



**Developing and Researching
an Accessible Reading
Assessment for Students
with Disabilities**



Partnership for Accessible
Reading Assessment

Developing and Researching an Accessible Reading Assessment for Students with Disabilities

Martha L. Thurlow, Deborah R. Dillon, Jamal Abedi, and Marsha Brauen

May 2012

All rights reserved. Any or all portions of this document may be reproduced and distributed without prior permission, provided the source is cited as:

Thurlow, M. L., Dillon, D. R., Abedi, J., & Brauen, M. (2012). *Developing and researching an accessible reading assessment for students with disabilities*. Minneapolis, MN: University of Minnesota, Partnership for Accessible Reading Assessment.



Partnership for Accessible Reading Assessment

University of Minnesota
207 Pattee Hall
150 Pillsbury Dr. SE
Minneapolis, MN 55455

<http://www.readingassessment.info>
readingassess@umn.edu

This work is supported, in part, by the U.S. Department of Education, Institute of Education Sciences, National Center for Special Education Research—Grant No. H324F040002. Opinions expressed do not necessarily reflect those of the U.S. Department of Education or offices within it.



UNIVERSITY
OF MINNESOTA
**Curriculum and
Instruction**



UCDAVIS



The University of Minnesota is committed to the policy that all persons shall have equal access to its programs, facilities, and employment without regard to race, color, creed, religion, national origin, sex, age, marital status, disability, public assistance status, veteran status, or sexual orientation.

This document is available in alternative formats upon request.

Executive Summary

The National Center on Educational Outcomes and the Department of Curriculum and Instruction at the University of Minnesota, the University of California, Los Angeles, the University of California, Davis, and Westat engaged in a collaborative program of research and development on an accessible reading assessment for students with disabilities. Called the Partnership for Accessible Reading Assessment (PARA), the project produced an approach to the assessment of reading proficiency believed to be more accessible than previous large-scale assessments used for school accountability purposes. This report summarizes the field test of this accessible assessment comparing it to an assessment considered to be less accessible for students with disabilities in grades 4 and 8.

The field test included special education students with Individualized Education Programs (IEPs) who were categorized as having learning disabilities (LD), speech-language impairments (SLI), intellectual disabilities (ID), or deafness/hard of hearing (D/HH). Students with disabilities on 504 accommodation plans were not included; thus we refer to special education students (SPED) in this report rather than the more general students with disabilities. The students in the comparison group were students not receiving special education services; they are referred to as non-special education students (non-SPED) in this report.

PARA Accessible Assessment and Comparison Proxy State Assessment

The nature of the assessment used in the field test as the *accessible assessment* was based on existing literature, findings from PARA's research program, and other considerations such as universal design guidelines, accommodation policies for reading assessments, and research on reading assessment accommodations. These are described in the report.

The PARA accessible assessment incorporated three accessibility features: (a) segmented text, in which questions were asked after the section of the passage for which they were relevant, and all integrative questions were asked after the entire reading passage; (b) high interest passages that were identified by students as most interesting during a calibration study; and (c) color illustrations. The control assessment, called the *proxy state assessment* did not include these critical characteristics of the PARA accessible reading assessment. Specifically, the proxy state assessment: (a) did not have segmented text (all questions were asked after the entire passage); (b) used low interest passages that were identified by students as least interesting during a calibration study; and (c) provided no color. Both the accessible and the proxy state assessments were presented via paper and pencil booklets of four reading passages with seven questions per passage. Universal design features were applied to both assessments as well, including white space, clearly marked questions, and unlimited time for taking the assessment.

General Methodology

Passages were selected for the assessments based on a review of research in children's and young adult literature that indicated topics of interest to students in grades 4 and 8. They included both *informational-expository* and *narrative-literary* text types to conform to the NAEP 2009 Reading Framework guidelines.

Passages were 500-700 words for grade 4 students and 700-1000 words for grade 8 students. Each text was given a holistic rating of either easy, medium, or hard, based on Fry readability data and Lexiles analyses, as well as consideration of topic familiarity, concept load or degree of difficulty of the ideas presented about a topic, and the extent to which the passage had accompanying pictures or visual items that might support the reader's understanding of the text.

Criteria were developed to guide decisions about the inclusion or exclusion of visual elements for each passage. These are described in the report and provided in detail in an appendix. Consensus on the visual elements to include or exclude was based on individual analyses followed by group analysis and discussion. Agreement was almost 95% prior to discussion and 100% after discussion. After passages were identified and analyzed, permission was secured from publishers and illustrators or photographers to use the texts in the assessments.

Three *situational motivation questions* were included after each passage. These questions focused on (a) how interesting the passage was, (b) how challenging the passage was to read, and (c) whether the student liked to read the passage because it was a story (narrative texts) or they learned new facts (expository texts). These questions were boxed so that it was clear that they were different from the multiple choice items.

All students were asked to complete the Motivations for Reading Questionnaire (MRQ) on a day before the administration of the accessible or proxy state assessment. The MRQ consisted of 54 items that covered 11 dimensions of motivation organized subscales, five of which we used for the field test analyses (Reading Efficacy, Reading Challenge, Reading Curiosity, Aesthetic Enjoyment of Reading, Importance of Reading)

Planned versus Implemented Field Test Design

The planned field test design called for testing to begin in January, 2010 and continue consecutively in two states. A total of 150 special education students in each disability category and 150 students without disabilities were to be tested in grade 4 and grade 8 on the accessible assessment and on the proxy state assessment. The total number of students was to be 750 special education and 750 non-special education students tested in each condition (accessible assessment, proxy state assessment) in each grade, for a total of 3000 students in the field test.

Because of delays in obtaining IRB approval for the field tests and a late start date in identifying participating schools, the field test actually started in April, 2010. After

data collection in April and May, only a few students in some disability categories had participated – this was particularly the case for students with intellectual disabilities and students who were deaf/hard of hearing.

A revised data collection plan was established to increase the field test sample. It focused on obtaining sufficient numbers of students with learning disabilities and with speech language impairments. State approval was obtained to enter the schools again in the fall of 2010 in grades 5 and 9. Identification of participating schools started in the summer of 2010. The main reason for selecting grades 5 and 9 in the fall instead of grades 4 and 8 was to obtain students with similar levels of content knowledge/instruction in reading; this was more likely to be the case for spring and fall students in consecutive grades.

The actual numbers of students with learning disabilities nearly reached the desired numbers in grade 4/5 (n = 131 in the accessible assessment; n = 113 in the proxy state assessment) and exceeded them in grade 8/9 (n = 401 in the accessible assessment; n = 448 in the proxy state assessment). The desired numbers of students with speech-language impairments was not reached in either grade 4/5 (n = 68 in the accessible assessment; n = 72 in the proxy state assessment) or grade 8/9 (n = 23; n = 30).

Overall, the numbers of special education students participating in the field test for all but students with learning disabilities was considerably lower than planned, totaling 185 students at grade 4/5 and 150 students at grade 8/9. The number of participating non-special education students met and exceeded desired numbers. The total number of non-special education students was 860 at grade 4/5 and 1158 at grade 8/9. Not all students had all types of data needed for inclusion in all analyses. The numbers of students in each analysis is included with the results of the analysis.

Research Questions

The purpose of the PARA field test was to determine the degree to which its accessible reading assessment provided for accessibility, validity, and comparability for students with LD, SLI, ID, and D/HH. Analyses could not be conducted for each disability category separately because of limited numbers of students; thus, LD, SLI, ID, and D/HH were combined for one analysis. Four primary research questions addressed:

1. Is the PARA accessible reading assessment more accessible than a proxy state reading assessment for students with LD, and for students with disabilities combined (LD, SLI, ID, and D/HH)?
2. Are scores from the PARA accessible reading assessment more reliable than scores from the proxy state assessment for students with LD, and for students with disabilities combined (LD, SLI, ID, and D/HH)?
3. Is there evidence of the validity of scores from the accessible reading assessment?

4. Does the PARA accessible reading assessment demonstrate the same characteristics (student performance, reliability, validity) for students in grades 4 and 8?

In addition to the basic research questions, several secondary research questions were addressed:

1. Does students' performance on the reading assessment correlate with the overall general measure of motivation, based on the results of the Motivations for Reading Questionnaire (MRQ)?
2. Do students perform better on the passages they rate as more interesting?
3. Do special education and non-special education students rate informational-expository passages or narrative-literary passages as more interesting, or are both rated as equally interesting?
4. Does the text type in reading assessment passages (narrative-literary or informational-expository) impact student performance in general and by groups of students?
5. Do students score better on passages they rate as "kind of challenging" or "pretty challenging" in terms of how challenging (difficult) the passage was to read?
6. Do students perform better on the *narrative-literary* passages when they indicated that they liked the passage because they knew something about it or when they indicated that they liked the passage because it was a story? How often did they mark both responses as reasons for liking the passage, and does doing so have an effect on performance?
7. Do students perform better on the *informational-expository* passages when they rated that they liked the passage because it provided them with information? Or did students perform better on the informational-expository passage when they rated that they liked the passage because they knew something about the passage? Or did they mark both responses as reasons for liking the passage?
8. Do students who take the PARA accessible assessment more often use fewer accommodations than assigned compared to students who take the proxy state assessments?
9. Does the number of omitted or not reached items differ for the accessible and proxy state assessments? Does the level of student motivation affect the number of omitted or not reached items?

Results

The PARA field test of its accessible assessment compared to a proxy state assessment was conducted in the spring and fall of 2010, and included over 3700 students (over 1500 special education students and over 2100 non-special education students). The accessible assessment had higher interest passages, segmented text in the passages, and color illustrations. The proxy state assessment had lower interest passages, no segmenting of text, and no color illustrations.

Equipercenile equating was used to equate results from the spring and fall testing. Using the combined data, we present for each research question the basic results of our analyses.

Primary Research Question 1 (RQ1): Is the PARA accessible reading assessment more accessible than a proxy state reading assessment for students with LD, and for students with disabilities combined (LD, SLI, ID, and D/HH)?

Results of analyses for RQ1 did not support the hypothesis that students with disabilities would perform significantly better on the accessible reading assessment compared to the performance of students with disabilities who took the proxy state assessment. This finding suggested that the addition of higher interest passages, segmented text, and color were not enough to increase performance beyond that potentially produced by the already interesting passages (although rated as less interesting than the high interest passages) and universal design elements of both the accessible assessment and the proxy state assessment.

Primary Research Question 2 (RQ2): Are scores from the PARA accessible reading assessment more reliable than scores from the proxy state assessment for students with LD, and for students with disabilities combined (LD, SLI, ID, and D/HH)?

Results of analyses for RQ2 did not support the hypothesis that the accessible reading assessment had higher reliability than the proxy state reading assessment. Item level analyses indicated that the internal consistency for the accessible assessment and the proxy state assessment were similar for special education students at both grades 4 and 8.

Principal component analysis indicated that the items on both tests were unidimensional. Both tests were reliably tapping into the construct of reading comprehension, as typically assessed when multiple choice items follow the reading passages and test takers hold the implicit purpose of reading to recall or draw inferences.

Primary Research Question 3 (RQ3): Is there evidence of the validity of scores from the accessible reading assessment?

Results of analyses for RQ3 supported the hypothesis that accessibility features of the PARA accessible assessment did not alter the construct of reading. When using scores

equated based just on the Non-SPED and LD sample, results did show higher scores for Non-SPED students who took the proxy state assessment, but all other analyses indicated that the PARA accessibility features did not alter the focal construct of the study.

Primary Research Question 4 (RQ4): Does the PARA accessible reading assessment demonstrate the same characteristics (student performance, reliability, validity) for students in grades 4 and 8?

Results of analyses for RQ4 indicated that the accessible reading assessment demonstrated the same characteristics for students in grades 4/5 and 8/9. In both grades, there were no significant differences between the accessible and proxy state assessments. The same consistency across grades was found for the validity of scores, with both grades showing that the accessibility features incorporated into the assessment did not alter the construct being assessed. Finally, analyses of reliability of the accessible and proxy state assessments were consistent across the two grades.

Secondary Research Question 1 (SRQ1): Does students' performance on the reading assessment correlate with the overall general measure of motivation, based on the results of the Motivations for Reading Questionnaire (MRQ)?

Results of analyses for SRQ1 indicated a small but significant correlation between students' overall motivation score and their performance on the reading assessment (both accessible and proxy state assessments for all students except students with learning disabilities).

Secondary Research Question 2 (SRQ2): Do students perform better on the passages they rate as more interesting?

Results of analyses for SRQ2 indicated that the scores for students were higher when they rated passages as more interesting; this was the same for the accessible and proxy state assessments. This was generally true across grades, except for grade 8/9 students taking the accessible assessment who performed significantly higher on those passages they rated as more interesting.

Secondary Research Question 3 (SRQ3): Do special education and non-special education students rate informational-expository passages or narrative-literary passages as more interesting, or are both rated as equally interesting?

Results of analyses for SRQ3 indicated that passage type was related to interest regardless of whether students were SPED or Non-SPED. SPED students generally reported higher interest than Non-SPED students. Some differences by grade and group also were observed between the accessible and proxy state assessments.

Secondary Research Question 4 (SRQ 4): Does the text type in reading assessment passages (narrative-literary or informational-expository) impact student performance in general and by groups of students?

Results of analyses for SRQ4 indicated that in general, text type did affect student performance. For grade 4/5 students, students performed significantly better on expository-informational passages, regardless of whether they were in the SPED or Non-SPED group. For grade 8/9 students, passage type did not significantly affect performance on the assessment for either the SPED or Non-SPED group.

Secondary Research Question 5 (SRQ5): Do students score better on passages they rate as “kind of challenging” or “pretty challenging” in terms of how challenging (difficult) the passage was to read?

Results of analyses for SRQ5 indicated that students who took the accessible and proxy state assessments and rated passages as “kind of challenging” or “not too challenging” after they read them, performed better than when they rated passages as “very challenging” or “pretty challenging.” This was true for students at both grade levels and for both Non-SPED and SPED students.

Secondary Research Question 6 (SRQ6): Do students perform better on the *narrative-literary* passages when they indicated that they liked the passage because they knew something about it or when they indicated that they liked the passage because it was a story? How often did they mark both responses as reasons for liking the passage, and does doing so have an effect on performance?

Results of analyses for SRQ6 indicated that both SPED and Non-SPED students in grade 4/5 taking the accessible assessment performed better on passages they reported liking because they were stories. SPED and Non-SPED students in grade 8/9 performed better on passages that they both reported liking because they were stories and because they knew something about them. For grade 4/5 students and grade 8/9 students taking the proxy state assessment, performance on the reading assessment did not depend on their reasons for liking the narrative-literary passages.

Secondary Research Question 7 (SRQ7): Do students perform better on the *informational-expository* passages when they rated that they liked the passage because it provided them with information? Or did students perform better on the informational-expository passage when they rated that they liked the passage because they knew something about the passage? Or did they mark both responses as reasons for liking the passage?

Results of analyses for SRQ7 indicated that for the grade 4/5 accessible assessment, performance was not related to students’ reasons for liking the passages. For the grade 8/9 accessible assessment, Non-SPED students performed lower when they reported liking passages because it provided them with information and they knew something about it. For the grade 4/5 proxy state assessment, there was no significant effect for liking the passage. For the grade 8/9 proxy state assessment, Non-SPED students scored lower when they reported liking the passage because they knew something about it, an unexpected finding.

Secondary Research Question 8 (SRQ8): Do students who take the PARA accessible reading assessment more often use fewer accommodations than assigned to them than students who take the proxy state assessment?

Results of analyses for SRQ8 indicated that greater numbers of accommodations were used by students in grade 8/9 compared to students in grade 4/5. Differences emerged between the two assessments in the numbers of students using versus assigned to use the extended time accommodation only in grade 8/9. This accommodation was used by more students in the proxy state assessment compared to the accessible assessment, consistent with the original hypothesis that students taking the accessible reading assessment would use fewer accommodations than assigned to them.

Secondary Research Question 9 (SRQ9): Does the number of omitted or not reached items differ for the accessible and proxy state assessments? Does the level of students' motivation affect the number of omitted or not reached items.

Results of analyses for SRQ9 indicated that the number of not reached items was very low. Because of the low number overall and by test, additional analyses could not be conducted. Still, the low omission rates were quite impressive.

Discussion

Each of the findings is discussed in terms of current research and practice. In addition, implications for further research needs are identified.

Table of Contents

Executive Summary	iii
Introduction.	1
PARA Research That Informed the Field Test	4
Other Factors That Informed the Field Test	9
Design of PARA Accessible Reading Assessment	11
Field Test Purpose and Research Questions	11
Methodology	13
Tools and Materials Developed for PARA Field Test	13
Summary of Field Test Reading Assessments.	18
Recruitment Procedures	19
Study Procedures Prior to Administration	20
Field Staff.	22
Administration Procedures.	24
Planned and Implemented Field Test Designs.	25
Results.	28
Primary Research Questions	28
Secondary Research Questions	33
Discussion	62
References	75
Appendices.	83
Appendix A: PARA Reports and Publications	85
Appendix B: Summary of Disability Categories	87
Appendix C: NAEP 2009 Text Types	95
Appendix D: Text Type and Difficulty Rating of Passages in the Accessible and Proxy State Reading Assessments	97
Appendix E: Visual Analysis Criteria (VIA)	99
Appendix F: NAEP 2009 Cognitive Targets and Percentages Outlined in 2009 NAEP Reading Framework	101
Appendix G: Sample Multiple Choice Item from PARA Accessible Reading Assessment	103
Appendix H: Test Layout Criteria.	105
Appendix I: Segmenting Criteria for PARA Accessible Assessment	107

Appendix J: Motivations to Read Questionnaire (MRQ) Scoring Information and Research	109
Appendix K: Assessor Training Manual	111
Appendix L: MRQ Administration and Logistical Instructions, Including Frequently Asked Questions and Answers	177
Appendix M: Equipercentile Test Equating Procedures Used to Equate Across Testing Times, Including Computational Approach and Conversion Tables	183
Appendix N: Analyses of Primary Research Questions for Scores Equated Separately for the Accessible and Proxy State Assessments	191

In 2004, the National Center on Educational Outcomes (NCEO) and the Department of Curriculum and Instruction at the University of Minnesota, along with its partners at the University of California Los Angeles, University of California Davis, and Westat were awarded a grant to develop and conduct research on a reading assessment intended to be accessible for students with disabilities. The project was called the Partnership for Accessible Reading Assessment (PARA). The culminating goal of the project was to develop and field-test instruments or methods for assessing reading proficiency that are suitable for large-scale administration for school accountability purposes. This assessment was designed with the intent that it would be more accessible than previous assessments for students who had disabilities that affect reading while maintaining the validity and comparability of the resulting scores.

This report summarizes the development and field testing of the assessment developed by PARA. Throughout the report, we use the terms “special education students” and “non-special education students” to refer to the two primary groups of students included in the field test. We make this distinction because the term “students with disabilities” is a broad term that may include both students with Individualized Education Programs (IEPs) who receive special education services and students with accommodation plans under Section 504 of the 1973 Rehabilitation Act Section 504 as amended. Section 504 students have disabilities that do not necessarily require that the student have an IEP. We use the term “non-special education students” to indicate that the comparison group of students do not receive special education services. (See www.nldline.com/iep_vs_504.htm for additional information on the distinctions between special education and 504 students.)

