor structure of a statewide science test under oral and standard administrations. *Journal of Psychoeducational Assessment*, *27*(4), 323–333. https://doi.org/10.1177/0734282908328632

Kinnunen, R., Vauras, M. & Niemi, P. (1998a). Comprehension monitoring in beginning readers. *Scientific Studies of Reading*, *2*(4), 353–375. https://doi.org/10.1207/s1532799xssr0204_4

Kinnunen, R., Vauras, M. & Niemi, P. (1998b). Comprehension monitoring in beginning readers. *Scientific Studies of Reading*, *2*(4), 353–375. https://doi.org/10.1207/s1532799xssr0204_4

Koon, S., Foorman, B. & Galloway, T. (2020). Identifying North Carolina students at risk of scoring below proficient in reading at the end of grade 3. *Washington, DC: Regional Educational Laboratory Southeast, National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education.*

Koon, S., Foorman, S., Galloway, B. & Galloway, T. (2020). Identifying North Carolina students at risk of scoring below proficient in reading at the end of grade 3. *Semantic Scholar*.

Korbey, H. & Arial, M. (2013). Why reading aloud to older children is valuable. Retrieved from *http://www.kqed.org/mindshift/2022/10/9/why-reading-aloud-to-older-childrenis-valuable.*

Koretz, D. & Hamilton, L. (2001). The performance of SWD on New York's revised regents' comprehensive examination in English. *UCLA National Center for Research on Evaluation, Standards and Student Testing (CRESST)*.

Koretz, D., Linn, R. L., Dunbar, S. B. & Shepard, L. A. (1991). The effects of high-stakes testing on achievement: Preliminary findings about generalization across tests. *Semantic Scholar*.

Kraft, M. A., Brunner, E. J., Dougherty, S. M. & Schwegman, D. J. (2020). Teacher accountability reforms and the supply and quality of new teachers. *Journal of Public Economics*, *188*, 104212. https://doi.org/10.1016/j.jpubeco.2020.104212

Kuhn, M. R. (2020a). Whole class or small group fluency instruction: A tutorial of four effective approaches. *Education Sciences*, *10*(5), 145. https://doi.org/10.3390/educsci10050145

Kuhn, M. R. (2020b). Whole class or small group fluency instruction: A tutorial of four effective approaches. *Education Sciences*, *10*(5), 145. https://doi.org/10.3390/educsci10050145

Kuhn, M. R., Schwanenflugel, P. J. & Meisinger, E. B. (2010). Aligning theory and assessment of reading fluency: Automaticity, prosody, and definitions of fluency. *Reading Research Quarterly*, *45*(2), 230–251. https://doi.org/10.1598/rrq.45.2.4

Lai, S. A. & Berkeley, S. (2012). High-stakes test accommodations: Research and practice. *Learning Disability Quarterly*, *35*(3), 158–169.

Laitusis, C. C. (2007). *Research designs and analysis for studying accommodations on assessments*. Retrieved 2023, from

https://www.ets.org/research/policy_research_reports/publications/chapter/2007/ddmg.html.

Laminack, L. & Wadsworth, R. (2006). *Learning under the influence of language and literature: Making the most of reads-aloud across the day*. Portsmouth, NH: Heinemann.

Laminack, L. & Wadworths, R. (2006). *Unwrapping the read aloud: Making every read aloud intentional and instructional (theory and practice in action)* (Pap/DVD ed.). Scholastic Teaching Resources (Theory and Practice).

Lane, H. B. & Wright, T. L. (2007). Maximizing the effectiveness of reading aloud. *The Reading Teacher*, *60*(7), 668–675. https://doi.org/10.1598/rt.60.7.7

Lang, S. C., Kumke, P. J., Ray, C. E., Cowell, E. L., Elliott, S. N., Kratochwill, T. R. & Bolt, D. M. (2005). Consequences of using testing accommodations: Student, teacher, and parent perceptions of and reactions to testing accommodations. *Assessment for Effective Intervention*, *31*(1), 49–62. https://doi.org/10.1177/073724770503100105

Lazarus, S. S., Thurlow, M. L., Rieke, R., Halpin, D. & Dillon, T. (2012). *Using cognitive labs to evaluate student experiences with the read-aloud accommodation in math* [PDF]. https://publications.ici.umn.edu/nceo/accommodations-toolkit/human-read-aloud-research. http://www.cehd.umn.edu/NCEO/OnlinePubs/Tech67/TechnicalReport67.pdf.