Introduction

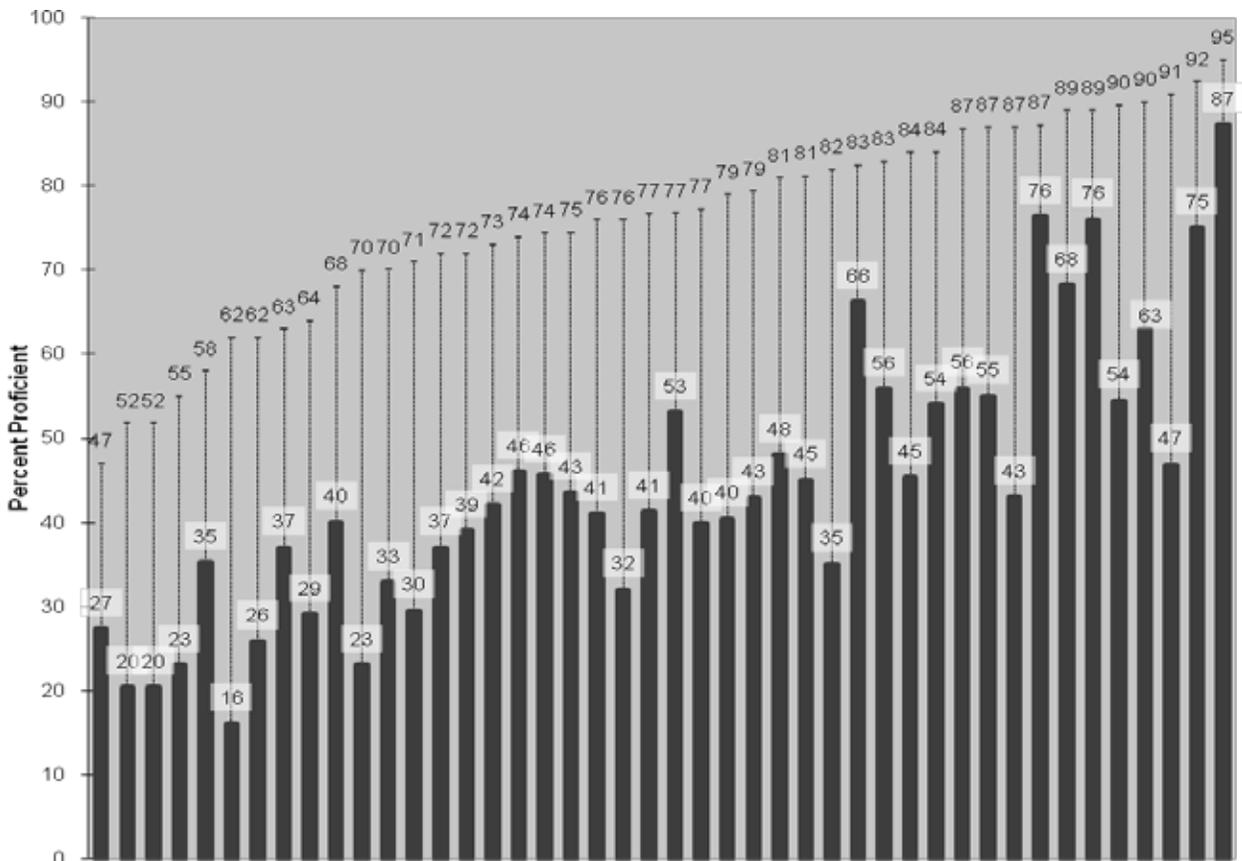
Learning to read and the demonstration of reading proficiency are cornerstones of educational success (Cunningham & Stanovich, 1997; Duke & Carlisle, 2011; Entwisle & Alexander, 1999; Pretorius, 2000; Slavin, Lake, Chambers, Cheung, & Davis, 2009) as well as other outcomes (DeWalt, Berkman, Sheridan, Lohr, & Pignone, 2004). Test results in recent years, however, have shown that an alarming number of students have difficulty mastering the reading skills tested (Chudowsky, Chudowsky, & Kober, 2009; Thurlow, Bremer, & Albus, 2011), even though there are some positive trends in reading on state tests overall since 2005 (Chudowsky & Chudowsky, 2010).

In 2003 the National Assessment of Educational Progress (NAEP) determined that 71% of students with disabilities in grade 4 were reading below a basic level of reading achievement compared to 35% of students without disabilities (National Center for Education Statistics, 2003). Despite statistically significant increases in the performance of students with disabilities on the NAEP reading assessment (National Center for Education Statistics, 2009), their performance continues to be much lower than the performance of students without disabilities. Nationally 66% of students with disabilities and 31% of students without disabilities were below basic on the 2009 NAEP reading

assessment in grade 4, while 78% of students with disabilities and 37% of students without disabilities were below basic on the 2009 NAEP reading assessment in grade 8. Gaps of less than 30 percentage points between students with and without disabilities at the basic or above level were seen in only 8 states in grade 4 and only 3 states in grade 8.

Similar results are seen in individual state assessments of reading (see Figure 1). For grade 4 reading tests in 2008-09, 29 of 44 states (66%) showed achievement gaps of 30 percentage points or more between non-special education students and those receiving special education services (Thurlow, Bremer, & Albus, 2011). For grade 8 reading tests in 2008-09, 40 of 46 states (87%) showed achievement gaps of 30 percentage points or more.

Figure 1. Elementary School Reading Performance on the Regular Assessment (2008-09)



Source: Thurlow, Bremer, & Albus, 2011.

Legend: Heavy Solid Bar = Students with disabilities percent proficient.

Dashed Line = Gap between students with disabilities and regular students. For some states the “regular students” comparison group may include students with disabilities because states report data differently.

Although students with various special education categorical labels have disabilities that affect reading, most research has focused only on those with learning disabilities. These students have been found to be challenged more than students without disabilities by text organization and text structures (Anderson and Armbruster, 1984; Meyer, Brandt, & Bluth, 1980). Less skilled, less fluent, and less practiced readers have difficulty perceiving, understanding, and negotiating major structures that organize expository texts (e.g., description, temporal sequence, explanation, compare/contrast, problem solution) (Gersten, Fuchs, Williams, & Baker, 2001). These less skilled readers are equally challenged in understanding the story grammars that organize narrative texts into familiar or predictable sequences of events, plot structures, character motives, and expected outcomes or resolutions.

The challenges of juggling cognitive subprocesses across a range of text types are amplified by the inclusion of increasing numbers of items on standardized assessments that require higher orders of inferencing (Fritschmann, Deshler, & Schumaker, 2007; Partnership for 21st Century Skills, 2006). This focus on a deeper cognitive demand is reflected in the cognitive targets of the NAEP 2009 Framework (National Assessment Governing Board, 2008)—a shift from assessing literal recall to more higher order cognitive processing.

Along with the inclusion of special education students in large-scale assessments has come the need to provide assurances that these assessments allow students to show their knowledge and skills rather than the effects of their disabilities. Thompson, Johnstone, and Thurlow (2002) and Thompson, Thurlow, and Malouf (2004) described ways in which the principles of universal design in architecture could apply to assessments, resulting in an approach called universal design of assessment (UDA). This approach can be incorporated to either remove assessment barriers without requiring accommodations or make testing accommodations easier to apply.

When assessments provide an accurate picture of students' knowledge and skills, and a group of students demonstrate low performance levels, test results can indicate areas in which test takers need additional instruction to improve their learning. However, when there are features of assessments that prevent students with disabilities from accurately demonstrating their knowledge and skills, the test results have little utility in guiding instructional efforts. In this case, the assessment needs to be improved to portray students' knowledge and skills more accurately.

Without paying attention to the extent to which assessments are accessible, the capabilities of a reader may elude assessment due to the inability or impracticality of assessing a range of intertwined subprocesses that contribute to reading comprehension (Stanovich, 1980, 1986, 1988; Vellutino, 2003). For example, accessible assessments maintain the construct of reading comprehension without privileging any one subprocess like decoding, encoding, lexical access, attention to text structure, and surface or deep level fluency that might impede less skilled readers. Similarly, accessible assessments address disengagement that may occur among students who otherwise can

demonstrate the necessary subprocesses. These students' reading ability is inaccessible because they are unlikely to engage in reading tasks due to years of frustration and negative perceptions of reading tasks typified in standardized assessments. In short, the ability to demonstrate what one knows and can do within the construct of reading involves the complex interaction of factors ranging from differences in knowledge, skills, and abilities that underlie word recognition, to language comprehension, to dispositions such as readers' motivation, goals, and purposes (Vellutino, 2003). The purpose of an accessible reading assessment is to yield more accurate, valid results in spite of these factors.

PARA Research That Informed the Field Test

PARA personnel engaged in a program of research to explore the definition of reading proficiency, study factors that affect achievement on large-scale assessments of reading, and develop (in collaboration with other projects) principles and guidelines for accessible reading assessments. (See Appendix A for a list of reports on these aspects of the project.) The purpose of conducting and gathering research evidence was to define the nature of the accessible assessment that would be explored through a large-scale field test. This research basis is summarized briefly here.

Definition of Reading Proficiency. In collaboration with the National Accessible Reading Assessment Projects (NARAP), three principles for defining reading proficiency were identified (NARAP, 2006):

- (1) Definitions of reading proficiency must be consistent with core NCLB provisions;
- (2) Reading proficiency must be defined in such a way that flexible expressions of reading are allowed while preserving the essential nature of reading; and
- (3) Definitions of reading proficiency must reflect both comprehension and foundational skills.

These principles were considered during the development of PARA's accessible assessment.

Exploring the Characteristics of Special Education Students. To develop an assessment approach that minimizes the need for accommodations or makes assessments more appropriate for accommodations, PARA researchers engaged in two activities. The first explored the characteristics of special education students with specific disability category labels. The second studied the characteristics of special education students whose performance on state reading assessments did not reflect their reading skills.

The first effort (Thurlow, Moen, Liu, Scullin, Hausmann, & Shyyan, 2009) was a literature review on the characteristics of students with specific learning disabilities, speech-language impairments, intellectual disabilities, emotional-behavioral disabilities, autism, deafness/hard of hearing, and visual impairments/blindness. In addition to

describing the students, it explored reading instruction and assessment implications for these students. (See Appendix B for a brief summary of reviews for all of these disability categories.) This work reinforced the need for research to identify ways in which assessments could be made more accessible for special education students, while preserving what the assessment is intended to measure. It also helped the project determine that it would focus its field test on students with learning disabilities, speech-language impairments, intellectual disabilities, and deafness/hard of hearing, categories that each show challenges for students demonstrating their knowledge and skills on reading assessments.

Figure 2 shows the prevalence of all disability categories in the U.S. in 2008. Students with specific learning disabilities comprised 42.9% of all students with disabilities (Data Accountability Center, 2010). Although the percentage of students in this category has decreased during the past several years (for example, the percentage was 45.5% in 2008 and 51% in 1999), these students still comprise the largest single category of students who receive special education services. Students with learning disabilities are followed in prevalence by students with speech/language impairments (19.1% of all students with disabilities) and students with intellectual disabilities, formerly referred to as students with mental retardation (8.1%). Among the less prevalent categories and for which reading is a significant challenge, are students with deafness and hard of hearing (1.2%). Students in all of these categories are challenged in various ways in demonstrating their knowledge and skills in reading.

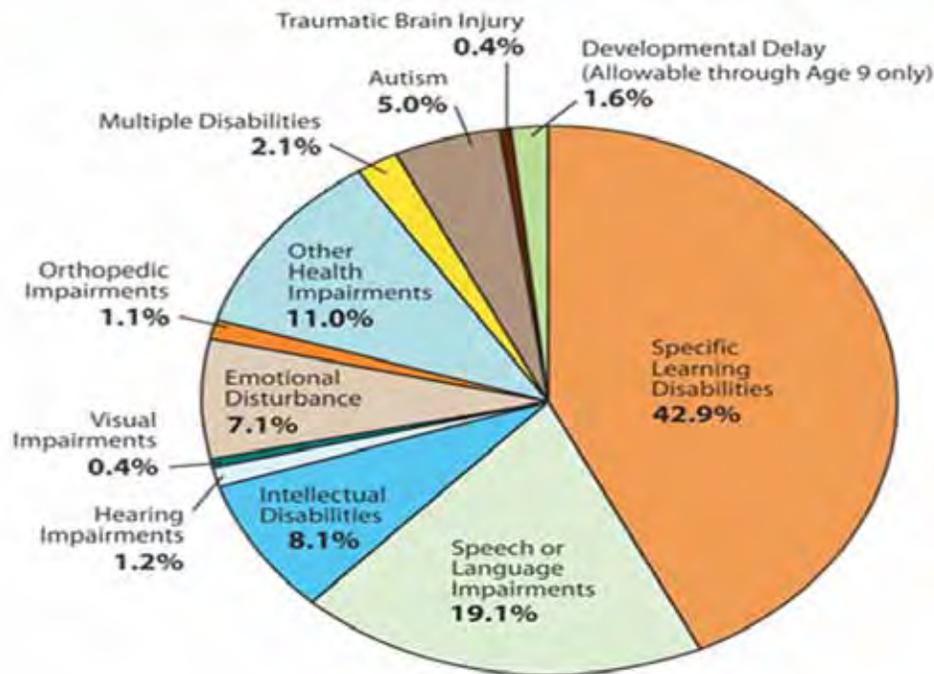
The second effort (Moen, Liu, Thurlow, Lekwa, Scullin, & Hausmann, 2009, 2010) was a study that obtained information from both students and their teachers in the state. This study found that factors other than simply a student's ability to decode or comprehend text could affect the student's performance on a state reading assessment. Another factor in low performance on the state assessment was affective, such as anxiety or low motivation. Most students appeared to have multiple barriers, and the effects of barriers were likely interconnected.

Exploring Existing State Assessments. An early PARA study of the content of states' reading assessments, as specified in their test blueprints and specifications, indicated that comprehension and analysis/interpretation items occurred approximately two times as frequently as foundational skills items (those that addressed vocabulary, word identification, word analysis, or fluency) (Johnstone, Moen, Thurlow, Matchett, Hausmann, & Scullin, 2007; see also Johnstone, Thurlow, Thompson, & Clapper, 2008). This confirmed the importance of having the PARA accessible assessment focus on comprehension and interpretation, and led to the exploration of the NAEP 2009 framework (National Assessment Governing Board, 2008) as a basis for the focus of PARA's assessment.

Three studies explored the performance of students with disabilities on existing state assessments (Abedi, Leon, & Kao, 2007a, b; Kato, Moen, & Thurlow, 2007, 2009). They examined how students respond to item distracters and explored differential item

Figure 2. Distribution of Students with Disabilities Receiving Special Education Services

Students Receiving Special Education Services by Disability Category



Source: www.IDEAdata.org
 Part B Child Count (2008). Students ages 6-21, 50 states, DC, PR, BIE schools

functioning, differential omission frequency, and response characteristic curves. The studies suggested that fatigue or frustration may play a role in the performance of students with disabilities on state assessments. The authors indicated that this conclusion was prompted by the finding that students did not complete all items and that they seemed to be giving random responses. They did not define “fatigue” or “frustration,” but noted the importance of conducting additional research. The authors also suggested that it is important to examine the extent to which there may be differential rates of omissions. These findings led to a study of the impact of segmented text and increasing motivation as ways to address fatigue and the tendency to not respond to all items.

PARA Intervention Studies. Four studies examined the effects of various approaches to increasing the accessibility of reading assessments. Two of these studies examined motivational factors in accessibility. Others examined the effects of segmenting the text in reading passages and the use of a reading pen as a partial auditory accommodation.

Motivational Factors. Students with disabilities often view the task of reading multiple passages and answering comprehension items on a test as insurmountable due to

low self-efficacy (O'Brien & Dillon, 2008). This, in turn, affects the amount of effort, willingness to persevere, and the control these students execute on a task (Bandura, 1986; Pajares & Schunk, 2002; Purdie, Hattie, & Douglas, 1996). In an accessible assessment, the goal is to design into the task, and into the students' stance toward the task, a higher expectation for success and a higher assigned value to the assessment. Guthrie and Wigfield (2005) have noted that altering text characteristics, providing student choice and more control in the setting, establishing goals, and altering other constructs related to reading tasks can influence situational motivation, affecting the assessment outcome. Yet, even though motivation and engagement contribute to reading comprehension, it is rare to find investigations of how these constructs play out in experimental studies of instruction, intervention, or assessments (Guthrie & Wigfield, 2000; Guthrie, Wigfield, & Perencevich, 2004).

Dillon et al. (2009a,b) examined several motivational factors for students with disabilities that affect reading. They found evidence of the benefits of careful selection of reading passages that are of high interest to students, including the use of visual items such as color pictures in contrast to current approaches that result in neutral and sometimes boring reading passages. The importance of these factors was confirmed in a second study that also included the choice of reading passage as a motivational factor (Dillon, O'Brien, Thurlow, & Lekwa, 2011). These studies supported the careful selection of reading passages of interest to students, the generation of a set of criteria to guide the inclusion of pictures and graphics in reading passages, and the use of color photos or graphics to support the text.

The impact of choice as an important motivational factor for students in grades 4 and 8 had not been confirmed at the time that the design of the PARA accessible assessment for the field test had to be completed. Yet, later analyses from the quantitative portion of the Motivation study indicated that allowing students to have their choice of reading passage did not lead to higher reading assessment scores. Still, mean scores favored Choice, but indicated that further research with a larger sample size is needed. One additional observation of importance in the analysis: Non-parametric comparisons (Kruskal-Wallis) of standard errors for choice and no choice conditions between special education and non-special education students suggested that allowing those in special education to choose the passages they read for a reading comprehension test reduced the variability in error estimates of their scores. The standard errors for non-special education students remained relatively unaffected. This was the case for students in both grades. Although this is not the differential boost hypothesized for the intervention of allowing choice, it is an important form of evidence favoring a more accessible test. Reliability of scores is essential to their accuracy (Dillon, O'Brien, Thurlow, & Lekwa, 2011).

Segmenting Text. Abedi et al. (2009) explored the segmenting of reading passages in ways that did not violate the intended constructs to be measured. Specific criteria were used to determine where breaks would occur, the questions that were appropriate to place after sections, and the questions that had to be at the end because of their integrative nature.

The results of this study indicated a significant improvement in the reliability of results for students with disabilities and no change in the validity of results. Because reliability was improved, Segmented Text was selected to be a feature of the PARA accessible assessment.

Reading Pen. Thurlow, Moen, Lekwa, and Scullin (2010) examined the effects of a reading pen as a partial auditory accommodation for reading assessments when the pen is used only to pronounce words on demand. Findings of this study indicated a significant between-group effect (students without disabilities performed higher than special education students) but no significant within-group differences. Although there were no group differences between those students who used the pen and those who did not, for some students the use of the pen did produce a significantly higher score compared to their performance without the reading pen. Still, because it was difficult to determine which students would benefit from the pen, it was not included as part of the PARA accessible assessment.

NARAP Principles and Guidelines. PARA and its sister projects (DARA—Designing Accessible Reading Assessments and TARA—Technology Assisted Reading Assessment) collaborated to produce the document *Accessibility Principles for Reading Assessments* (Thurlow, Laitusis et al., 2009). The five principles in this document were the result of a process of first identifying principles drawn from research, existing standards, and the consensus of experts in educational assessment, and then subjecting the principles to a series of reviews. The principles, each of which is supported by implementation guidelines, are:

1. Reading assessments are accessible to all students in the testing population, including students with disabilities.
2. Reading assessments are grounded in a definition of reading that is composed of clearly specified constructs, informed by scholarship, supported by empirical evidence, and attuned to accessibility concerns.
3. Reading assessments are developed with accessibility as a goal throughout rigorous and well-documented test design, development, and implementation procedures.
4. Reading assessments reduce the need for accommodations, yet are amenable to accommodations that are needed to make valid inferences about a student's proficiencies.
5. Reporting of reading assessment results is designed to be transparent to relevant audiences and to encourage valid interpretation and use of these results.

These principles helped to guide the development of PARA's accessible assessment for the field test.

Other Factors That Informed the Field Test

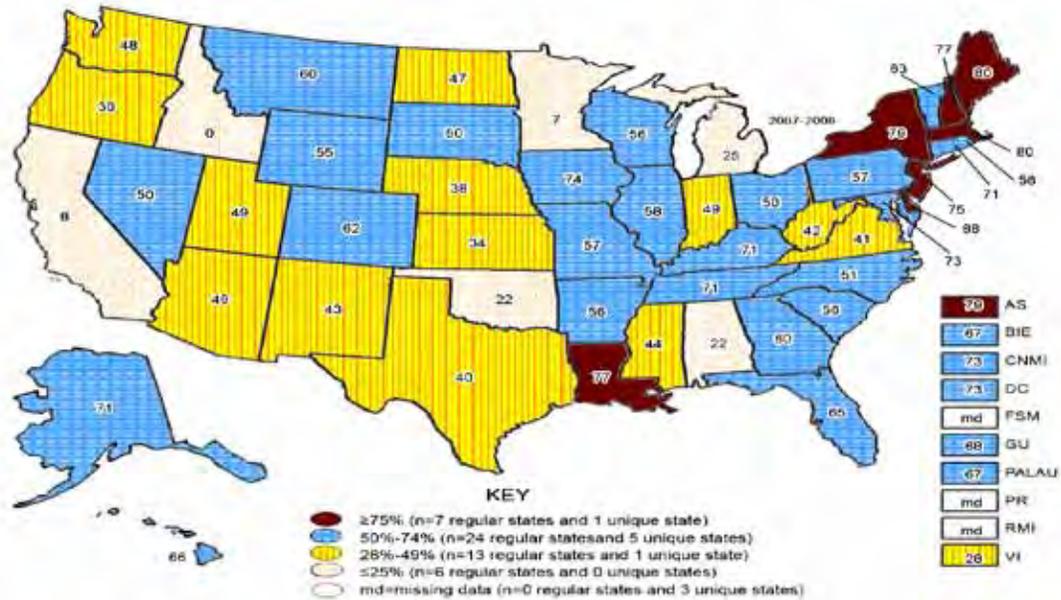
PARA researchers considered several other factors as they began the development of the accessible reading assessment. Specifically, they focused on universal design considerations for accommodation policies for reading assessments, and previous research on the effects of accommodations.

Universal Design Considerations. Universal design of assessment considerations were created by Thompson, Johnstone, Anderson, and Miller (2005). They highlighted ways to address each of the elements of universal design described by Thompson et al. (2002, 2004). All items developed for the assessments included in the PARA field test (both treatment and control) were reviewed using the universal design considerations.

Accommodation Policies for Assessments of Reading. Across the 50 states, the policies on accommodations permitted for reading assessments are varied, with some accommodations viewed as more controversial than others (Thurlow & Larson, 2011). The ways in which the test materials are presented generally are the most controversial accommodations (i.e., have the greatest variability in policies) for state reading assessments. Among the accommodations categorized as presentation accommodations, the four most controversial accommodations for reading assessments in 2009 were *read aloud passages*, *sign interpret passages*, *read aloud questions*, and *sign interpret questions*, each with from about 40% to 50% of states prohibiting the accommodation and the rest allowing them in some way or simply not having a specific policy.

Despite the controversy surrounding some of the accommodations used in reading assessments, there are many accommodations that are available to students with disabilities. The number documented in states' policies is over 70 different accommodations that involve presentation, response, equipment, setting, and timing/scheduling. Although states are not required to publicly report on the specific accommodations that students use, they are required to report on the total number of students receiving accommodations for each assessment. The percentages of students with disabilities using accommodations for state reading assessments in 2007-08 is striking, ranging in elementary school from just over 0% to 88% of students with disabilities (see Figure 3). Individual states indicate that some of the most frequently used accommodations are those that are most controversial.

Figure 3. Percentages of Special Education Students Using Accommodations for Elementary Reading Assessments in 2007-08



Source: Altman, Thurlow, & Vang, *Annual Performance Reports: 2007-2008 State Assessment Data* (www.nceo.info/OnlinePubs/APRreport2007-2008.pdf). Reprinted with permission from the National Center on Educational Outcomes.

For the PARA field test, the decision was made to follow the accommodation policies in the states in which the field test was conducted. The decision was also made to monitor the accommodations that the IEPs of special education students indicated they should receive during state assessments, as well as the accommodations that they actually used during the field test. It was posited that the need for accommodations might be less because of the greater accessibility of the PARA assessment.

Research on Reading Assessment Accommodations. Research on the effects of accommodations has increased over the years as states have developed and implemented their policies for assessment accommodations (see Cormier, Altman, Shyyan, & Thurlow, 2010; Johnstone, Altman, Thurlow, & Thompson, 2006; Thompson, Blount, & Thurlow, 2002; Zenisky & Sireci, 2007). A range of approaches to studying the effects of accommodations emerged as the importance of accommodations for state assessments was recognized (Thurlow, McGrew, Tindal, Thompson, Ysseldyke, & Elliott, 2000). Yet today the most common approach to studying the effects of accommodations is to look for a differential boost in performance (Laitusis, 2007). The differential boost hypothesis states that although an accommodation may have some effect on the performance of students who do not need the accommodation (generally interpreted to mean non-special education students), the effect for those students needing the accommodation must be greater than it is for those not needing it. This approach is consistent with

early recommendations that the “purpose of an accommodation is not to raise scores per se, but rather to make them fairer and more accurate. To do this, one must be able to identify the difficulties caused by the disability that are irrelevant to the construct measured and to design accommodations that will offset those specific difficulties” (Koretz, 1997, p. 76).

Although the research on accommodations has increased dramatically, not all of it has focused on reading assessments, or on students with disabilities. In an examination of studies published in peer review journals or defended in doctoral dissertations, investigators at NCEO have summarized the focus and findings of accommodations research. Research that focused on assessments other than regular assessments or that focused on ELLs were excluded (research involving ELLs with disabilities were included). In 2005-06 (Zenisky & Sireci, 2007), 14 of 32 studies (44%) focused on reading, whereas in 2007-08 (Cormier, Altman, Shyyan, & Thurlow, 2010), 18 of 40 studies (45%) focused on reading. Looking at just those studies published in 2008 (n = 11), only five included students with disabilities and the school levels varied considerably (1 elementary, 2 middle school, 1 high school, and 1 both elementary and middle school). The accommodations included in the research were segmented text, extended time, computer-based administration with directions read, and multiple accommodations at once (2 studies).

Based on the available research, the decision was made to incorporate segmented text as part of the accessible approach. Extended time and reading of directions were incorporated for all students, regardless of whether they participated in the accessible assessment or a comparison assessment called the “proxy state assessment.”

Design of PARA Accessible Reading Assessment

On the basis of its own and others’ research, PARA determined that its accessible assessment would incorporate three features: (a) segmented text; (b) high interest passages identified by students as most interesting during a calibration study; and (c) color. The “proxy state assessment” did not include these critical characteristics of the PARA accessible reading assessment. Details on the creation of the intervention assessments (accessible assessment) and control assessments (proxy state assessment) are provided in the Methodology section.

Field Test Purpose and Research Questions

The purpose of the PARA field test was to determine the degree to which its accessible reading assessment provided for accessibility, validity, and comparability for students with learning disabilities (LD), speech-language impairments (SLI), intellectual disabilities (ID), and deafness/hard of hearing (D/HH). To address these assessment qualities, the PARA accessible assessment was compared to a proxy state assessment. Four basic research questions were the primary focus of interest for the PARA field test:

1. Does the PARA accessible reading assessment produce higher performance than a proxy state reading assessment for students with LD, and for students with disabilities combined (LD, SLI, ID, and D/HH)?
2. Are scores from the PARA accessible reading assessment more reliable than scores from the proxy state reading assessment for students with LD, and for students with disabilities combined (LD, SLI, ID, and D/HH)?
3. Does the PARA accessible reading assessment produce higher performance of non-special education students, thereby suggesting that it alters the focal construct?
4. Does the PARA accessible reading assessment demonstrate the same characteristics (student performance, reliability) for students in grades 4 and 8?

In addition to the basic research questions, several secondary research questions were addressed:

1. Does students' performance on the reading assessment correlate with the overall general measure of motivation, based on the results of the Motivations for Reading Questionnaire (MRQ)?
2. Do students perform better on the passages they rate as more interesting?
3. Do special education and non-special education students rate informational-expository passages or narrative-literary passages as more interesting, or are both rated as equally interesting?
4. Does the text type in reading assessment passages (narrative-literary or informational-expository) impact student performance in general and by groups of students?
5. Do students score better on passages they rate as "kind of challenging" or "pretty challenging" in terms of how challenging (difficult) the passage was to read?
6. Do students perform better on the *narrative-literary* passages when they indicated that they liked the passage because they knew something about it or when they indicated that they liked the passage because it was a story? How often did they mark both responses as reasons for liking the passage, and does doing so have an effect on performance?
7. Do students perform better on the *informational-expository* passages when they rated that they liked the passage because it provided them with information? Or did students perform better on the informational-expository passage when they rated that they liked the passage because they knew something about the passage? Or did they mark both responses as reasons for liking the passage?

8. Do students who take the PARA accessible assessment more often use fewer accommodations than assigned to them than students who take the proxy state assessments?
9. Does the number of omitted or not reached items differ for the accessible and proxy state assessments? Does the level of student motivation affect the number of omitted or not reached items?

Each of PARA's primary and secondary research questions and our hypotheses are provided in the Results section.

Methodology

The PARA field test involved the administration of accessible and proxy state assessments to samples of students in grades 4 and 8, and comparison of the performance of special education and non-special education students. In this section, we provide information on the tools and materials developed for the study; recruitment of states, districts, and schools; a summary of the two approaches to reading assessment that were tested; study procedures prior to administration; field staff (including training); and administration procedures.

Tools and Materials Developed for PARA Field Test

Special passages and multiple choice items were developed for the reading assessment—both the “Proxy State Assessment” for the control group, and the “Accessible Reading Assessment” for the treatment group. The passages were selected and items developed based on specific criteria discussed here, and they were calibrated in a large-scale study conducted in 2008. (See Dillon et al., 2009a,b for the results of the calibration study.)

Text types used for the assessments. Passages on the proxy state and accessible assessments conform to the NAEP 2009 Reading Framework guidelines (National Assessment Governing Board, 2008) that specify several different text types that should be assessed on reading tests. The NAEP Reading Framework was adopted because it provides a national perspective of reading content standards rather than state specific content standards. (See Appendix C for the list of NAEP 2009 text types.) In designing field test assessments, we included two of the suggested text types: *informational-expository* and *narrative-literary*. These two text types require that students show their ability to read typical texts they encounter in school.

Motivational characteristics of the texts. To achieve the goal of using motivating texts, we first selected passages based on a review of research in children's and young adult literature that indicated topics of interest to students at the targeted grade levels, that is, fourth and eighth grades, and expert advice from children's literature scholars. We actively sought high-interest texts from popular children and youth publications such as *Cricket* magazine and *Ranger Rick* or short stories based on topics that adolescents

could relate to from real life situations. For example, passages included thought-provoking stories with themes that students could relate to (how to overcome fears) or informational texts based on unusual or weird events or animals (finding a dead body preserved in a bog). We also included the pictures, photos, and illustrations that accompanied the original passages. Selecting passages that were highly interesting over passages considered to be rated as high quality literature or from a traditional canon was done intentionally to first and foremost engage youth (Schiefele,1999) and influence the possible completion of the reading assessment with more valid results. (For information on motivating readers see O'Brien and Dillon, 2008.)

Passage lengths and readability criteria. The passages selected for the assessment were 500-700 words in length for 4th graders and 700-1000 words in length for 8th graders. Passage lengths were determined based on our review of the technical specifications for existing state assessments, thereby meeting the grant requirement to explore assessments that mirrored typical large-scale reading assessments. Our goal was to use intact passages or excerpted text from short book chapters. Finding passages of particular lengths that were intact and still retained a rich message was key to the ecological validity of the study. In a few cases, we carefully excerpted passages from longer pieces, but we did not rewrite any passage text. Next, we calculated or gauged the difficulty of passages by using standard readability measures including the Fry Readability Graph (Fry, 1977) and Lexiles (Lexile Framework, MetaMetrics, Inc). The Fry data and information from the Lexile analyses for each passage were examined by reading experts and a final determination of whether a passage was considered grade-level appropriate was rendered by them. In addition, reading experts assigned each passage a difficulty rating of easy, medium, or hard. This holistic rating was based on the readability analyses but included consideration of topic familiarity, concept load or the degree of difficulty of the ideas presented about a topic, and the extent to which the passage had accompanying pictures or visual items which might support a reader's understanding. We provided holistic ratings for the texts because the calibration study required that we assess students within a band of the two grade levels for which we were designing assessments. That is, we assessed students in grades 3-5 to develop the 4th grade reading test and we assessed students in grades 7-9 to develop the 8th grade reading test. (See Appendix D for a listing of the passages on the proxy state and accessible reading assessments listed by text type and difficulty rating.)

Criteria for including visual items with passages. Criteria were developed to make decisions about the inclusion or exclusion of visual items (e.g., photos, pictures, creative fonts, maps, sketches, illustrations, or diagrams) that accompanied each of the reading passages selected for the PARA assessment. Our goal was to retain the original visual elements (photos, pictures, creative fonts, maps, sketches, illustrations, or diagrams) that accompanied each passage because we believed these features would engage students, motivate them to read, and as others have found, support their comprehension (Brookshier, Scharff, & Moses, 2002; Levie & Lentz, 1982). The criteria we developed to include or exclude visuals were generated using the research literature, expert opinions from a noted scholar in the area of children's literature, reading professors, and expert

K-12 classroom teachers (Dillon O'Brien et al., 2009a,b). (See Appendix E for a listing of the visual analysis criteria.) In creating the criteria, we were also mindful of research findings that suggest that visuals can distract some students from the print message, compromising their comprehension (e.g., Rose, 1986). Pairs of reading researchers used the criteria with each passage to determine, through individual analysis followed by joint analysis/discussion, which visual items would be “retained” and which would be “deleted.” An example illustrates criterion #2—that the picture “...is important to *understanding the concepts discussed in the passage. If the author has included a picture or map to help readers understand a concept that is key to the passage or explained in the text, we will retain the visual item to support the reader’s understanding of the concept and thus their overall comprehension.*” An example of this criterion in action was when we retained a picture map showing how to make rice balls. The text described the process, but because most readers would not have prior knowledge of what rice balls are or the process of making them from scratch, we retained the pictures. Consensus was reached through individual analysis followed by group analysis and discussion. Agreement was almost 95% for all visual items prior to discussion between researchers and 100% after discussion. After passages were identified and analyzed, permission was secured from publishers and illustrators and/or photographers to use the texts in the PARA accessible reading assessment.

Multiple choice item development process. Seven multiple-choice items followed each passage; the items measured cognitive targets. We asked students to read passages and then to answer questions in which they were required to (a) Locate/Recall, (b) Integrate/Interpret, and (c) Critique/Evaluate. We used cognitive targets in percentages appropriate for each grade level. (See Appendix F for the list of NAEP 2009 cognitive targets and percentages of each type for particular grade levels as outlined in the NAEP Reading Framework.) One item for each passage on the assessment measured contextual vocabulary knowledge instead of measuring vocabulary words used in isolation. To create high quality multiple choice questions or items that followed each passage, we hired two professional item writers with experience in writing items for reading assessments. To ensure that they maintained high standards in writing items and consistency across their work and with other item writers across the nation, the item writers developed a code book that guided how they constructed and worded items. The item writers also worked as a team, analyzing each others’ items before the items were sent to research staff where the items underwent additional scrutiny from reading professionals at the university and public school levels for quality and potential bias. The item writers wrote a total of 10 items for each passage. As a result of the calibration process, the 10 items were reduced to the best 7 items for each passage (See Appendix G for a sample of a multiple choice item from the PARA accessible assessment.)

Test layout. Although there is little definitive evidence on how to best design the layout of test pages to make them more readable and accessible (Chall & Squire, 1996; Waller, 1996), we perused criteria from syntheses of research in readability and document analyses (Oakland & Lane, 2004) and universal design applied to large scale assessments

(Thompson, Johnston, & Thurlow, 2002) to inform our design and layout. (See Appendix H for a listing of the test layout criteria developed by PARA staff.)

Test passages were reproduced by digital layout experts working with assessment professionals so that the assessment passages looked as much as possible like the passages as they appeared in the original publications with just a few modifications. A two-column format was used to present the passages to assure that passages were consistently presented with a lack of clutter and ample white space per page (at variance from the column layout range found among the originals). We obtained graphics and photo files from publishers and placed these visual elements in the passages as they naturally occurred in the original texts if an item was retained after the visual analysis process. The 7 multiple choice items were arranged on the page to make the answer choices easy to read and to contrast the answer section with other text on the page. Each test booklet contained 4 reading passages, two informational-expository and two narrative-literary passages, each with 7 multiple choice questions. Students mark answers directly in the test booklet.

Situational motivation questions. After students completed the passages and multiple choice questions for each passage, they answered generic 4-point Likert-scale questions that required them to rate each passage on these criteria: (a) how interesting the passage was; (b) how challenging the passage was to read; (c) whether they liked the passage because it was a story (narrative texts), or they learned new facts (expository texts), and whether they knew something about the passage (a rating of prior knowledge after reading the passage). The layout of the test allowed us to place the situated motivation questions into a special box so it was clear that they were different from the multiple choice items that followed and focused on the specific content of the passage.

Because the Likert items immediately followed each passage, students were able to remember passage content well enough to rate how interesting or challenging the material was before they continued on to the next passage. Still, the two brief questions were not so intrusive as to slow down momentum needed to complete the test. The following are the passage interest and difficulty scales:

- 4--This passage was very interesting to read.
 - 3--This passage was pretty interesting to read.
 - 2--This passage was kind of interesting to read.
 - 1--This passage was not at all interesting to read.
-
- 4--This passage was very challenging to read.
 - 3--This passage was pretty challenging to read.
 - 2--This passage was kind of challenging to read.
 - 1--This passage was not at all challenging to read

Questions about whether they liked the passage because it was a story (for the narrative texts) or they learned new facts (for the expository texts), and whether they knew

something about the passage (a rating of prior knowledge after reading the passage) were included. These questions helped us answer research questions about the role of students' motivation via their interest in a passage after reading it, their perceptions of the difficulty of a passage, and their thoughts about learning from various text types in comparison to their performance on the comprehension items following a passage.