Lee, S. & Tsai, S.-F. (2016a). Experimental intervention research on students with specific poor comprehension: A systematic review of treatment outcomes. *Reading and Writing*, *30*(4), 917–943. https://doi.org/10.1007/s11145-016-9697-x

Lee, S. & Tsai, S. F. (2016b). Experimental intervention research on students with specific poor comprehension: A systematic review of treatment outcomes. *Reading and Writing*, *30*(4), 917–943. https://doi.org/10.1007/s11145-016-9697-x

Lee Swanson, H. (2011). Dynamic testing, working memory, and reading comprehension growth in children with reading disabilities. *Journal of Learning Disabilities*, *44*(4), 358–371. https://doi.org/10.1177/0022219411407866

Leko, M. M. & Brownell, M. T. (2011). Special education preservice teachers' appropriation of pedagogical tools for teaching reading. *Exceptional Children*, *77*(2), 229–251. https://doi.org/10.1177/001440291107700205

Levine, A. (2005). Educating school leaders. *Education Schools Project*. Retrieved September 12, 2021, from

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), 3–16. https://doi.org/10.1111/emip.12027

Libresco, A. S., Balantic, J. & Kipling, J. C. (2011). *Every book is a social studies book*. Santa Barbara, CA: Libraries Unlimited.

Lickteig, M. & Russell, J. (1993). Elementary teachers' read-aloud practices. *Reading Improvement, 30*(4), 202.

Lipsey, M. W. & Wilson, D. B. (2001). *Practical meta-analysis*. London: Sage.

Liu, K. K., Ward, J. M., Thurlow, M. L., Christensen, L. L. & Albus, D. (2015). Large-scale assessment and English language learners with disabilities. *Educational Policy*, *31*(5), 551–583. https://doi.org/10.1177/0895904815613443

Lopez-Fernandez, O. & Molina-Azorin, J. F. (2011). The use of mixed methods research in the field of behavioral sciences. *Quality & Quantity*, *45*(6), 1459–1472. https://doi.org/10.1007/s11135-011-9543-9

Lovett, B. J. (2010). Extended time testing accommodations for SWD. *Review of Educational Research*, *80*(4), 611–638. https://doi.org/10.3102/0034654310364063

Lunneborg, C. E. (1979). Book review: Psychometric theory: Second edition Jum C. Nunnally New York: McGraw-Hill, 1978, 701 pages. *Applied Psychological Measurement*, *3*(2), 279–280. https://doi.org/10.1177/014662167900300216

Lyster, S. A., Lervåg, A. & Hulme, C. (2016a). Preschool morphological training produces long-term improvements in reading comprehension. *Reading and Writing*, *29*(6), 1269–1288. https://doi.org/10.1007/s11145-016-9636-x

Lyster, S. A., Lervåg, A. & Hulme, C. (2016b). Preschool morphological training produces long-term improvements in reading comprehension. *Reading and Writing*, *29*(6), 1269–1288. https://doi.org/10.1007/s11145-016-9636-x

Madaus, G., Russell, M. & Higgins, J. (2009). Paradoxes of high-stakes testing. *Journal of Education*, *190*(1–2), 21–30. https://doi.org/10.1177/0022057410190001-205

Maheady, L. & Gard, J. (2010). Classwide peer tutoring: Practice, theory, research, and personal narrative. *Intervention in School and Clinic*, *46*(2), 71–78. https://doi.org/10.1177/1053451210376359

Maheady, L., Jabot, M., Rey, J. & Michielli-Pendl, J. (2007). An early field-based experience and its impact on pre-service candidates' teaching practice and their pupils' outcomes. *Teacher Education and Special Education: The Journal of the Teacher Education Division of the Council for Exceptional Children*, *30*(1), 24–33. https://doi.org/10.1177/088840640703000103

Marsh, H., Bornmann, L., Mutz, R., Daniel, H. D. & O'Mara, A. (2009). Gender effects in the peer reviews of grant proposals: A comprehensive meta-analysis comparing traditional and multilevel approaches. *Review of Educational Research*, *79*(3), 1290–1326. https://doi.org/10.3102/0034654309334143

Thurlow, M. L. (2014a). Accommodation for challenge, diversity and variance in human characteristics. *The Journal of Negro Education*, *83*(4), 442. https://doi.org/10.7709/jnegroeducation.83.4.0442

Thurlow, M. L. (2014b). Accommodation for challenge, diversity and variance in human characteristics. *The Journal of Negro Education*, *83*(4), 442. https://doi.org/10.7709/jnegroeducation.83.4.0442

Martin, J. B. (2011). *Brown bear, brown bear, what do you see? (brown bear and friends)*. Henry Holt and Co. (BYR).