Segmenting of the Passages. The segmenting of passages criteria developed for the PARA field test draws on the work of Abedi and colleagues (Abedi et al., 2009) which focused on dividing up a reading passage into manageable segments or chunks of the story to address fatigue/frustration and lack of completion of assessments that was more common among students with disabilities. The field test was built on and in some cases adapted Abedi and his colleagues' procedures. The new criteria were influenced by reading research in the area of text structures (e.g., Meyer et al., 1980; Rumelhart, 1975), such as top-level structure analysis with narrative text and micro and macro propositional analysis. Because the two text types (informational and narrative) have different structures, they require several different segmenting criteria. Finally, the new criteria and the subsequent analyses of the reading assessment passages for the field test were also based on an additional idea: The introduction of breaking long reading passages into several segments by *dividing the texts in places that create more motivation for students to want to continue to read remaining portions of the passage*. Specifically, we divided the highly interesting narrative texts by identifying naturally occurring "cliff hangers." (See Dillon & O'Brien, 2009, for a discussion of motivation factors impacting this work.) The rank-ordered list of criteria, developed for use in segmenting the passages for the PARA accessible assessment, is broken into four categories including overall required procedures for segmenting (these were used with all passages), preferred procedures (used with all passages if at all possible), and specific procedures for both narrative and expository texts. The passages were divided into 2 or 3 segments for both grade 4 and grade 8. (See Appendix I for the Segmenting Criteria for the PARA accessible assessment.)

Motivations to Read Questionnaire (MRQ). The MRQ (Wigfield, Guthrie, & McGough, 1996) consists of 54 items that measure different aspects of elementary school-age children's reading motivations. The 54 items are divided into 11 dimensions or constructs of motivation. These include: Reading Efficacy, Reading Challenge, Reading Curiosity, Aesthetic Enjoyment of Reading, Importance of Reading, Compliance, Reading Recognition, Reading for Grades, Social Reasons for Reading, Reading Competition, and Reading Work Avoidance. These constructs are organized and reported out after analysis into 5 subscales: Reading Efficacy, Reading Challenge, Reading Curiosity, Aesthetic Enjoyment of Reading, and Importance of Reading and one subscale we elected not to report on because it was not relevant to the study (Compliance). The MRQ can be used to discover the nature of children's motivations for reading, as well as some of the things about reading children do not find motivating. Examples of items include: *I like hard, challenging books. I read to learn new information about topics that interest me.*

The MRQ can be administered in 20-25 minutes. This assessment was placed into a survey tool that students generally completed prior to (but on a different day) from the reading assessment to which they were assigned (either accessible or proxy state). In a few cases, students completed the MRQ on the day after the reading assessment was administered. Scoring information for the items as well as the research undergirding the MRQ is described in Appendix J.

Summary of Field Test Reading Assessments

The accessible assessment and the proxy state assessment were similar in several aspects, but differed in the use of segmented text, higher interest passages, and color illustrations. Table 1 shows the similarities and differences in the assessment conditions.

Table 1. Features of the Accessible and Proxy State Assessments

Accessible Assessment	Proxy State Assessment	Both Assessments
<ul style="list-style-type: none"> Segmented Text Passages Rated as Most Interesting Color Visual Elements 	<ul style="list-style-type: none"> No Segmented Text Passages Rated as Least Interesting Black and White Visual Elements 	<ul style="list-style-type: none"> Four passages with seven questions each (2 informational-expository and 2 narrative-literary) Interesting Passages Only Visual Elements that Met Criteria for Inclusion Ample White Space Two Column Format Clear Answer Choices No Time Limits Accommodations

Accessible Assessment. As noted in Table 1, the accessible form of the assessment contained 4 passages (2 informational-expository texts and 2 narrative-literary texts) that students rated during a large-scale calibration study as the *most* interesting of all texts to read (Dillon et al., 2009a,b; Dillon et al., 2011). The passages were arranged in the booklet with an expository passage first, followed by a narrative passage, then an expository passage, and then another narrative passage. Passage order was the same for each booklet. Besides including the most interesting passages, two other features were unique to the PARA accessible assessment:

- Each passage contained selected visual items from the original texts (if the visual elements existed) presented *in color*.
- *Passages were each divided into sections or segments* with questions interspersed after each section to help students read and understand the passage better as they progressed through the assessment.

This is the first time that the two features (segmenting and the use of student-rated high interest passages) were combined into one assessment.

Proxy State Assessment (non-accessible). The proxy state assessment also contained 4 passages (2 informational-expository texts and 2 narrative-literary texts). The passages were ones that students rated as the *least* interesting of all texts read during a previous study (Dillon et al., 2011). Passages were arranged in each booklet with the same passage type order as the Accessible Assessment (expository passage, narrative passage, expository passage, narrative passage). Besides including the *least* interesting passages, two other features were unique to the PARA proxy state assessment:

- Each passage contained visual items from the original texts (if they existed) but they were presented in *black and white* (not in color)
- Passages were *not divided up into segments*. Instead, the entire passage was presented, followed by all 7 multiple choice questions.

Recruitment Procedures

State Recruitment. States identified for recruitment in the study were large states with relatively small English Language Learner (ELL) populations. ELLs were not included in the study, even if they were also receiving special education services. ELLs were excluded from the sample to avoid possible confounding effects of language with the content, and the impact of such confounding on their assessment results. States with these characteristics were identified. Project personnel then contacted the state director of assessment and the state special education director about participation in the study. The project director spoke with the directors in each state to discuss the details of the study. The parameters of the study were clarified, including the focus on students in the regular assessment (not the alternate assessment based on alternate achievement standards), and the exclusion of ELLs.

The field test details also included the timing of the study. The ideal testing time was identified for each state; in both states the desire was to not conflict with statewide testing schedules. Original plans were to implement the field test in January in one state and in March in the other. In each state, contact personnel were identified; these individuals helped with IRB processes and recruitment of districts and schools.

Because of delays in IRB approvals and identification of schools with the targeted groups of students (students with learning disabilities, speech/language impairments, intellectual disabilities, and deafness/hard of hearing), timelines had to be shifted, with testing conducted in April and May. Insufficient numbers of students were obtained during that time, so the field test was also implemented in the fall with students in grades 5 and 9. These grades were chosen in an attempt to keep the knowledge level approximately the same for students in the spring and fall administrations. No changes

were made in the content of the assessments for the fall administration. Students in the spring and fall testing were from different schools.

Recruitment of Districts and Schools. For the spring data collection, Westat recruited school districts in the two states that agreed to participate in the field test. Both states provided us with the numbers of students with disabilities in 4th and 8th grades for each of the 4 disabilities to be assessed. Districts with the largest counts of students with learning disabilities, speech/ language impairments, intellectual disabilities, and hearing impairments in 4th and 8th grades were selected to be recruited.

The Superintendents of these school districts were contacted by letter and follow-up phone call to request approval for us to contact their schools. After obtaining approval from the Superintendent of each district, we selected schools to be recruited using the same criterion as we used to recruit districts. That is, we selected the schools with the largest numbers of students with learning disabilities, speech/language impairments, intellectual disabilities, and hearing impairments in 4th and 8th grades. We then sent a recruitment letter to the principals of these schools describing the study and requesting their participation. Then we contacted the principals by phone to obtain their decisions. Nine schools in State 1 and 18 in State 2 agreed to participate in the spring data collection.

For the fall data collection, we used the same procedures, except that we targeted 5th and 9th grade students. We obtained agreements to participate in the fall data collection from 22 schools in State 1 and 11 schools in State 2.

For both the spring and fall data collections, there were two primary reasons that districts or schools declined to participate. First, the timing of the field test conflicted with various activities at the beginning or end of the school year, for example, state testing. Second, districts or schools maintained that students with disabilities were tested too often. Smaller numbers of districts and schools indicated that they would not participate because too much staff time would be needed to participate and there were other research projects going on at the same time.

All participating schools identified a school contact to assist with the collection of MRQ data, as well as to assist in the identification of potential participants. The school contact was provided an incentive of from \$200 to \$500, depending on the number of students, to cover the time needed to complete these tasks. Participating students were provided an incentive of a gift card; 4th graders received a \$15 gift card and 8th graders received a \$25 gift card.

Study Procedures Prior to Administration

Five tasks were completed before the MRQ and the reading assessment could be administered in participating schools. The same procedures were completed prior to the spring and fall data collections. The only difference between the collections was the

numbers of non-special education students sampled to participate in the assessments. (See below for further explanation.)

First, principals who agreed to participate were asked to designate a school contact to carry out a variety of tasks to assist us in implementing the study. The school contact collected background data on each student and coordinated, with Westat, the administration of the MRQ survey and the reading assessment.

Second, the school contact gathered information about students who would be eligible to participate. For all students, these data included the student's name, gender, and ethnicity. For special education students, their primary disability and necessary test accommodations were also provided. To facilitate the collection of these data, we sent the school contacts blank rosters, and they provided these data to Westat prior to the administration of the MRQ so that we could generate a sample of non-special education students to participate in the study. (A copy of the roster and accompanying instructions can be found in the assessor training manual, which is included as Appendix K of this report.) In addition, the school contacts provided Westat with the sampled students' reading scores for the most recent reading assessment and classroom performance in reading. The latter was a rating provided by each student's reading/language arts teacher as to whether the student was reading below grade level, at grade level, or above grade level. For the fall collection, school contacts also provided students' mathematics scores from the most recent state assessment.

Third, the selection of students to participate in the assessment occurred. For each disability category, all of the students in each participating school were selected to participate in the assessment. For general education students, the number of students sampled varied depending on the state. For the Spring 2010 data collection in State 1, 30 4th graders and 43 8th graders were selected per school; in State 2, 50 students were selected per school regardless of grade (unless a school had both grades, in which case 25 of each grade were selected). For the Fall 2010 data collection in both states, 35 students were selected per school each in grades 5 and 9 (no schools had both grades 5 and 9). The selection of non-special education students was conducted randomly using a database application. Rosters of all general education students in the targeted grades were collected from schools, in any format they preferred to send. Because entering all students in a grade would have been time consuming for large schools, the total number of students in the grade was entered into the database application. Whole random numbers between 1 and this maximum number were generated by the application and displayed as slot numbers. Data entry staff keyed the name (or ID) of the student in that slot or position on the roster into the application to the sample of students.

Fourth, after receiving the rosters and randomly selecting non-special education students to participate, students were randomly assigned one of two conditions: the PARA accessible assessment, or the proxy state assessment (see description of these under "Summary of Field Test Reading Assessments"). Random assignment was accomplished via the database; as students were entered into the database, the data entry application

assigned a random number (between 0 and 1 exclusive) to each. To divide the students evenly into two groups, students with random numbers greater than or equal to .5 were assigned to the proxy state assessment. Those students with numbers below the threshold were assigned to the PARA accessible assessment.

Fifth, student-level recruitment began by sending parent/guardian letters to the school for students to take home to their parent/guardian. Parents could decline to have their child participate in the study. The informational letter included a description of the study, its purposes, and the procedures for granting consent or refusal. Parents who did not want their child to participate were asked to indicate their refusal for their child by calling an 800 telephone number. If a voicemail was reached, the caller was requested to leave detailed information, including their name, their child's name and grade, and the school the child attended. All students whose parents refused participation were excluded from the data collection and field test.

Field Staff

This section provides information on how we hired the individuals to administer the reading assessments, hereafter called the “assessors,” how they were trained, and the procedures they used to administer the assessments.

Hiring of assessors. The following skills and experience were required for all PARA assessors:

- A bachelor's degree in education/administration or psychology,
- Professional education experience,
- Good communication skills,
- Detail oriented,
- Successful completion of the training program,
- Passing a standard Westat background check as well any other background checks required by the school districts, and
- Prior experience administering an educational assessment in a school setting either for Westat or another research firm.

Each of the PARA assessors had to follow four basic principles in collecting PARA data. These are (1) collecting data according to study procedures; (2) observing high standards of personal conduct; (3) adhering to study protocols and procedures; and (4) maintaining the privacy of respondents and the confidentiality of individual data collected.

All assessors for this study signed a confidentiality agreement to demonstrate their commitment to the principle of the privacy of respondents and data obtained during the data collection. This agreement required that they keep completely confidential the names and addresses of respondents, all information or opinions collected in the course of interviews, and any information learned incidentally about individual respondents, responding organizations, or the places and organizations where respondents work and live. They were also to exercise care to prevent access by others to data in their possession.

Training of assessors. Westat developed in-depth extensive training that all PARA assessors had to complete before they began to administer the assessments. The training was designed to ensure that data collectors were knowledgeable about the purpose of the study and the data collection tools that were being used. All assessors hired for the PARA study had extensive teaching and assessor experience and had the skills needed to successfully administer the field test assessments. Test administrators were recruited from Westat's extensive pool of experienced test administrators.

A one-day training was held in each state for individuals who would be administering the field test assessments. The training was conducted by the co-principal investigators, the field supervisor, and staff from Westat. It included both (a) presentations on the study materials, and (b) interactive exercises designed to familiarize the test administrators with the assessment. Test administrators were provided with hands-on experience with the assessment booklets and were encouraged to ask questions throughout the training.

The topics of the presentations at these trainings were:

1. Overview of the PARA project including the study sample, the study design, and the activities that took place prior to the data collection;
2. Data collection procedures including confidentiality issues;
3. Administering the assessment;
4. Testing accommodations;
5. Overview of the data collection materials; and
6. Professional conduct.

The assessors were provided a training manual that addressed all of these issues along with frequently asked questions on general testing issues and accommodations for students with disabilities. A copy of the manual is presented in Appendix K.

A field supervisor worked closely with the assessors during the data collection period to resolve any data collection challenges that occurred. Each assessor had a weekly scheduled phone call with the field supervisor during which he or she reported on the

status of his or her assignment. In addition to the scheduled calls, the field supervisor was available by phone whenever a question or issue needed immediate attention.

Administration Procedures

Administering the MRQ. The MRQ was administered to groups of students via a computer-based platform using an online survey tool. The school contact helped students gain access to computers and the internet to take the MRQ. School contacts were provided MRQ administration and logistical instructions as well as a list of frequently asked questions and answers. These are included in Appendix L. (Note that this information was not included in the assessor training manual because assessors did not administer the MRQ.) The administration of the MRQ was the same for the spring and fall data collections.

Administering the reading assessment. Administration of assessments was the same for the spring and fall data collections. Prior to administering the reading assessment, students had to consent to their participation in the study. For both the spring and fall assessments, each 4th/5th grader was given a student assent form. They could participate only if they affirmed verbally to the assessor that they understood the purposes of the assessment. The 8th/9th graders could participate only if they read and signed the consent form confirming that they understood purposes of the assessment.

After the assessors obtained consent from students to participate, they administered the assessment. They:

1. Read the sample items pages to the students,
2. Collected the sample items from the students,
3. Read the directions for the assessment to the students,
4. Began testing, and
5. Marked the start time to record on test booklets.

Students were instructed to close their test booklets when they completed the assessment. The assessor then reviewed these students' test booklets for completeness and prompted the child to finish, if necessary. Note that a test was considered to be complete if a student attempted work for 30 minutes. The assessor then marked the start and end time on the booklet. Students received their gift cards after they signed a receipt for the cards. Finally, the assessor completed an accommodation form for students with disabilities (see below for the accommodations available to students) or an assessment form for general education students. A copy of the accommodation form is included in the assessor training manual (see Appendix K).

Accommodations for special education students. For this study, special education students were provided the accommodations they received when participating in standardized state assessments. School contacts listed the accommodations needed by each student on the school roster. The contact made the arrangements needed to provide these accommodations. The following accommodations were available to special education students:

1. Read aloud directions,
2. Sign interpret directions,
3. Repeat/paraphrase directions,
4. Magnification equipment,
5. Proctor/scribe,
6. Extended time (Students were allowed up to 2 hours for this assessment),
7. Breaks,
8. Individual testing, and
9. “Other” accommodations.

Planned and Implemented Field Test Designs

The PARA field test was designed with careful consideration of optimal conditions for the implementation of an experimental study design. Due to several factors identified previously (smaller numbers of participants than planned, need to conduct the field test during two different school years), the implemented field test design was different from the planned field test design.

Planned Field Test Design. PARA was scheduled to start its field test in January, 2010, first in State 1 and then in State 2, with all data collection to be completed by April, 2010. Special education students with learning disabilities, speech/language impairments, intellectual disabilities, and deafness/hard of hearing, as well as non-special education students, were targeted for participation. In addition to taking the assessments (either the accessible or the proxy state assessment), information was collected on each student’s characteristics (gender, age, reading level), most recent state reading test score, and responses to a motivation survey. Students selected for participation were randomly assigned to either the accessible assessment or the proxy state assessment condition.

The original field trial design required a total of 1,500 students in grade 4 and 1,500 students in grade 8. The desired distribution of students for the field test is shown in the Table 2.

Table 2. Planned Distribution of Students for the PARA Field Test

Student Group	Accessible Assessment	Proxy State Assessment
Grade 4		
Special Education Students		
Students with LD	150	150
Students with SLI	150	150
Students with ID	150	150
Students who are Deaf/HH	150	150
Non-Special Education Students	150	150
Grade 4 Total n = 1500	750	750
Grade 8		
Special Education Students		
Students with LD	150	150
Students with SLI	150	150
Students with ID	150	150
Students who are Deaf/HH	150	150
Non-Special Education Students	150	150
Grade 8 Total n = 1500	750	750

Implemented Field Test. Because of delays in obtaining IRB approval and a late start date in identifying participating schools, the field test actually started in April, 2010. After the data collection in April and May, only a few students in some disability categories had participated – this was particularly the case for students with intellectual disabilities and students who were deaf/hard of hearing.

Table 3. Actual Distribution of Students for the PARA Field Test¹

Student Group	Accessible Assessment	Proxy State Assessment
Grade 4		
Special Education Students		
Students with LD	131	113
Students with SLI	68	72
Students with ID	4	8
Students who are Deaf/HH	17	16
Non-Special Education Students	447	413
Grade 4 Total n = 1911	1289	622
Grade 8		
Special Education Students		
Students with LD	401	448
Students with SLI	23	30
Students with ID	29	32
Students who are Deaf/HH	27	29
Non-Special Education Students	588	570
Grade 8 Total n = 2177	1068	1109

¹These numbers reflect those students taking the assessment, not necessarily the number of students included in the analyses (which may have been less because of missing MRQ or state test data).

A revised data collection plan was established to supplement the field test sample. It focused on obtaining sufficient numbers of students with learning disabilities and with speech language impairments. State approval was obtained to enter the schools again in the fall of 2010 in grades 5 and 9 to supplement the field test data. Identification of participating schools started in the summer of 2010. The main reason for selecting grades 5 and 9 in the fall instead of grades 4 and 8 was to obtain students with similar levels of content knowledge/instruction in reading; this was more likely to be the case for spring and fall students in consecutive grades.

Results

As noted in the Introduction, research questions were grouped into two categories: (a) Primary Research Questions (PRQ), which focused on the effectiveness, reliability, and validity of the accessible and proxy state reading assessments, and (b) Secondary Research Questions (SRQ), which provided additional evidence about various aspects of the accessible reading assessments and the students who took them, including questions about different types of text, different levels of student motivation, and the use of accommodations.