McCormack, A. (2019). *An investigation into the use of the "tests read" accommodation for the New York State 4th grade ELA assessment, the impact on reported scores, and the reported knowledge and beliefs of administrators regarding the assignment and purpose of accommodations*. National Center for Educational Outcomes. https://nceo.info/references/4PrTvxA8RPKR0TDOu4Z_LQ.

McCormick, M. K. & McTigue, E. M. (2011). Teacher reads-aloud make science come alive, *Science Scope, 34*(5), 45–49.

McCormick, S. (1977a). *Should you read aloud to your children?* JSTOR. http://www.jstor.org/stable/41404495

McCormick, S. (1977b, February). *Should you read aloud to your children?* https://www-jstor-org.jerome.stjohns.edu/stable/pdf/41404495.pdf?refreqid=excelsior%3Ab6f21945a45e260e4627c8f20393d06b. Retrieved September 12, 2021, from http://www.jstor.org/stable/41404495

McDonnell, L. M., McLaughlin, M. J. & Morison, P. (1997). *Educating one and all*. National Academies Press. https://doi.org/10.17226/5788

McGee, L. M. & Schickedanz, J. (2007). *Repeated interactive read alouds in preschool and kindergarten*. Reading Rockets. https://www.readingrockets.org/article/repeated-interactive-read-alouds-preschool-and-kindergarten.

McKevitt, B. C. & Elliott, S. N. (2003). Effects and perceived consequences of using read-aloud and teacher-recommended testing accommodations on a reading achievement test. *School Psychology Review*, *32*(4), 583–600. https://doi.org/10.1080/02796015.2003.12086222

McMahon, D., Wright, R., Cihak, D. F., Moore, T. C. & Lamb, R. (2015). Podcasts on mobile devices as a read-aloud testing accommodation in middle school science assessment. *Journal of Science Education and Technology*, *25*(2), 263–273. https://doi.org/10.1007/s10956-015-9591-3

Meloy, L. L., Deville, C. & Frisbie, D. A. (2002a). 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*(4), 248–255. https://doi.org/10.1177/07419325020230040801

Meloy, L. L., Deville, C. & Frisbie, D. A. (2002b). The effect of a read aloud accommodation on test scores of students with and without a learning disability in reading. *Remedial & Special Education, 23*(4), 248–255.

Miller, D. (2012). *Reading with meaning* (2nd ed.). Routledge.

Miller, J. & Schwanenflugel, P. J. (2008). A longitudinal study of the development of reading prosody as a dimension of oral reading fluency in early elementary school children. *Reading Research Quarterly*, *43*(4), 336–354. https://doi.org/10.1598/rrq.43.4.2

Morris, S. B. & DeShon, R. P. (2002). Combining effect size estimates in meta-analysis with repeated measures and independent-groups designs. *Psychological Methods*, *7*(1), 105–125. https://doi.org/10.1037/1082-989X.7.1.105

Morrow, L. M. (1997). *The literacy center: Contexts for reading and writing*. Stenhouse Pub.

Nation, K. (1999). Reading skills in hyperlexia: A developmental perspective. *Psychological Bulletin*, *125*(3), 338–355. https://doi.org/10.1037/0033-2909.125.3.338

National center for education statistics (NCES) home page, a part of the U.S. department of education. (2022). https://nces.ed.gov/.

National institute of child health and human development. (2000). Report of the national reading panel. teaching children to read: an evidence-based assessment of the scientific research literature on reading and its implications for reading instruction: reports of the subgroups. *NIH Publication No. 00–4754. Washington, DC: U.S. Government Printing Office.*

National reading panel - teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction [PDF]. (2008). https://www.nichd.nih.gov/sites/default/files/publications/pubs/nrp/Documents/report.pdf.

National reading panel (historical/for reference only). (n.d.). https://www.nichd.nih.gov/. https://www.nichd.nih.gov/research/supported/nrp.

National reading panel publications. (n.d.). https://www.nichd.nih.gov/. https://www.nichd.nih.gov/about/org/der/branches/cdbb/nationalreadingpanelpubs

Nolet, V. & McLaughlin, M. J. (2005). *Accessing the general curriculum: Including SWD in standards-based reform*. Corwin Press. https://doi.org/10.4135/9781483329253

Noortgate, W. V. D. & Onghena, P. (2003). Multilevel meta-analysis: A comparison with traditional meta-analytical procedures. *Educational and Psychological Measurement*, *63*(5), 765–790.

Oakhill, J. (2020a). Four decades of research into children's reading comprehension: A personal review. *Discourse Processes*, *57*(5–6), 402–419. https://doi.org/10.1080/0163853x.2020.1740875

Oakhill, J. (2020b). Four decades of research into children's reading comprehension: A personal review. *Discourse Processes*, *57*(5-6), 402–419. https://doi.org/10.1080/0163853x.2020.1740875

Oczkus, L. D. (2012). The power of reading aloud to your students. In *best ever literacy survival tips: 72 lessons you can't teach without* (pp. 19–28). New York, NY: Scholastic.

Pangatungan, M. (2018, June 22). *The impact of read-aloud accommodations of fourth- and fifth-grade elementary students with and without learning impairment: A descriptive case study*. DigitalCommons@CSP. https://digitalcommons.csp.edu/cup_commons_grad_edd/198/.

Paterson, P. O. & Elliott, L. N. (2006). Struggling reader to struggling reader: High school students' responses to a cross-age tutoring program. *Journal of Adolescent & Adult Literacy*, *49*(5), 378–389. https://doi.org/10.1598/jaal.49.5.2

Prater, M. & Sileo, T. W. (2004). Fieldwork requirements in special education preparation: A national study. *Teacher Education and Special Education: The Journal of the Teacher Education Division of the Council for Exceptional Children*, *27*(3), 251–263. https://doi.org/10.1177/088840640402700305

Pressley, M., El-Dinary, P., Gaskins, I., Schuder, T., Bergman, J. L., Almasi, J. & Brown, R. (1992). Beyond direct explanation: Transactional instruction of reading comprehension strategies. *The Elementary School Journal*, *92*(5), 513–555. https://doi.org/10.1086/461705

Randall, J. & Engelhard, G. (2009). Performance of students with and without disabilities under modified conditions. *The Journal of Special Education*, *44*(2), 79–93. https://doi.org/10.1177/0022466908331045

Raudenbush, S. W. & Bryk, A. S. (n.d.). Empirical Bayes meta-analysis. *Journal of Educational Statistics*, *10*(2), 75–98. https://doi.org/10.3102/10769986010002075

Raudenbush, S. W. & Bryk, A. S. (2002). *Hierarchical linear models: applications and data analysis methods* (2nd ed.). London: Sage.

Reads-aloud in the classroom: A pilot study of teachers' self-reporting practices - document - gale academic onefile. (n.d.). https://link.gale.com/apps/doc/A506829124/AONE?

Reutzel, D. R. & Cooter, R. B. (2008). *Teaching children how to read: The teacher makes the difference*. Upper Saddle River, NJ: Pearson Education.

Richardson, J. S. (2006). *Read it aloud! using literature in the secondary content classroom*. International Reading Association.

Riley, R. D., Lambert, P. C. & Abo-Zaid, G. (2010). Meta-analysis of individual participant data: Rationale, conduct, and reporting. *BMJ*, *340*(feb05 1), c221–c221. https://doi.org/10.1136/bmj.c221

Risko, V. J. & Reid, L. (2019). What really matters for literacy teacher preparation? *The Reading Teacher*, *72*(4), 423–429. https://doi.org/10.1002/trtr.1769

Rogers, C. M. (2019). *A summary of the research on effects of test accommodations: 2015–2016 (NCEO report 412)* [PDF]. https://nceo.umn.edu/docs/OnlinePubs/NCEOReport412.pdf.