Data for this study were obtained from two states at two times (spring and fall). Somewhat different population characteristics emerged in the spring and fall testings. To control for the spring-fall differences, we used a equipercentile test equating procedure to equate scores across testing times. This procedure, along with the computational approach and conversion tables, is described in Appendix M. We also conducted analyses for the Primary Research Questions for scores equated separately for the accessible and proxy assessments. These analyses, which were consistent with those presented here, are provided in Appendix N.

Primary Research Questions

Primary Research Question 1 (RQ1): Is the PARA accessible reading assessment more accessible than a proxy state reading assessment for students with learning disabilities (LD) and students with disabilities combined (LD, SLI, MR, and D/HH)?

We hypothesized that students with LD and special education students (SPED) combined (LD, SLI, MR, and D/HH) would perform better on the accessible assessment than the proxy state assessment. Group comparisons were between (a) LD accessible assessment vs. LD proxy state assessment, and (b) SPED accessible assessment vs. SPED proxy state assessment. To control for initial differences between students in the treatment (accessible) and control (proxy state assessment) groups, students' *state* reading test scores from 2010 were used as a covariate. The dependent variable in this analysis was the reading assessment (accessible and proxy) score.

Table 4 presents a summary of descriptive analyses of the performance of each group (Non-SPED, LD, SLI, ID, D/HH) on the accessible and proxy state assessments. The data in Table 4 suggest substantial differences between groups of students, but not a large difference between performance on the accessible and the proxy state assessment. For example, the overall mean assessment score (average of the accessible and proxy) for Non-SPED students was 19.46, compared to an overall mean of 11.95 for students with LD. As seen in Table 4, the performance difference between the accessible and proxy state assessments for these students was quite small (e.g., 12.27 vs. 11.67 for students with LD). For groups other than Non-SPED, SPED, and LD comparisons of the accessible and proxy state assessments may not be meaningful due to the small sample sizes.

Table 4. Mean Scores for Reading Score by Test Form, Disability Status, and Grade

		GRADE 4/5			GRADE 8/9		
		Mean	Std. Deviation	N	Mean	Std. Deviation	N
Non-SPED	Proxy	19.36	4.59	375	19.65	4.99	485
	Accessible	19.56	4.56	376	19.46	5.07	508
	Total	19.46	4.58	751	19.55	5.03	993
LD	Proxy	12.27	4.56	99	12.87	5.08	373
	Accessible	11.67	4.34	113	12.35	4.80	331
	Total	11.95	4.44	212	12.63	4.96	704
SLI	Proxy	18.91	5.09	69	14.38	5.29	22
	Accessible	18.93	4.68	59	15.24	6.39	17
	Total	18.92	4.89	128	14.75	5.73	39
ID	Proxy	9.21	2.20	7	10.82	5.11	23
	Accessible	9.25	2.36	4	10.71	5.38	17
	Total	9.22	2.14	11	10.77	5.16	40
D/HH	Proxy	9.05	5.14	13	10.82	6.87	16
	Accessible	12.75	6.20	12	11.47	6.23	19
	Total	10.83	5.86	25	11.17	6.44	35
Total	Proxy	17.70	5.59	563	16.40	6.14	919
	Accessible	17.70	5.61	564	16.40	6.15	892
	Total	17.70	5.60	1127	16.40	6.14	1811

* All scores in this table have been equated. The table includes only those students who had both a state test score and either an accessible or a proxy state assessment score.

The results of analyses are reported separately for grade 4/5 and grade 8/9. As data in Table 4 suggest, the pattern of data was very similar across the two grades. Similar to the data for grade 4/5, there was a substantial difference between the performance of the SPED and Non-SPED students in grade 8/9, but little difference in performance between the accessible and proxy state assessments. Analyses showed that there were no significant differences between the scores of students who took the accessible assessment and those of students who took the proxy state assessment.

In summary, the results of analyses for RQ1 did not support the hypothesis that students with disabilities perform significantly better on the accessible reading assessment compared to the performance of students with disabilities who took the proxy assessment.

Primary Research Question 2 (RQ2): Are scores from the PARA accessible reading assessment more reliable than scores from the proxy state reading assessments for students with LD and for students with disabilities combined (LD, SLI, MR, and D/HH)?

We hypothesized that the reliability of scores would be higher for the accessible reading assessment compared to the proxy state assessment for students with LD, and for students with disabilities combined (LD, SLI, MR, and D/HH). To test this hypothesis, reliability (internal consistency, coefficient alpha) of the PARA accessible and proxy state assessments were calculated using item level data for (a) students with LD, (b) SPED students, and (c) Non-SPED students. Prior to the computation of alpha coefficient, principal components analyses were conducted to examine dimensionality of items because coefficient alpha is extremely sensitive to multidimensionality (Cortina, 1993). Results of the principal components analyses indicated that the items in both assessments (accessible and proxy state) were unidimensional.

Table 5 presents the reliability coefficients for the accessible and proxy state assessments by groups (LD, SPED, and Non-SPED) for students in grades 4/5 and 8/9. As the data in Table 5 show, reliability for the proxy state assessment for students with LD in grade 4/5 was .75 and for the accessible assessment was .67. For the SPED group, the reliability coefficients for both tests were slightly higher for the proxy state assessment (.87) than the accessible (.83) assessment. For the Non-SPED group, reliability was .83 for the proxy state assessment and .79 for the accessible assessment in grade 4/5. Reliability was .81 for the proxy state assessment and .82 for the accessible assessment in grade 8/9 for Non-SPED students.

Table 5. Reliability (Internal Consistency) Coefficient for the Accessible and Proxy State Assessment for Non-SPED, SPED, and LD Students in Grade 4/5 and 8/9

Grade 4/5	Non-SPED		SPED		LD	
	Proxy	Accessible	Proxy	Accessible	Proxy	Accessible
Internal Consistency	.83	.79	.87	.83	.75	.67
N	375	376	188	188	99	113
Grade 8/9	Non-SPED		SPED		LD	
	Proxy	Accessible	Proxy	Accessible	Proxy	Accessible
Internal Consistency	.81	.82	.77	.76	.76	.74
N	485	508	434	384	373	331

Reliability coefficients for students in grade 8 were similar to those for students in grade 4/5. For the accessible assessment, the reliability coefficient for LD students was .74 compared to .76 for the proxy state assessment. For SPED group, the reliability coefficient

for the accessible assessment was .76 compared to .77 for the proxy state assessment. Thus, the reliability coefficients for the proxy assessment were slightly higher than the coefficients for the accessible assessment. The higher reliability for the SPED group could be due to the larger number of students in this group and higher true score variance in the SPED group (Allen & Yen, 1979; Thorndike, 2005). Similarly, the lower reliability coefficient for the accessible assessment might be partly due to the lower variability in the accessible assessment scores compared with the distribution of the proxy state assessment scores.

In summary, the results of analyses for RQ2 did not support the hypothesis that the accessible reading assessment had higher reliability than the proxy state reading assessment.

Primary Research Question 3 (RQ3): Does the PARA accessible reading assessment improve performance for non-special education students, thereby suggesting that it alters the focal construct?

We hypothesized that non-special education students would perform the same on the accessible reading assessment and the proxy state assessment. To test this hypothesis, the total test scores of Non-SPED students on the PARA accessible assessment were compared with the total test scores of Non-SPED students on the proxy state assessment using the student’s score on the most recent state reading assessment as a covariate (see Table 6). The Non-SPED group included those students assessed in the spring and fall after scores had been equated and adjusted for differences between the two administrations of the assessments.

Table 6. Descriptive Statistics for Grade 4/5 Non-SPED and LD Students for Accessible and Proxy State Reading Assessments

Assessment Form	Non-SPED Students			LD Students		
	<i>M</i>	<i>SD</i>	<i>N</i>	<i>M</i>	<i>SD</i>	<i>N</i>
Proxy	19.56	4.56	474	11.67	4.34	113
Accessible	19.36	4.59	489	12.27	4.56	99

We used the reading assessment scores for the Non-SPED and LD groups to investigate whether the pattern of results was different. A one-way analysis of covariance was performed to compare the means of Non-SPED grade 4/5 students who took the proxy state assessment and those who took the accessible assessment. In this analysis we included the student’s most recent state reading test score as a covariate. The results showed that there was a significant difference in the scores of students who took the proxy state assessment ($M = 19.56, SD = 4.56$) versus the accessible assessment ($M = 19.36, SD = 4.59$), $F(1, 748) = .34, p = .56$. These results suggested that Non-SPED

grade 4/5 students performed equally on the proxy state assessment and the accessible assessment.

The same analysis was performed for LD students in grade 4/5 to compare their performance on the accessible and proxy state assessments. The results showed no differences between the means on the proxy state ($M = 11.67, SD = 4.34$) and accessible ($M = 12.27, SD = 4.56$) assessments, $F(1, 209) = 0.87, p = .35$.

In summary, the results of analyses for RQ3 supported the hypothesis that accessibility features of the PARA assessment did not alter the construct of reading. When using the scores equated based just on the Non-SPED and LD sample, the results showed higher scores for Non-SPED students who took the proxy state assessment. For grade 8/9 students the results of analyses did not reveal any significant difference between Non-SPED students who took the accessible assessment and those who took the proxy state assessment indicating that the PARA accessibility features did not alter the focal construct of this study (i.e., reading).

Primary Research Question 4 (RQ4): Does the PARA accessible reading assessment demonstrate the same characteristics (accessibility, reliability, validity) for students in grades 4 and 8?

For this question, we hypothesized that in both grades 4/5 and 8/9, (a) students with LD and SPED students (LD, SLI, MR, and D/HH) would perform better on the accessible assessment than the proxy state assessment; (b) students with LD and SPED students would obtain scores with higher reliability on the PARA accessible reading assessment than on the proxy state assessment; and (c) Non-SPED students would perform the same on the PARA accessible assessment and the proxy state assessment.

Group performances were compared across grade levels and disability status using an ANCOVA design in which state reading test scores were used as a covariate. Results of analyses by grade (grades 4/5 and 8/9) showed consistent findings across the two grades. In terms of how effective the accessible assessment was in increasing the scores of SPED students, the results were consistent in the two grades, suggesting no significant difference between the proxy state and accessible assessments.

The results of analyses in terms of the validity of scores from accessible reading assessments were also consistent across the two grades and suggested that the accessibility features incorporated into the assessments did not alter the construct being assessed (i.e., reading). Therefore, the outcomes of the accessible assessments for Non-SPED students were comparable with the outcomes of the proxy state assessment in both grades 4/5 and 8/9. This was not the case when the analyses were performed using the equated scores across fall and spring based only on the Non-SPED and LD samples. In this case, there were significant differences in the means of the proxy state and accessible assessments for Non-SPED grade 4/5 students. The analyses for the Non-SPED grade 8/9 students also revealed a significant difference in the mean of the accessible ($M =$

16.84, $SD = 5.79$) and proxy state ($M = 17.57$, $SD = 6.60$) assessments, $F(1, 990) = 4.83$, $p < .05$. This suggested that those Non-SPED students taking the proxy state assessment performed better than those who took the accessible assessment. Analyses for LD students showed no differences in their performance for the accessible and proxy state assessments.

Results of analyses on the reliability of the accessible and proxy state assessments were also consistent across the two grades and suggested small discrepancies in the reliability of the assessments. The proxy state assessment for the grade 4 students showed slightly higher reliability than the accessible assessment, potentially due to score heterogeneities.

In summary, the results of analyses by grade showed consistent findings across the two grades. In both grades 4/5 and 8/9 there were no significant differences between the accessible and proxy state assessments. The same consistency across grades was found for the validity of scores, with both grades showing that the accessibility features incorporated into the assessment did not alter the construct being assessed. Finally, analyses of reliability of the accessible and proxy state assessments were consistent across the two grades.

Secondary Research Questions

Results of analyses for the Secondary Research Questions are presented here. To address the comparability of scores across the type of assessment, assessment locations, and assessment times, data from different times/locations were equated using equipercentile equating with a less restricted assumption. Raw scores were converted to equipercentile equated scores and analyses were conducted on the converted scores under the hierarchical linear model with students' most recent scores on the state's reading test as a covariate.

Secondary Research Question 1 (SRQ1): Does students' performance on the reading assessment correlate with the overall measure of motivation, based on the results of the MRQ survey?

Hypothesis. We hypothesized that in each grade, there would be a significant correlation between students' overall motivation score and their performance on both the accessible and proxy state assessments for: (a) students with LD, (b) SPED students, (c) non-SPED students, and (d) all students overall.

Analyses. Correlations were calculated between overall reading assessment raw scores and overall MRQ scores for: (a) students with LD, (b) SPED students, (c) non-SPED students, and (d) all students. These analyses were restricted to those students with complete MRQ data. Prior to conducting these correlational analyses, we examined distributional characteristics (skewness and kurtosis for any indication of restriction of range and other related characteristics) to check possible violations to the assumptions of the Product Moment correlation (such as homoscedasticity, etc.). Results did not show any major violations of the assumptions.

The MRQ questionnaire had 11 subscales (see the section under “Instrument” for detailed discussion of these categories). The 11 subscales had high reliability (internal consistency coefficient). The results of a confirmatory factor analysis model also confirmed that the subscales were unidimensional and internally consistent. Results of a hierarchical confirmatory factor model indicated that a higher-order factor could explain a substantial amount of the variance of the 11 subscales. Therefore, a single “Motivation” latent variable was assumed and this higher-order factor was correlated with the reading assessment scores.

Results. Results of the analyses for this question indicated that there was a significant correlation between students’ overall motivation score (based on the MRQ) and their performance on the reading assessment (across proxy and accessible assessments) for all student groups, except students with LD (see Table 7). Although these correlations were significant, their magnitude was small. The proportion of variance in reading scores explained by the MRQ ranged from 1.2 percent for SPED students to 5.7 percent for all grade 8 students.

Table 7. Correlations between Reading Scores and MRQ Overall Motivation Scores

Sample	<i>r</i> (based on higher-order CFA)	<i>Sig</i>
Students with LD	0.05	.19
SPED Students	0.11	< .01
Non-SPED sStudents	0.20	< .01
All Students	0.24	< .01
Grade 4 Students	0.19	< .01
Grade 8 Students	0.24	< .01

Secondary Research Question 2 (SRQ2). Do students perform better on the passages they rate as more interesting?

Hypothesis. We hypothesized that students would perform better on reading assessment passages that they had rated as more interesting on the Situated Motivation Questions.

Analyses Analyses were performed at the item level. Reading test items were divided into several testlets, each with the test items related to one of the passages. These testlets were then grouped based on the level of interest in those passages. Student performances were compared across the testlets using an HLM model.

Results. Results are presented separately for: (a) grade 4/5 accessible assessment students, (b) grade 4/5 proxy state assessment students, (c) grade 8/9 accessible assessment students, and (d) grade 8/9 proxy state assessment students.

Sample: Grade 4/5 Accessible Assessment. As the data in Table 8 suggest, estimated mean performance was higher when the perceived interest level of the passage was higher. For each unit increase in interest there was .03 increase in the reading assessment score. Students who had average performance on the state reading test and who reported the highest level of interest in the passages, scored on average 4.45 out of a possible 7 on the reading assessment testlets. At the lowest level of interest, the mean accessible assessment testlet score was 4.37.

Table 8. Grade 4 Accessible Assessment—Student Performance Across Interest Levels (0 = highest level of interest, 3 = lowest level of interest)

Interest Level	Estimated Mean Test Score
0 (Highest)	4.45
1	4.42
2	4.40
3 (Lowest)	4.37

For grade 4/5 students who took the accessible assessment, lower reported interest on the passages did not affect significantly their score on the reading assessment testlets. If they had scored higher than average on the state test, their reading assessment testlet score increased by 0.11 (see Table 9).

Table 9. Grade 4/5 Accessible—State Reading, Interest, and Interactions

	Value	Std Error	df	t-value	p-value
(Intercept)	4.45	0.05	1680	85.77	0.00
State Reading	0.11	0.01	562	14.26	0.00
Interest	-0.03	0.04	1680	-0.75	0.46
State Reading X Interest	0.00	0.00	1680	0.03	0.97

Sample: Grade 8/9 Accessible Assessment. As the data in Table 10 show, similar to the results for students in grade 4, the mean score for students in grade 8 was systematically higher with each unit increase in interest level. For example, at the lowest level, the mean accessible assessment reading testlet score for grade 8 students was 3.84 out of 7. This increased by 0.15 to 3.99 at the level 2 interest and further increased to 4.14 and 4.28 at levels 1 and 0 (the highest level of interest) respectively.

Table 10. Grade 8/9 Accessible—Student Performance across Interest Levels (0 = highest level of interest, 3 = lowest level of interest)

Interest Level	Mean Assessment Score
0 (Highest)	4.28
1	4.14
2	3.99
3 (Lowest)	3.84

The increase in the assessment score was statistically significant ($p < .000$). Thus, students reporting the least amount of interest in the passage scored 3.84 on average when performance was conditional on the state reading score. Thus, those who had higher scores on the state reading test had a slightly greater decrease in their accessible reading assessment testlet score ($b = -0.01$, $p < .01$) as their interest in the passages dropped. And those who had lower scores on the state reading test had less of a decrease in their accessible reading assessment testlet score as their interest in the passages dropped (see Table 11).

Table 11. Grade 8 Accessible—State Reading, Interest, and Interactions

	Value	Std Error	df	t-value	p-value
(Intercept)	4.28	0.05	2665	78.73	0
State Reading	0.07	0.00	890	14.90	0
Interest	-0.15	0.02	2665	-6.20	0
State Reading X Interest	-0.01	0.00	2665	-5.44	0

Sample: Grade 4/5 Proxy State Assessment. Grade 4/5 students who reported the lowest level of interest in the passages, scored on average 4.41 out of a possible 7 on the proxy state reading assessment testlets (see Table 12). If they had scored higher than the average on the state reading test, their reading assessment testlet score was higher by 0.12.

Table 12. Grade 4 Proxy State Assessment—Student Performance Across Interest Levels (0 = highest level of interest, 3 = lowest level of interest)

Interest Level	Mean Assessment Score
0 (Highest)	4.86
1	4.71
2	4.56
3 (Lowest)	4.41

For this group of students, reporting lower interest in the passages affected significantly their score on the proxy state reading assessment testlet (see Table 13). For each unit of lower interest, their scores on the proxy state reading assessment dropped 0.15 units on average. Those reporting the highest level of interest in the passage scored 4.86 on average.

Table 13. Grade 4 Proxy—State Reading, Interest, and Interactions

	Value	Std Error	df	t-value	p-value
(Intercept)	4.84	0.06	1680	85.98	0.00
State Reading	0.13	0.01	561	18.01	0.00
Interest	-0.15	0.03	1680	-4.98	0.00
State Reading X Interest	0.01	0.00	1680	1.41	0.16

Sample: Grade 8/9 Proxy State Assessment. Students who had an average score on the state reading test and who reported the highest level of interest in the passages scored on average 3.79 out of a possible 7 on the proxy state reading assessment (see Table 14). If they had scored higher than average on the state test, their reading assessment score increased by 0.12 units.

Table 14. Grade 8/9 Proxy State Assessment—Student Performance Across Interest Levels (0 = highest level of interest and 3 = lowest level of interest)

Interest Level	Mean Assessment Score
0 (Highest)	3.79
1	3.72
2	3.65
3 (Lowest)	3.58

For this group of students, reporting lower interest in the passages related significantly to their score on the proxy state reading assessment testlets (see Table 15). For each unit of decrease in interest, scores on the state proxy reading assessment testlets dropped 0.07 on average. Those reporting the least amount of interest in the passage scored 3.51 on average.