Rogers, C. M., Christian, E. M. & Thurlow, M. L. (2012). A summary of the research on the effects of test accommodations: 2009-2010 (Technical Report 65). *Minneapolis, MN: University of Minnesota, National Center on Educational Outcomes.*

Rosenblad, A. (2009). Introduction to meta-analysis by Michael Borenstein, Larry v. hedges, Julian P.T. Higgins, Hannah r. Rothstein. *International Statistical Review*, *77*(3), 478–479. https://doi.org/10.1111/j.1751-5823.2009.00095_15.x

Routman, R. (2003). Reading essentials. *Portsmouth, NH: Heinemann.*

Samuels, S. J. & Wu, Y. (2004). How the amount of time spent on independent reading affects reading achievement. Proceedings of the 2004 annual convention of the international reading association, *International Reading Association. Reno, Nevada, 124–130.*

Santoro, L., Chard, D. J., Howard, L. & Baker, S. K. (2008). Making the very most of classroom reads-aloud to promote comprehension and vocabulary. *The Reading Teacher*, *61*(5), 396–408. https://doi.org/10.1598/rt.61.5.4

Schulte, A. G., Elliott, S. N. & Kratochwill, T. R. (2001). Effects of testing accommodations on standardized mathematics test scores: An experimental analysis of the performances of students with and without disabilities. *School Psychology Review*, *30*(4), 527–547. https://doi.org/10.1080/02796015.2001.12086133

Serafini, F. & Giorgis, C. (2003). *Reading aloud and beyond: Fostering the intellectual life with older readers*. Portsmouth, NH: Heinemann.

Serafini, F., Moses, L., Kachorsky, D. & Rylak, D. (2020). Incorporating multimodal literacies into classroom-based reading assessment. *The Reading Teacher*, *74*(3), 285–296. https://doi.org/10.1002/trtr.1948

Sireci, S. G., Scarpati, S. E. & Li, S. (2005). Test accommodations for SWD: An analysis of the interaction hypothesis. *Review of Educational Research*, *75*(4), 457–490. https://doi.org/10.3102/00346543075004457

Slay, L. E. & Morton, T. B. (2020). Engaging pre-service teachers in read alouds. *Changing English*, *27*(1), 64–82. https://doi.org/10.1080/1358684x.2019.1710696

Spencer, R. R. (2011, November). *Reading aloud and student achievement*. Digital Commons@Lindenwood University. https://digitalcommons.lindenwood.edu/dissertations/509.

Spiel, C., Evans, S. W. & Harrison, J. R. (2019). Does reading standardized tests aloud meet the scientific definition of an accommodation? *Journal of Applied School Psychology*, *35*(4), 380–399. https://doi.org/10.1080/15377903.2019.1601145

Stanovich, K. E. (1980). Toward an interactive-compensatory model of individual differences in the development of reading fluency. *Reading Research Quarterly*, *16*(1), 32. https://doi.org/10.2307/747348

Strickland, D. & Riley-Ayers, S. (2006). Early Literacy: Policy and practice in the preschool years. *National Institute for Early Education Research (NIEER)*.

Strickland, D. S. (1989). *Emerging literacy: Young children learn to read and write*. Intl Literacy Assn.

Strickland, D. S. & Morrow, L. M. (1989). *Emerging literacy: Young children learn to read and write*.

Suggate, S. P. (2014a). A meta-analysis of the long-term effects of phonemic awareness, phonics, fluency, and reading comprehension interventions. *Journal of Learning Disabilities*, *49*(1), 77–96. https://doi.org/10.1177/0022219414528540

Suggate, S. P. (2014b). A meta-analysis of the long-term effects of phonemic awareness, phonics, fluency, and reading comprehension interventions. *Journal of Learning Disabilities*, *49*(1), 77–96. https://doi.org/10.1177/0022219414528540

Swanson, E., Vaughn, S., Wanzek, J., Petscher, Y., Heckert, J., Cavanaugh, C., Kraft, G. & Tackett, K. (2011a). A synthesis of read-aloud interventions on early reading outcomes among preschool through third graders at risk for reading difficulties. *Journal of Learning Disabilities*, *44*(3), 258–275. https://doi.org/10.1177/0022219410378444

Swanson, E., Vaughn, S., Wanzek, J., Petscher, Y., Heckert, J., Cavanaugh, C., Kraft, G. & Tackett, K. (2011b). A synthesis of read-aloud interventions on early reading outcomes among preschool through third graders at risk for reading difficulties. *Journal of Learning Disabilities*, *44*(3), 258–275. https://doi.org/10.1177/0022219410378444

Tallal, P., Merzenich, M. M., Miller, S. & Jenkins, W. (1998). Language learning impairments: Integrating basic science, technology, and remediation. *Experimental Brain Research*, *123*(1–2), 210–219. https://doi.org/10.1007/s002210050563

Taylor, E. (2017). *The influence of testing accommodations on academic performance and disciplinary rates* (ED582587). ERIC. https://eric.ed.gov/?id=ED582587

Terblanche, L. (n.d.). Read-alouds: Do they enhance students' ability to read. *(Unpublished dissertation).* Retrieved September 12, 2021, from

Thurlow, M., Elliott, J. L. & Ysseldyke, J. E. (1998). *Testing SWD: Practical strategies for complying with district and state requirements*. Corwin Press.

Thurlow, M. L., Lazarus, S. S., Thompson, S. J., Morse, A., Christensen, L. L. & Braam, M. (2009). State policies on assessment participation and accommodations for SWD. *The Journal of Special Education*, *38*(4), 232–240. https://doi.org/10.1177/00224669050380040401

Thurlow, M. L., Moen, R. E., Lekwa, A. J. & Scullin, S. B. (2010). *Examination of a reading pen as a partial auditory accommodation for reading assessment*. Minneapolis, MN: University of Minnesota, Partnership for Accessible Reading Assessment.

Tindal, G., Heath, B., Hollenbeck, K., Almond, P. & Harniss, M. (1998). Accommodating SWD on large-scale tests: An experimental study. *Exceptional Children*, *64*(4), 439–450. https://doi.org/10.1177/001440299806400401

Toffalo, D. & Milke, R. M. (2007). Test reviews: Fuchs, L., Fuchs, D., Eaton, S. & Hamlett, C. (2003). Dynamic assessment of test accommodations. San Antonio, Tx: Psychological corporation. *Journal of Psychoeducational Assessment*, *26*(1), 83–91. https://doi.org/10.1177/0734282907300568

Trelease, J. (2001). *The read aloud handbook* (5th ed.). Penguin Books.

Trelease, J. (2006). *The read aloud handbook*. Noura Books.

Turner, F. (2010a). Evaluating the effectiveness of fluency-oriented reading instruction with increasing Black and Latino reading fluency, as compared to Asian and

white second-grade students' reading fluency. *The Journal of Negro Education, 79*(2), *112–124.*

Turner, F. D. (2010b). Evaluating the effectiveness of fluency-oriented reading instruction with increasing black and Latino reading fluency, as compared to Asian and white second-grade students' reading fluency. *The Journal of Negro Education*, *79*(2), 112–124. https://doi.org/10.2307/20798330

U.S. Dept of Health and Human Services. (2000). *Report of the national reading panel: Teaching children to read: An evidence-based assessment of the scientific research lit on reading and its implications for reading inst (reports of the subgroups)* (2nd ed.). National Inst of Child Health & Human Development.

Urdan, T. C. (n.d.). *Statistics in plain English* (4th ed.). Psychology Press. https://doi.org/10.4324/9781410612496

Vanchu-Orosco, M. (2012). *A meta-analysis of testing accommodations for SWD: Implications for high-stakes testing (unpublished dissertation)*. University of Denver, Denver CO.

Wadsworth. (1986). *Over in the meadow*.

Wiseman, A. (2010). Interactive read alouds: Teachers and students constructing knowledge and literacy together. *Early Childhood Education Journal*, *38*(6), 431–438. https://doi.org/10.1007/s10643-010-0426-9

Wooten, K. (2009). Novice Teachers' Perceptions of Preparedness to Teach by Certification Route. *Doctoral Diss: ProQuest LLC.* Retrieved September 12, 2021, from

Wright, K. B. (1992). *A fourth grade read aloud program*. ERIC.

Zarnowski, M. & Turkel, S. (2012). Creating new knowledge: books that demystify the process. *Journal of Children's Literature, 38*(1), 28–34.

Zehr, M. A. (2010). Reading aloud to teens gains favor among teachers. *Education Week, 76*(1), 1–7.

Zenisky, A. L. & Sireci, S. G. (2007). *A summary of the research on the effects of test accommodations: 2005–2006 (technical report 47)*. Minneapolis, MN: University of Minnesota, National Center on Educational Outcomes.

Zuriff, G. E. (2000). Extra examination time for students with learning disabilities: An examination of the maximum potential thesis. *Applied Measurement in Education*, *13*(1), 99–117. https://doi.org/10.1207/s15324818ame1301_5

Vita

Name	*Sarah Elizabeth Ryan*
Baccalaureate Degree	*Bachelor of Arts Degree, Christopher Newport University, Newport News, VA, Major: Art, Minor: Psychology*
Date Graduated	*January 1993*
Other Degrees and Certificates	*Master of Education, University of San Diego, San Diego, CA, Major: Educational Leadership.*
Date Graduated	*August 2018*