Table 15. Grade 8/9 Proxy—State Reading, Interest, and Interactions

	Value	Std Error	df	t-value	p-value
(Intercept)	3.79	0.05	2735	74.10	0.00
State Reading	0.12	0.01	916	21.53	0.00
Interest	-0.07	0.03	2735	-2.51	0.01
State Reading X Interest	-0.00	0.00	2735	-0.16	0.87

Secondary Research Question 3 (SRQ3) Do special education and non-special education students rate informational-expository passages or narrative-literary passages as more interesting or are both rated as equally interesting?

Hypothesis: We hypothesized that students would rate informational-expository passages and narrative-literacy passages as equally interesting. Also, it was hypothesized that students who took the accessible reading assessment would rate all four passages as more interesting than would students who took the proxy state assessment.

Analyses. An HLM model was used. Because there were four forms of the assessment, all students did not take the same test items. Comparisons were thus made on test items under passages that were similar in content (referred to as *testlets*).

Results. Results are presented separately for: (a) grade 4/5 accessible assessment students, (b) grade 4/5 proxy state assessment students, (c) grade 8/9 accessible assessment students, and (d) grade 8/9 proxy state assessment students.

Sample: Grade 4/5 Accessible Assessment. The level of interest of SPED and Non-SPED students by the type of passage was similar (see Table 16). The average level of interest reported for narrative passages by non-SPED students was 1.77; for informational passages the average level of interest was 1.69. For SPED students, the average level of interest for narrative passages was 1.84; for informational passages, the average level of interest was 1.76.

Table 16. Grade 4/5 Accessible Assessment—Mean Interest Level (by Passage Type/ Disability Status (0 = Highest Level of Interest, 3 = Lowest Level of Interest))

	SPED	Non-SPED
Informational	1.76	1.69
Narrative	1.84	1.77

The minimal differences in interest levels by type of passages were not statistically significant (see Table 17).

Table 17. Grade 4/5 Accessible Assessment—Interest Level by Disability Group and Passage Type

	Value	Std Error	df	t-value	p-value
(Intercept)	1.77	0.04	1680	46.98	0.00
Informational	-0.08	0.04	1680	-1.80	0.07
SPED	0.07	0.07	562	1.05	0.28
Informational X SPED	-0.01	0.08	1680	-0.19	0.85

Sample: Grade 8/9 Accessible Assessment. Table 18 shows the average interest level reported by grade 8/9 SPED and Non-SPED students by type of passage. On average, the level of interest reported by Non-SPED students for narrative passages was 2.35 (0 = most interest, 3 = least interest). When informational passages were rated, the interest level decreased by 0.25 units (lower rating corresponds with higher interest). In other words, students’ interest increased when informational passages were presented.

Table 18. Grade 8/9 Accessible Assessment—Mean Interest Level by Disability Group and Passage Type (0 = Highest Interest Level, 3 = Lowest Interest Level)

	SPED	Non-SPED
Informational	2.18	2.11
Narrative	2.27	2.35

As indicated by the data in Table 18, the type of passage did have an influence on the interest level reported by grade 8/9 students who took the accessible assessment (see Table 19). This was true regardless of whether students were in the SPED or the Non-SPED group.

Table 19. Grade 8/9 Accessible Assessment—Interest Level by Disability Group and Passage Type

	Value	Std Error	df	t-value	p-value
(Intercept)	2.35	0.04	2665	66.10	0.00
Informational	-0.25	0.04	2665	-5.56	0.00
SPED	0.09	0.05	890	1.67	0.10
Informational X SPED	-0.00	0.07	2665	-0.05	0.96

*Fixed effects: A8Interest = Informational + SPED + Informational * SPED*

Sample: Grade 4/5 Proxy State Assessment. The average level of interest reported for narrative passages by Non-SPED students was 2.01 (see Table 20). The average interest level reported for narrative passages for SPED students was 2.12. These were similar to each other, and similar to the ratings for informational passages (1.97 and 2.07).

Table 20. Grade 4/5 Proxy State Assessment—Mean Interest Level by Disability Group and Passage Type (0 = Highest Level of Interest, 3 = Lowest Level of Interest)

	SPED	Non-SPED
Informational	2.07	1.97
Narrative	2.12	2.01

Analyses indicated that there were no differences in the interest ratings for SPED or Non-SPED students, regardless of the type of passage (see Table 21).

Table 21. Grade 4/5 Proxy State Assessment—Interest Level by Disability Group and Passage Type

	Value	Std Error	df	t-value	p-value
(Intercept)	2.01	0.04	1680	47.74	0.00
Informational	-0.04	0.05	1680	-0.90	0.37
SPED	0.11	0.07	561	1.50	0.13
Informational X SPED	-0.14	0.08	1680	-1.67	0.10

Fixed effects: P4Interest = Informational + SPED + Informational * SPED

Sample: Grade 8/9 Proxy State Assessment. On average, the level of interest reported by Non-SPED students for narrative passages was 2.53 (0 = most interest, 3 = least interest). When informational passages were rated, the interest rating was lower by 0.25 units (see Table 22).

Table 22. Grade 8/9 Proxy State Assessment—Mean Interest Level (by Disability Group and Passage Type (0 = Highest level of Interest, 3 = Lowest Level of Interest))

	SPED	Non-SPED
Informational	2.17	2.38
Narrative	2.38	2.53

Level of interest reported by students varied with the type of passage and whether the student group was SPED or Non-SPED (see Table 23). On average, the level of interest reported by Non-SPED students for narrative passages was 2.53 while for informational passages it was 0.25 units lower. Because a lower rating indicated higher interest, this

means that students' interest increased when informational passages were presented. For SPED students, interest was only 0.08 units higher for informational passages compared to narrative passages. Regardless of the type of passage, SPED students reported more interest than Non-SPED students (interest rating was 0.15 units lower for SPED students, indicating higher interest).

Table 23. Grade 8/9 Proxy State Assessment—Interest Level by Disability Group and Passage Type

	Value	Std Error	df	t-value	p-value
(Intercept)	2.53	0.04	2735	71.17	0.00
Informational	-0.25	0.04	2735	-5.91	0.00
SPED	-0.15	0.05	916	-2.84	0.00
Informational X SPED	0.18	0.06	2735	2.93	0.00

Fixed effects: $P8Interest = \text{Informational} + \text{SPED} + \text{Informational} * \text{SPED}$

Secondary Research Question 4 (SRQ4). Does the text type in reading assessment passages (narrative-literary or informational-expository) impact student performance in general and by groups of students?

Hypothesis. We hypothesized that students would perform equally well on narrative-literary and informational-expository passages.

Analyses. An HLM model was used to compare student performance for different types of passages. Student state assessment scores were used as the covariate.

Results. Results are presented separately for: (a) grade 4/5 accessible assessment students, (b) grade 4/5 proxy state assessment students, (c) grade 8/9 accessible assessment students, and (d) grade 8/9 proxy state assessment students.

Sample: Grade 4/5 Accessible Assessment. Table 24 shows the performance of SPED and Non-SPED students on the two passage types (informational-expository and narrative-literary).

Table 24. Grade 4/5 Accessible Assessment—Mean Performance by Disability Group and Passage Type

	SPED	Non-SPED
Informational-expository	3.84	4.83
Narrative-literary	3.77	4.53

Passage type did impact student performance in general ($b = .30, p < .01$), and having a disability did not change this pattern ($b = -0.003, p = ns$) (see Table 25). Students who had an average score on the state reading test and were in the Non-SPED group scored on average 4.53 on the narrative passages of the reading assessment out of a possible 7 points. For informational passages, this score on average was 0.30 units higher. SPED students scored on average 0.76 units lower overall. If they had scored higher than average on the state tests, their accessible assessment reading score was 0.09 units higher.

Table 25. Grade 4/5 Accessible Assessment—Effects of Reading Passage Type on Student Performance

	Value	Std Error	df	t-value	p-value
(Intercept)	4.53	0.06	1689	71.30	0.00
State Reading	0.09	0.01	561	13.87	0.00
SPED	-0.76	0.12	561	-6.52	0.00
Informational	0.30	0.07	1689	4.39	0.00
State Reading X Informational	0.01	0.01	1689	0.88	0.38
SPED X Informational	0.00	0.13	1689	-0.02	0.98

Fixed effects: $A4 = \text{State Reading} + \text{SPED} + \text{Informational} + \text{State Reading} * \text{Informational} + \text{Informational} * \text{SPED}$

Sample: Grade 8/9 Accessible Assessment. Table 26 shows the performance of SPED and Non-SPED students who took the accessible assessment on the two passage types (informational-expository and narrative-literary).

Table 26. Grade 8 Accessible Assessment—Mean Performance by Disability Group and Passage Type

	SPED	Non-SPED
Informational-expository	3.27	4.77
Narrative-literary	3.27	4.67

Passage type did not impact student performance in general ($b = 0.10, p = ns$) nor for SPED students ($b = -0.10, p = ns$) (see Table 27). Students who were non-SPED and had an average score on the state reading test when taking the narrative passages, scored on average 4.67 on the reading assessment out of a possible 7. SPED students had scores 1.40 units lower, on average. SPED students who scored higher than average on the state test had a score on the accessible reading assessment that was 0.03 units higher.

Table 27. Grade 8/9 Accessible—Effects of Reading Passage Type on Student Performance

	Value	Std Error	df	t-value	p-value
(Intercept)	4.67	0.06	2673	74.58	0.00
State Reading	0.03	0.00	889	8.41	0.00
SPED	-1.40	0.10	889	-13.94	0.00
Informational	0.10	0.06	2673	1.74	0.08
State Reading X Informational	-0.003	0.00	2673	-0.91	0.37
SPED X Informational	-0.10	0.09	2673	-1.10	0.27

Fixed effects: A8 = State Reading + SPED + Informational + State Reading * Informational + -Informational * SPED

Sample: Grade 4/5 Proxy State Assessment. Table 28 shows the performance of SPED and Non-SPED students who took the proxy state assessment for the two passage types (informational-expository and narrative-literary).

Table 28. Grade 4/5 Proxy State Assessment—Mean Performance by Disability Group and Passage Type

	SPED	Non-SPED
Informational-expository	4.28	4.94
Narrative-literary	4.31	4.88

Passage type did not affect student performance in general ($b = 0.07, p = ns$) nor for SPED students ($b = -0.09, p = ns$) (see Table 29). Students in the Non-SPED group who had an average score on the state reading test, when taking the narrative passages, scored on average 4.88 on the reading assessment out of a possible 7. SPED students had scores that were 0.57 units lower, on average. SPED students with higher than average scores on the state reading test had higher scores on the proxy state reading assessment, but there was also an interaction of passage type and state reading score. For those students who scored higher than average on the state reading test, their scores on the proxy state reading assessments were 0.02 units lower.

Table 29. Grade 4/5 Proxy State Assessment—Effects of Reading Passage Type on Student Performance

	Value	Std Error	df	t-value	p-value
(Intercept)	4.88	0.06	1686	78.84	0.00
State Reading	0.13	0.01	560	19.70	0.00
SPED	-0.57	0.11	560	-5.08	0.00
Informational	0.06	0.06	1686	1.07	0.28
State Reading X Informational	-0.02	0.01	1686	-3.46	0.00
SPED X Informational	-0.09	0.11	1686	-0.82	0.41

Fixed effects: P4 = State Reading + SPED + Informational + State Reading * Informational + Informational * SPED

Sample: Grade 8/9 Proxy State Assessment. Table 30 shows the performance of SPED and Non-SPED students who took the proxy state assessment on the two passage types (informational-expository and narrative-literary).

Table 30. Grade 8/9 Proxy State Assessment—Mean Performance by Disability Group and Passage Type

	SPED	Non-SPED
Informational-expository	3.58	4.42
Narrative-literary	3.16	3.63

Passage type did impact student performance in general ($b = 0.79, p < .01$) as well as for SPED students ($b = -0.50, p < .01$) (see Table 31). Students in the Non-SPED group who had an average score on the state reading test, when taking the narrative passages, scored on average 3.63 on the reading assessment out of a possible 7. SPED students had scores that were significantly lower, by 0.47 units on average. With informational passages, students' reading assessment scores were 0.79 units higher than for narrative passages. If students had scored higher than average on the state test, then their scores were 0.03 units higher. Higher than average scores on the state reading test were associated with higher reading assessment scores. For SPED students on informational passages, the reading assessment score was 0.50 units lower.

Table 31. Grade 8/9 Proxy State Assessment—Effects of Reading Passage Type on Student Performance

	Value	Std Error	df	t-value	p-value
(Intercept)	3.63	0.06	2754	58.53	0.00
State Reading	0.08	0.01	916	15.79	0.00
SPED	-0.47	0.10	916	-4.73	0.00
Informational	0.79	0.07	2754	11.17	0.00
State Reading X Informational	0.03	0.01	2754	3.91	0.00
SPED X Informational	-0.50	0.11	2754	-4.39	0.00

Fixed effects: P8 = State Reading + SPED + Informational + State Reading * Informational + Informational * SPED

Secondary Research Question 5 (SRQ5). Do students score better on passages they rate as “kind of challenging” or “pretty challenging” in terms of how challenging (difficult) the passage was to read?

Hypothesis. We hypothesized that students would score better on passages rated as “C” or “D” on the challenging scale than they would on passages rated as “A” or “B.”

Analyses. An HLM model was used to compare students’ performance on passages (i.e., testlets) that were rated differently on the “challenging” scale. Students’ state test scores were used as a covariate.

Results. Results are presented separately for: (a) grade 4/5 accessible assessment students, (b) grade 4/5 proxy state assessment students, (c) grade 8/9 accessible assessment students, and (d) grade 8/9 proxy state assessment students.

Sample: Grade 4/5 Accessible Assessment. The difficulty of the passage was rated by students for each passage and the average assessment score calculated for students with the same ratings (see Table 32).

Table 32. Grade 4/5 Accessible Assessment—Mean Performance by Disability Group and Difficulty Rating

	SPED	Non-SPED
Very Challenging	3.27	3.98
Pretty Challenging	3.36	4.27
Kind of Challenging	3.45	4.56
Not at all Challenging	3.53	4.84

The rating the students gave to passages affected student performance negatively ($b = 0.29, p < .01$), and having a disability did not change this pattern ($b = 0.009, p = ns$) (see Table 33). When Non-SPED students who had an average score on the state reading test rated passages as the most challenging, they scored on average 3.98 out of a possible 7. This score was 0.11 units higher if they had scored higher than average on the state reading test. It decreased by 0.71 for SPED students. For each unit the passage was rated less challenging, the reading assessment score was 0.29 units higher.

Table 33. Grade 4/5 Accessible Assessment—Student Performance on Reading Passages by Difficulty Level

	Value	Std Error	df	t-value	p-value
(Intercept)	3.98	0.15	1663	26.22	0.00
State Reading	0.11	0.14	561	7.42	0.00
SPED	-0.71	0.24	561	-2.94	0.00
A4Hard.c	0.29	0.06	1663	5.13	0.00
State Reading X A4Hard.c	-0.01	0.01	1663	-1.24	0.22
SPED X A4Hard.c	0.01	0.09	1663	0.11	0.92

Fixed effects: A4 = State Reading + SPED + A4Hard.c + A4Hard.c * State Reading + A4Hard.c * SPED

Sample: Grade 8/9 Accessible Assessment. The difficulty of the passage was rated by students for each passage and the average assessment score calculated for students with the same ratings (see Table 34).

Table 34. Grade 8/9 Accessible Assessment—Mean Performance Level by Disability Groups and Difficulty Rating

	SPED	Non-SPED
Very Challenging	3.13	3.60
Pretty Challenging	3.37	4.05
Kind of Challenging	3.61	4.50
Not at all Challenging	3.85	4.95

The rating the students gave to the passages was negatively related to student performance ($b = 0.45, p < .01$), and having a disability made the impact even stronger ($b = -0.34, p < .01$) (see Table 35). When Non-SPED students who had an average score on the state reading test rated the passages as the most challenging, they scored on average, 3.60 out of a possible 7. This score was 0.09 units higher if they had scored higher than average on the state reading test. It decreased by 0.47 for SPED. For each unit the passage was rated less challenging, the reading assessment score was 0.45 units higher. Also, for those who scored higher than average on the state reading test, for every unit the challenge of the passage decreased, students’ reading assessment scores increased slightly less than those who rated the passage as very challenging ($b = -0.02, p < .01$). Finally, SPED students who rated a passage as less challenging did not show as much an effect as for Non-SPED students ($b = -0.34, p < .01$).

Table 35. Grade 8/9 Accessible Assessment—Student Performance on Reading Passages by Difficulty Level

	Value	Std Error	df	t-value	p-value
(Intercept)	3.60	0.13	2658	27.40	0.00
State Reading	0.09	0.01	889	7.33	0.00
SPED	-0.47	0.21	889	-2.26	0.02
A8Hard.c	0.44	0.05	2658	9.46	0.00
State Reading X A8Hard.c	-0.02	0.00	2658	-5.11	0.00
SPED X A8Hard.c	-0.34	0.08	2658	-4.58	0.00

Fixed effects: A8 = State Reading + SPED + A8Hard.c + A8Hard.c * State Reading + A8Hard.c * SPED

Sample: Grade 4/5 Proxy State Assessment. The difficulty of the passage was rated by students for each passage and the average assessment score calculated for students with the same ratings (see Table 36).

Table 36. Grade 4/5 Proxy State Assessment—Mean Performance by Disability Group and Difficulty Rating

	SPED	Non-SPED
Very Challenging	3.48	4.43
Pretty Challenging	3.49	4.62
Kind of Challenging	3.50	4.81
Not at all Challenging	3.51	5.00

The rating the student gave to passages was negatively related to student performance ($b = 0.19, p < .01$), and having a disability made this relation even stronger ($b = -0.95, p < .01$) (see Table 37). When Non-SPED students who had an average score on the state reading test rated passages as the most challenging, they scored on average 4.43 on the reading assessment out of a possible 7. This score was 0.08 units higher if they had scored higher than average on the state reading test and it decreased by 0.95 units for SPED students. For each unit the passage was rated less challenging, the reading assessment score was 0.19 units higher. Also, for those students who scored higher than average on the state reading tests, for every unit of decrease in challenge of the passage, their reading assessment score increased slightly more than those who rated the passage as very challenging ($b = 0.01, p < .01$). Finally, SPED students performed better on the passages they rated as less challenging than did the Non-SPED students ($b = 0.17, p < .01$).

Table 37. Grade 4/5 Proxy State Assessment—Student Performance on Reading Passages by Difficulty Level

	Value	Std Error	df	t-value	p-value
(Intercept)	4.43	0.14	1672	31.44	0.00
State Reading	0.08	0.01	560	6.14	0.00
SPED	-0.95	0.21	560	-4.50	0.00
P4Hard.c	0.19	0.05	1672	3.64	0.00
State Reading X P4Hard.c	0.01	0.01	1672	2.73	0.01
SPED X P4Hard.c	0.17	0.08	1672	2.09	0.04

Fixed effects: P4 = State Reading + SPED + P4Hard.c + P4Hard.c * State Reading + P4Hard.c * SPED

Sample: Grade 8/9 Proxy State Assessment. The difficulty of the passage was rated by students for each passage and the average assessment score calculated for students with the same ratings (see Table 38).

Table 38: Grade 8/9 Proxy State Assessment—Mean Performance by Disability Group and Difficulty Rating

	SPED	Non-SPED
Very Challenging	2.81	3.54
Pretty Challenging	2.86	3.74
Kind of Challenging	2.92	3.94
Not at all Challenging	2.97	4.14

The rating the student gave to passages was negatively related to student performance ($b = 0.20, p < .01$), and having a disability made this impact even stronger ($b = -0.73, p < .01$) (see Table 39). When Non-SPED students who had an average score on the state reading test rated passages as the most challenging, they scored on average 3.54 on the reading assessment out of a possible 7. This score was 0.08 units higher if they had scored higher than average on the state reading test and it decreased by 0.73 for SPED students. For each unit the passage was rated less challenging, the reading assessment score was 0.20 units higher.

Table 39: Grade 8/9 Proxy State Assessment—Student Performance on Reading Passages by Difficulty Level

	Value	Std Error	df	t-value	p-value
(Intercept)	3.53	0.14	2724	25.75	0.00
State Reading	0.08	0.01	915	6.88	0.00
SPED	-0.73	0.20	915	-3.72	0.00
P8Hard.c	0.20	0.05	2724	3.76	0.00
State Reading X P8Hard.c	0.01	0.00	2724	1.88	0.06
SPED X P8Hard.c	0.03	0.08	2724	0.41	0.68

Fixed effects: P8 = State Reading + SPED + P8Hard.c + P8Hard.c * State Reading + P8Hard.c * SPED

Secondary Research Question 6 (SRQ6). Do students perform better on the narrative-literary passages when they rate that they liked the passage because they knew something about it or when they rate that they liked the passage because it was a story? How often did they mark both responses as reasons for liking the passage, and does doing so have an effect on performance?

Hypothesis. We hypothesized that most students would indicate that they liked a passage because they knew something about it and because it was a story. We also

hypothesized that passages rated this way would show the highest performance, followed by those rated as liked because it was a story, followed by those rated as liked because they knew something about it.

Analysis. An HLM model was used for the analysis for this question. The repeated measures included performance on items under different types of passages.

Results. Results are presented separately for: (a) grade 4/5 accessible assessment students, (b) grade 4/5 proxy state assessment students, (c) grade 8/9 accessible assessment students, and (d) grade 8/9 proxy state assessment students.

Sample: Grade 4/5 Accessible Assessment. Table 40 shows the average performance on the grade 4/5 accessible assessment as a function of the way in which students rated the reasons for liking the narrative-literary passage that they read.

Table 40. Grade 4/5 Accessible Assessment—Mean Performance on Assessment by Passage Ratings (Narrative-Literary)

<i>Reason for liking the passage</i>	Average performance on assessment	
	SPED	Non-SPED
Not for the reasons provided	3.18	4.53
It was a story	3.84	5.07
Reader knew something about it	3.34	4.31
It was a story and the reader knew something about it	2.73	4.36

For the narrative-literary passages, when Non-SPED students reported liking the passage (but not for the reasons provided), they scored on average 4.53 on the reading assessment out of a possible 7 (see Table 41). Students performed higher (when compared to Non-SPED who liked the passage but not for the reasons provided) if they reported liking the passage because it was a story ($b = 0.54, p < .05$). Liking the passage because they knew something about it was not related to the student’s performance on the reading assessment. Similarly, endorsing both reasons for liking the passage was not related to the reading assessment score. A main effect for SPED students was found. That is, SPED students scored significantly lower on the reading assessment than Non-SPED students ($b = -1.35, p < .01$). There was no interaction of SPED status with reasons for liking the passage.

Table 41: Grade 4/5 Accessible Assessment—Full Regression Table for Student Performance by Passage Ratings (Narrative-Literary)

	Value	Std Error	df	t-value	p-value
(Intercept)	4.53	0.22	464	20.18	0.00
A4LikeA	0.54	0.25	336	2.17	0.03
A4LikeB	-0.22	0.31	336	-0.71	0.48
A4Both	-0.17	0.26	336	-0.66	0.51
SPED	-1.35	0.40	464	-3.41	0.00
A4LikeA:SPED	0.12	0.45	336	0.26	0.79
A4LikeB:SPED	0.38	0.51	336	0.73	0.46
A4Both:SPED	-0.28	0.45	336	-0.63	0.53

Fixed effects: A4 ~ A4LikeA + A4LikeB + A4Both + SPED + SPED * A4LikeA + SPED * A4LikeB + SPED * A4Both

Sample: Grade 8/9 Accessible Assessment. Table 42 shows the average performance on the grade 8 accessible assessment as a function of the way in which students rated the reasons for liking the narrative-literary passage they read.

Table 42: Grade 8/9 Accessible Assessment—Mean Performance on Assessment by Passage Ratings (Narrative-Literary)

Reason for liking the passage	Average performance on assessment	
	SPED	Non-SPED
Not for the reasons provided	2.96	4.32
It was a story	3.72	5.13
Reader knew something about it	2.69	4.98
It was a story and the reader knew something about it	2.89	5.09

For the narrative-literary passages, when Non-SPED students reported liking the passage (but not for the reasons provided), they scored on average 4.32 on the reading assessment (see Table 43). Non-SPED students performed higher if they reported liking the passage because it was a story only ($b = 0.81, p < .01$), because they knew something about it only ($b = 0.66, p < .05$), or because of both reasons ($b = 0.77, p < .05$). The same was not true for SPED students. That is, when these students reported liking the passage but not for the reasons provided, their score was significantly lower than it was for Non-SPED students who also liked the passage but not for the reasons provided ($b = -1.36, p < .01$). Similarly, if SPED students reported liking the passage because they knew something about it, their score was even lower ($b = -0.93, p < .05$), as well as when

they reported liking the passage for both reasons ($b = -0.84, p < .05$). For SPED students, liking the passage because it was a story did not have a significant effect on their reading assessment scores.

Table 43: Grade 8/9 Accessible Assessment—Full Regression Table for Student Performance by Passage Ratings (Narrative-Literary)

	Value	Std Error	df	t-value	p-value
(Intercept)	4.32	0.25	596	17.34	0.00
A8LikeA	0.81	0.27	272	3.04	0.00
A8LikeB	0.66	0.31	272	2.15	0.03
A8Both	0.77	0.28	272	2.79	0.01
SWD	-1.36	0.33	596	-4.15	0.00
A8LikeA:SWD	-0.05	0.37	272	-0.13	0.90
A8LikeB:SWD	-0.93	0.43	272	-2.16	0.03
A8Both:SWD	-0.84	0.38	272	-2.23	0.03

Fixed effects: $A8 \sim A8LikeA + A8LikeB + A8Both + SPED + SPED * A8LikeA + SPED * A8LikeB + SPED * A8Both$

Sample: Grade 4/5 Proxy State Assessment. Table 44 shows the average performance on the grade 4/5 proxy state assessment as a function of the way in which students rated the reasons for liking the narrative-literary passage they read.

Table 44. Grade 4/5 Proxy State Assessment—Mean Performance on Assessment by Passage Ratings (Narrative-Literary)

<i>Reason for liking the passage</i>	Average performance on assessment	
	SPED	Non-SPED
Not for the reasons provided	4.25	5.06
It was a story	4.10	5.23
Reader knew something about it	3.56	5.12
It was a story and the reader knew something about it	3.48	5.02

For grade 4/5 students taking the proxy state assessment, performance on the reading assessment did not depend on their reasons for liking the narrative-literary passages. There was a significant main effect of SPED students (see Table 45), with SPED performance lower overall than Non-SPED performance. Non-SPED students who liked the passages (but not because of the reasons provided), scored on average 5.06 out of a

possible 7 on the reading assessment, but SPED students scored significantly lower with 4.25 out of a possible 7 ($b = -0.81, p < .05$).

Table 45. Grade 4/5 Proxy State Assessment—Full Regression Table for Student Performance by Passage Ratings (Narrative-Literary)

	Value	Std Error	df	t-value	p-value
(Intercept)	5.06	0.24	422	20.83	0.00
P4LikeA	0.17	0.26	271	0.65	0.52
P4LikeB	0.06	0.32	271	0.20	0.84
P4Both	-0.04	0.28	271	-0.13	0.90
SPED	-0.81	0.36	422	-2.26	0.02
P4LikeA:SPED	-0.32	0.42	271	-0.75	0.45
P4LikeB:SPED	-0.75	0.50	271	-1.50	0.14
P4Both:SPED	-0.73	0.43	271	-1.69	0.09

Fixed effects: $P4 \sim P4LikeA + P4LikeB + P4Both + SPED + SPED * P4LikeA + SPED * P4LikeB + SPED * P4Both$

Sample: Grade 8/9 Proxy State Assessment. Table 46 shows the average performance on the grade 8/9 proxy state assessment as a function of the way in which students rated the reasons for liking the narrative-literary passage they read.

Table 46. Grade 8/9 Proxy State Assessment—Mean Performance on Assessment by Passage Ratings (Narrative-Literary)

Reason for liking the passage	Average performance on assessment	
	SPED	Non-SPED
Not for the reasons provided	3.00	3.83
It was a story	2.79	3.85
Reader knew something about it	2.37	3.91
It was a story and the reader knew something about it	2.58	3.71

For grade 8/9 students taking the proxy state assessment, performance on the reading assessment was not related to their reasons for liking the narrative-literary passages. There was a main effect of SPED students on reading assessment scores ($b = -0.83, p < .05$) (see Table 47), with the overall scores of SPED students lower than those of Non-SPED students. Non-SPED students who liked the passages (but not because of the reasons provided), scored on average 3.83, while SPED students scored on average 3.00 out of a possible 7 on the reading assessment.

Table 47: Grade 8/9 Proxy State Assessment—Full Regression Table for Student Performance by Passage Ratings (Narrative-Literary)

	Value	Std Error	df	t-value	p-value
(Intercept)	3.83	0.25	601	15.23	0.00
P8LikeA	0.02	0.27	236	0.07	0.95
P8LikeB	0.08	0.37	236	0.21	0.83
P8Both	-0.12	0.30	236	-0.38	0.70
SPED	-0.83	0.35	601	-2.38	0.02
P8LikeA:SPED	-0.23	0.39	236	-0.59	0.55
P8LikeB:SPED	-0.71	0.50	236	-1.43	0.15
P8Both:SWD	-0.30	0.42	236	-0.72	0.47

Fixed effects: $P8 \sim P8LikeA + P8LikeB + P8Both + SPED + SPED * P8LikeA + SPED * P8LikeB + SPED * P8Both$

Secondary Research Question 7 (SRQ7). Do students perform better on informational-expository passages when they rate that they liked the passage because it provided them with information? Or did students perform better on informational-expository passage when they rated that they liked the passage because they knew something about the passage? Or did they mark both responses as reasons for liking the passage?

Hypothesis. Students may prefer a passage because they know something about it or because it provides them with information. Passages that are rated this way show the highest performance, followed by those rated as liked because it provided them with information, followed by those rated as liked because they knew something about it.

Analysis. As for other similar questions, multiple scores were computed for each student and an HLM model was used for these analyses. Multiple scores were computed for each student based on the passages and items (testlets) that were informational-expository.

Sample: Grade 4/5 Accessible Assessment. Table 48 shows the average performance on the grade 4/5 accessible assessment as a function of the way in which students rated the reasons for liking the informational-expository passage that they read.

Table 48: Grade 4/5 Accessible Assessment—Mean Performance on Assessment by Passage Ratings (Informational-Expository)

<i>Reason for liking the passage</i>	Average performance on assessment	
	SPED	Non-SPED
Not for the reasons provided	2.85	4.71
Provided reader with information	3.74	5.13
Reader knew something about it	4.19	4.82
Provided reader with information and the reader knew something about it	3.56	5.17

For the informational-expository passages, performance on the reading assessment did not relate to the students’ reasons for liking the informational-expository passages (see Table 49). A main effect was found for SPED status, with overall SPED performance lower than Non-SPED performance. Non-SPED students who liked the passages (but not because of the reasons provided) scored on average 4.71 out of a possible 7 while Non-SPED students who liked passages scored on average 2.85 on the reading assessment.

Table 49: Grade 4/5 Accessible Assessment—Full Regression Table for Student Performance by Passage Ratings (Informational-Expository)

	Value	Std Error	df	t-value	p-value
(Intercept)	4.71	0.43	470	10.90	0.00
A4LikeA	0.42	0.44	337	0.97	0.33
A4LikeB	0.11	0.51	337	0.22	0.83
A4Both	0.46	0.44	337	1.04	0.30
SWD	-1.86	0.72	470	-2.58	0.01
A4LikeA*SPED	0.47	0.73	337	0.64	0.53
A4LikeB*SPED	1.23	0.85	337	1.44	0.15
A4Both*SPED	0.25	0.74	337	0.33	0.74

Fixed effects: A4 ~ A4LikeA + A4LikeB + A4Both + SPED + SPED * A4LikeA + SPED * A4LikeB + SPED * A4Both

Sample: Grade 8/9 Accessible Assessment. Table 50 shows the average performance on the grade 8/9 accessible assessment as a function of the way in which students rated the reasons for liking the informational-expository passage they read.

Table 50: Grade 8/9 Accessible Assessment—Mean Performance on Assessment by Passage Ratings (Informational-Expository)

<i>Reason for liking the passage</i>	Average performance on assessment	
	SPED	Non-SPED
Not for the reasons provided	3.39	5.16
Provided reader with information	3.30	5.29
Reader knew something about it	2.82	4.77
Provided reader with information and the reader knew something about it	3.08	4.60

For the informational-expository passages, when Non-SPED students reported liking the passage (but not for the reasons provided), they scored on average 5.16 on the reading assessment (out of a possible 7). This score was significantly lower if the students reported liking the passage for both reasons ($b = -0.56, p < .05$) (see Table 51). SPED students had significantly lower reading assessment scores ($b = -1.77, p < .01$).

Table 51: Grade 8/9 Accessible Assessment—Full Regression Table for Student Performance by Passage Ratings (Informational-Expository)

	Value	Std Error	df	t-value	p-value
(Intercept)	5.16	0.22	698	23.44	0.00
A8LikeA	0.13	0.23	366	0.55	0.58
A8LikeB	-0.39	0.34	366	-1.16	0.25
A8Both	-0.56	0.25	366	-2.28	0.02
SWD	-1.77	0.32	698	-5.48	0.00
A8LikeA:SPED	-0.22	0.35	366	-0.64	0.52
A8LikeB:SPED	-0.18	0.49	366	-0.38	0.71
A8Both:SPED	0.25	0.36	366	0.71	0.48

Fixed effects: $A8 \sim A8LikeA + A8LikeB + A8Both + SPED + SPED * A8LikeA + SPED * A8LikeB + SPED * A8Both$

Sample: Grade 4/5 Proxy State Assessment. Table 52 shows the average performance on the grade 4 proxy state assessment as a function of the way in which students rated the reasons for liking the informational-expository passage they read.

Table 52: Grade 4/5 Proxy State Assessment—Mean Performance on Assessment by Passage Ratings (Informational-Expository)

<i>Reason for liking the passage</i>	Average performance on assessment	
	SPED	Non-SPED
Not for the reasons provided	2.56	4.90
Provided reader with information	4.26	5.11
Reader knew something about it	2.38	5.06
Provided reader with information and the reader knew something about it	3.82	5.25

For grade 4/5 students taking the proxy state assessment, when Non-SPED students reported liking the passage (but not for the reasons provided), they scored on average 4.90 on the reading assessment. For Non-SPED students, liking the passage for any reason did not affect their reading assessment score. There was a main SPED effect. That is, SPED students had a significantly lower score ($b = -2.34, p < .01$) (see Table 53). For SPED students who reported liking the passage because it provided them with information, the scores were not quite as low ($b = 1.49, p < .05$).

Table 53: Grade 4 Proxy State Assessment—Full Regression Table for Student Performance by Passage Ratings (Informational-Expository)

	Value	Std Error	df	t-value	p-value
(Intercept)	4.90	0.45	468	10.82	0.00
P4LikeA	0.21	0.46	288	0.47	0.64
P4LikeB	0.16	0.58	288	0.27	0.79
P4Both	0.35	0.47	288	0.75	0.45
SWD	-2.34	0.62	468	-3.74	0.00
P4LikeA:SPED	1.49	0.64	288	2.33	0.02
P4LikeB:SPED	-0.34	0.79	288	-0.43	0.67
P4Both:SPED	0.91	0.65	288	1.40	0.16

Fixed effects: $P4 \sim P4LikeA + P4LikeB + P4Both + SPED + SPED * P4LikeA + SPED * P4LikeB + SPED * P4Both$

Sample: Grade 8/9 Proxy State Assessment. Table 54 shows the average performance on the grade 8/9 proxy state assessment as a function of the way in which students rated the reasons for liking the informational-expository passage they read.

Table 54: Grade 8/9 Proxy—Average Performance on Assessment by Passage Ratings (Informational-Expository)

<i>Reason for liking the passage</i>	Average performance on assessment	
	SPED	Non-SPED
Not for the reasons provided	2.75	5.08
Provided reader with information	3.12	5.17
Reader knew something about it	2.04	3.20
Provided reader with information and the reader knew something about it	3.03	5.38

For grade 8/9 students taking the proxy state assessment, Non-SPED students who reported liking the informational-expository passage (but not for the reasons provided), scored on average 5.08 on the reading assessment. If Non-SPED students reported liking the passage because they knew something about it, then their score was significantly lower ($b = -1.88, p < .01$) (see Table 55). This was not the case for SPED students ($b = 1.17, p < .07$). There was a significant main negative effect on the reading assessment score for SPED ($b = -2.33, p < .01$).

Table 55: Grade 8/9 Proxy State Assessment—Full Regression Table for Student Performance by Passage Ratings

	Value	Std Error	df	t-value	p-value
(Intercept)	5.08	0.33	660	15.22	0.00
P8LikeA	0.09	0.34	339	0.28	0.78
P8LikeB	-1.88	0.48	339	-3.94	0.00
P8Both	0.30	0.37	339	0.81	0.42
SWD	-2.33	0.44	660	-5.26	0.00
P8LikeA:SPED	0.28	0.45	339	0.61	0.54
P8LikeB:SPED	1.17	0.61	339	1.90	0.06
P8Both:SPED	-0.02	0.49	339	-0.05	0.96

Fixed effects: $P8 \sim P8LikeA + P8LikeB + P8Both + SWD + SWD * P8LikeA + SPED * P8LikeB + SPED * P8Both$

Secondary Research Question 8 (SRQ8). Do students who take the PARA accessible reading assessment more often use fewer accommodations than assigned to them than students who take the proxy state assessment?

Hypothesis. We hypothesized that students who took the PARA accessible assessment would use fewer accommodations than assigned to them whereas students who took the proxy state assessment would use approximately the same number of accommodations as they were assigned.

Analysis. Fisher’s exact test was used to compare the number of students who used, and did not use, assigned accommodations across accessible and proxy assessment forms.

Seven accommodations were used by SPED students who participated in this PARA field study. They were: read aloud directions, sign interpret, repeat/paraphrase, magnification, proctor/scribe, extended time, multiple breaks, and individual testing. Data on assigned accommodations were pulled from student rosters on the number of students for whom each assessment accommodation was assigned. Data on assessment accommodations that students used were pulled from the test forms. Differences in the numbers of students assigned accommodations and the number of students using them were calculated and compared.

Results. Table 56 presents types and frequencies of accommodations used for grade 4/5 students in the accessible assessment and the proxy state assessment. No significant differences were found in the use of any of the accommodations assigned, across the accessible and proxy state reading assessments. Notably, fewer students than were assigned special accommodations actually used them, unless the accommodation was “sign interpret directions.”

Table 56. Grade 4/5 Accommodations Assigned and Used by Students in the Accessible and Proxy State Assessments

	Accessible Assessment		Proxy State Assessment	
	Students Assigned Accommodation	Students Using Accommodations	Students Assigned Accommodation	Students Using Accommodations
Read Aloud Directions	140	71	127	62
Sign Interpret Directions	13	13	13	13
Repeat/Paraphrase Directions	113	32	112	27
Magnification	6	1	8	2
Proctor/Scribe	24	1	24	3
Extended Time	153	19	144	11
With Breaks	73	12	69	11
Individual	16	6	21	7

Table 57 shows the same data for students in the grade 8/9 accessible and proxy state assessments. The pattern of results from these comparisons was almost identical to that from the grade 4/5 comparisons. That is, no significant differences were found across the test forms in the number of students who used the accommodations that were assigned to them. There was one exception: the extended time accommodation was used by relatively more proxy state assessment students for whom it was assigned than by accessible assessment students ($p = 0.051$). Similar to the grade 4/5 students, many fewer students in grade 8/9 used accommodations than students who were assigned to accommodations.

Table 57. Grade 8/9 Accommodations Assigned and Used by Students in the Accessible and Proxy State Assessments

	Accessible Assessment		Proxy State Assessment	
	Students Assigned Accommodation	Students Using Accommodations	Students Assigned Accommodation	Students Using Accommodations
Read Aloud Directions	329	116	372	132
Sign Interpret Directions	21	20	24	24
Repeat/Paraphrase Directions	248	42	274	58
Magnification	11	0	19	0
Proctor/Scribe	49	0	49	0
Extended Time	399	19	440	36
With Breaks	110	7	125	12
Individual	40	8	40	8

Secondary Research Question 9 (SRQ9). Does the number of omitted or not reached items differ for the accessible and proxy state assessments? Does the level of student motivation affect the number of omitted or not-reached items?

Hypothesis. We hypothesized that students taking the accessible assessment would omit or not reach fewer items than students taking the proxy state assessment. Further, we hypothesized that as the level of student motivation increased, the number of omitted items and the number of not-reached items would decrease.

Analysis. A Poisson regression was performed to investigate whether the level of motivation was related to the number of omitted or not-reached items.

Results. The frequencies of different numbers of omitted items by grade level (4/5 and 8/9) for the accessible reading assessment and the proxy state assessment are presented

in Table 58. As shown in the table, students rarely omitted more than 1 or 2 items, regardless of the assessment or their grade.

Table 58. Frequencies of Different Numbers of Omitted Items for the Accessible and Proxy State Assessments in Grades 4/5 and 8/9

# Omitted		1	2	3	4	5	6	7	8	9	10	11	12	14	19
	Grade														
Accessible	4/5	53	32	9	3	3	0	1	0	0	0	0	1	0	1
Accessible	8/9	57	21	9	7	1	0	0	0	0	0	0	1	1	0
Proxy	4/5	111	25	7	2	1	0	1	0	1	0	0	0	1	0
Proxy	8/9	66	20	3	2	0	1	2	2	0	2	0	0	1	0

The estimates from the Poisson regression model are presented on Table 59, along with robust standard errors (Cameron & Trivedi, 2009). The estimates showed that total MRQ score was not related to the number of omitted items, but the type of student was. After computing the incident rates for the number of omitted items in the assessment, we found that the incident rate for students with LD was 1.43 times the incident rate for Non-SPED students.

Table 59. Poisson Regression Predicting Number of Omitted Items Based on MRQ and Student Group (LD vs. Non-SPED)

<i>Predictors</i>	<i>b</i>	<i>SE</i>	<i>t</i>	<i>p</i>
(Intercept)	-1.867	0.538	-3.471	<.001
Total MRQ Score	0.001	0.004	0.186	0.852
Type of Student (0=Non-SPED, 1=LD)	0.356	0.138	2.577	<.01

Note: Using a Zero-Inflated Poisson model did not result in a better fit to the data, thus the regular Poisson model results are presented.

Table 60 shows the number of not-reached items for the accessible reading assessment and the proxy state assessments for grades 4/5 and 8/9. As is evident in the table, the number of not-reached items was very low. There was only one student with two not-reached items for the grade 4/5 accessible assessment, the grade 4/5 proxy state assessment, and the grade 8/9 proxy state assessment; two students had two not-reached items for the grade 8/9 accessible assessment. Several students had the very last item missing. Because it is unlikely that these students ran out of time, these are more appropriately considered to be omitted items (see Tables 58 and 59).

Table 60. Frequencies of Different Numbers of Not-Reached Items for the Accessible and Proxy State Assessments in Grades 4/5 and 8/9

# Not Reached		0	1	2	3
	Grade				
Accessible	4	626	13	1	1
Accessible	8	1013	11	2	0
Proxy	4	621	1	1	0
Proxy	8	1059	4	1	1

Further analysis of these data was not conducted. This was due primarily to the small numbers. It was also due to the fact that none of the test administrators indicated that students did not finish the assessment. This was the case even though there was a question the administrator was to answer when an accommodated student did not finish the assessment.

Discussion

The PARA field test of its accessible assessment compared to a proxy state assessment was conducted in the spring and fall of 2010 and included over 3700 students (over 1500 special education students and over 2100 non-special education students). Preliminary research had indicated the potential of high interest reading passages, segmented text, and selected color illustrations for increasing the accessibility of a reading assessment. For the field test, two four-passage assessments with seven items per passage were created – one was the PARA accessible assessment and the other was a proxy state assessment. The accessible assessment had higher interest passages, segmented text in the passages, and color illustrations. The proxy state assessment had lower interest passages, no segmenting of text, and no color illustrations.

Both the accessible assessment and the proxy state assessment were created from passages of authentic texts that were selected by experts in reading to be of high interest to young readers (see Dillon et al., 2009a,b), consistent with the NAEP 2009 Reading Framework (National Assessment Governing Board, 2008), and consistent with the suggestion of Guthrie and Wigfield (2005) that altering text characteristics can influence situational motivation and assessment outcomes (see also O’Brien & Dillon, 2008). The illustrations for all passages were subjected to review criteria to eliminate those that were distracting or irrelevant to the passage (see Appendix H). In addition, the layout of the passages, questions, and responses adhered to the principles of universal design. Students were allowed to use the accommodations that they typically used for state assessments.

The PARA field test addressed four primary research questions and seven secondary research questions. Overall, the results indicated that the accessibility features

included in the PARA accessible assessment are ones that can be included as features of state assessments without concern about compromising the construct tested by the assessment. Despite this positive result, the field test did not support the hypothesis that the accessible assessment would improve the performance of special education students more than it increased the performance of non-special education students.

This conclusion is evident in the findings for the primary research questions. Each of these questions is discussed here, along with the findings for the secondary research questions.

Primary Research Question 1 (RQ1): Does the PARA accessible reading assessment produce higher performance than a proxy state reading assessment for students with learning disabilities (LD) and students with disabilities combined (LD, SLI, ID, and D/HH)?

Field test results indicated that there was no difference between performance on the accessible assessment and the proxy state assessment for students with LD or for all special education students (LD, SLI, ID, and D/HH combined). This finding suggests that the addition of higher interest passages, segmented text, and color were not enough to increase performance beyond that potentially produced by the already high interest passages and universal design elements of both the accessible assessment and the proxy state assessment. In short, engagement with the interesting passages on both the accessible and proxy assessments enabled readers to comprehend the passages.

The expected benefits of segmenting the passages to reduce possible disengagement from the task due to the perceived difficulty of longer running text did not account for improved engagement and performance overall. Previous research (Abedi et al., 2009) had indicated that segmenting the text in reading assessment passages did not change the constructs being measured but did significantly increase the reliability of results for students with disabilities. The expected benefit of more engaging color photos and illustrations may have been masked by the stronger engagement factor of the high interest of the passages. Guthrie and Wigfield (2005) and others (O'Brien & Dillon, 2008) have emphasized the role that motivational factors play in the engagement of students with tasks and their ultimate performance outcomes. It is also possible that the stance the readers take toward standardized assessments overall based on their repeated encounter with this reading genre, is a more potent predictor of performance than any one accessibility feature (see O'Brien & Dillon, 2008). This stance toward the genre may have fixed students' expectation for tests regardless of the design features that varied considerably from other such assessments. Finally, it is also possible that the content assessed (consistent with the NAEP 2009 framework) was content to which special education students had not been exposed. This could explain not only their lower performance overall but also the inability of enhancements, such as motivating and segmented text, to overcome their knowledge and skill deficits. Considerable

evidence indicates the lower performance levels of special education students on state reading assessments overall (Thurlow et al., 2011) and on NAEP assessments specifically (National Center for Education Statistics, 2009).

Descriptive data of the results for each disability category suggest that the results obtained for LD students and disabilities combined (LD, SLI, ID, D/HH) hold true for students with speech-language impairments and students with intellectual disabilities. Although the numbers are too small to conduct analyses and draw conclusions, students who are deaf/hard of hearing seemed to show higher performance on the accessible assessment compared to the proxy state assessments, particularly in grade 4/5. Further exploration of the features of the accessible assessment (e.g., high interest reading passages, segmented text, colored illustrations) may be worthwhile. Previous investigations of reading and students who are deaf/hard of hearing have highlighted the lower performance levels of these students (particularly prevalent in older research, for example, Bennett, Rock, & Kaplan, 1987; Centra, 1986) and the effects of sign interpretation accommodations for these students (e.g., Cawthon, Winton, Garberoglio, & Gobble, 2011; Johnson, Kimball, & Brown, 2001; Russell, Kavanaugh, Masters, Higgins, & Hoffman, 2009). Consideration of approaches that increase the accessibility of reading assessments for these students appear to be well worth further research.

Primary Research Question 2 (RQ2): Are scores from the PARA accessible reading assessment more reliable than scores from the proxy state reading assessments for students with LD and for students with disabilities combined (LD, SLI, MR, and D/HH)?

Item level analysis indicated that the internal consistency for the accessible assessment and the proxy state assessment were similar for special education students at both grades 4 (about .85) and 8 (about .76). For students with learning disabilities, internal consistency was similar for the accessible assessment and the proxy state assessment at grade 8 (about .75), but slightly less at grade 4 (.67 versus .75). The lower reliabilities might be explained to some extent by the lower variability in scores for these students (and particularly for students with LD in the accessible assessment).

A principal component analysis indicated that the items on both tests were unidimensional. Both tests were reliably tapping into the construct of reading comprehension, as typically assessed when multiple choice items follow the reading of passages and test takers hold the implicit purpose of reading to recall or draw inferences. The reliability of the proxy assessment was slightly higher for the combined disabilities group in both grades 4 and 8, which could have been due to the larger sample in this group and the higher true score variance in the combined disabilities group. Similarly, the lower reliability coefficient for the accessible assessment group might, in part, be due to the lower variability in the accessible assessment scores compared with the distribution of the scores on the proxy assessment. The lower variability in the accessible assessment scores for students with disabilities can be interpreted positively since is

suggests that students reading the passages and responding to the assessment questions more consistently understood the passages and were possibly more consistently engaged with the task.

Primary Research Question 3 (RQ3): Does the PARA accessible reading assessment produce higher performance than a proxy state reading assessment for non-special education students, thereby suggesting that it alters the focal construct?

Field test results indicated that there was no difference between performance on the accessible assessment and the proxy state assessment for non-special education students. As for special education students, the addition of higher interest passages, segmented text, and color did not increase performance beyond that produced by the already high interest passages and universal design elements of both the accessible assessment and the proxy state assessment.

The finding of no difference between the accessible assessment and the proxy state assessment for non-special education students indicates that the addition of the accessibility features does not change the construct being assessed. When reading proficiency, as measured by the most recent performance on mandated state tests, was entered as a covariate on performance on the accessible and proxy assessments, the variance accounted for by reading proficiency was not significantly changed regardless of the assessment format. The results suggest that readers used reading subprocesses, skills, and strategies, much to the same outcome, on both the accessible and proxy assessment, further indicating that both assessments were measuring the same construct. This is an important finding, one that suggests that creating a reading assessment that includes high interest segmented passages with color elements does not change the constructs being measured, and therefore the added features are acceptable for inclusion in state reading assessments.

Primary Research Question 4 (RQ4): Does the PARA accessible reading assessment demonstrate the same characteristics (accessibility, reliability, validity) for students in grades 4 and 8?

Results indicated that there were no significant differences between the proxy and accessible assessments for both grade 4 and 8 with respect to making reading assessments more accessible for students with disabilities. Despite these results, we discovered useful findings about the accessible assessments we developed. First, the results were consistent across the two grades with respect to accessibility, validity, and for the most part, reliability (e.g., the proxy state assessment for 4th graders showed slightly higher reliability than the accessible assessment, potentially due to score heterogeneities). What these findings suggest is that researchers and test developers can comfortably incorporate more accessible features such as interesting passages and color pictures into assessments for both 4th and 8th graders because it will not change important test features and structures required for large scale testing. This is important from a researcher's and test developer's standpoint because sometimes it is assumed that younger readers

or students with disabilities need more interesting content to capture their attention and sustain their reading. Instead, it could be argued that all readers of varying developmental levels benefit from interesting test content because reading passages that capture their attention may allow them to achieve to their highest potential. Students persevere through passages and questions, complete the assessment, and provide a more valid representation of their abilities.

Secondary Research Question 1 (SRQ1): Does students' performance on the reading assessment correlate with the overall general measure of motivation, based on the results of the Motivations for Reading Questionnaire (MRQ)?

To calculate the correlation between the overall reading assessment raw scores and the overall MRQ score we sought to reduce the 11 subscales in the MRQ down to one. Thus, we used a hierarchical confirmatory factor model that indicated that one higher-order factor could explain a substantial amount of the variance of the MRQ subscales. This allowed us to use a single motivation latent variable to correlate with the reading assessment scores. Results indicated that there was a significant correlation between students' overall motivation to read and their performance on the proxy and the accessible reading assessments for all students except those identified as having LD. This finding is similar (except for the students with LD performance) to results secured by other reading researchers when working with elementary level students (e.g., Wigfield & Guthrie, 1997).

When examining reading motivation we used Guthrie and Wigfield's (2000) definition: "the individual's personal goals, values, and beliefs with regard to the topics, processes, and outcomes of reading" (p. 405). Part of this definition centers on students' beliefs about themselves as readers. Research has found that students who have high self-efficacy, or the belief that they are capable of completing tasks and the connection of this belief to a specific activity, impact students' behaviors, including their performance on reading assessments (Bandura, 1997). Why is this finding important? If educators seek to address motivational issues in the curriculum they use, the tasks they design, and the ways they interact with students to build their self-efficacy, then these perspectives carry over into how students take up, and perform, on assessments. Further, if assessments mirror the best of what we know about how to design motivational materials and tasks—for example, interesting passages with vivid details and color pictures (McLoyd, 1979) and passages broken up into manageable segments (with questions interspersed as the task to be completed)—then students may feel more self-efficacious and thus willing to undertake, persevere, and complete the reading activity. It is clear that students' perceptions about themselves and the tasks they are to undertake impact performance.

Secondary Research Question 2 (SRQ2): Do students perform better on the passages they rate as more interesting?

The data reported in this section explain students' situational motivation, which is more specific than their general motivation to read. Situational motivation seeks to understand

how students' feelings about performing well on the assessment are tied to *specific* passage topics, or how difficult or easy the passage is to read, or the tasks they are asked to complete after reading passages on the reading assessment (e.g., the nature of the questions/items). On the accessible assessment, our results indicated that when grade 4/5 students perceived the passages as interesting, their performance was higher.

We also found that when the students rated the passages as not being as interesting, it did not significantly impact their reading scores. We attribute this second finding to the fact that both the accessible and the proxy state assessments were comprised of passages that were specially selected as interesting to students – per a calibration study conducted early in our research program that indicated student ratings of the top passages (Dillon, O'Brien, Kato, Scharber, Kelly, Beaton, & Biggs, 2009). Another, but likely less significant reason students failed to find more interesting passages more comprehensible is that years of test taking might support a stance in students that tests are neither interesting nor engaging, thereby lowering their expectation that passages on tests could be interesting enough to make a difference in understanding. In short, findings related to the relation between interest and comprehension outside of assessments might not apply here. However, the most interesting passages were included in the accessible assessment and the least interesting (of all the interesting passages) were included on the proxy state assessment. Both the proxy and the accessible assessment passages included pictures, although the proxy test pictures were in black and white.

The grade 8/9 data indicated similar results to the grade 4/5 data: the scores for students were higher when they rated the passages as more interesting. The one difference was that the scores for grade 8/9 students on the accessible assessment were statistically significant, in the area of students scoring higher on passages rated as more interesting, and students scored lower on passages they found least interesting. This finding is interesting in light of the fact that often as students progress up through the grades they become conditioned to taking large-scale assessments with passages they do not find to be interesting. They complete the assessments because they know they have to, but often they do not place value on the task or give it their full effort. In the case of the accessible assessment, we see that when passages are on topics students value, they may engage more with the reading, reading more deeply, and perform at a better level.

The students' performance on the proxy assessment indicated similar trends with respect to students performing better on passages they rated as more interesting. For grade 4/5 students, if they reported low interest in the passage, this significantly affected their score on the reading assessment. This same finding was true for the grade 8/9 students.

Overall, across both assessments, students' interest or situational motivation related to the reading assessment, impacted their performance, and in many cases, in a significant way. This finding is similar to the work of other reading scholars who found that students engage in deeper reading processes when they find texts interesting (Schiefele, 1999). This finding provides evidence for the need to secure narrative and expository passages that students' rate as interesting (not just those that experts or that research believe are

interesting). This is important for students across grade levels. Criteria for selecting interesting passages should include attention to a range of topics, including content that appeals to both genders and multiple ethnicities and that considers the varying interests of students at different grade levels.

Whenever possible (and if appropriate for the target population and purpose of the test), passages used in assessments should retain the characteristics of the original texts, including the visual elements (e.g., maps, charts, pictures) and the layout characteristics, such as the juxtaposition of text and visual elements in the original text. Students found the passages interesting, in part, because of these features. Finally, test developers could offer students some choices among several texts of similar genres and difficulty to read and respond to as also recommended by Scott Paris and his colleagues (2004). Offering choice to students represents challenges and there are controversies surrounding this idea in the research literature (e.g., Wainer, & Thissen, 1994). However, the benefit may be that some choice in the testing situation can enhance students' situational motivation and provide students with a sense of control and autonomy that could enhance their performance on a reading assessment (as noted in our data). As we noted above, however, the relation between interest and comprehension established outside of assessment contexts might not materialize in assessments unless assessments are carefully framed as not only including interesting passages but are purposely framed as not being like typical assessments to change possibly deeply engrained expectations of students that assessments are not supposed to be interesting.

Secondary Research Question 3 (SRQ3): Do special education and non-special education students rate informational-expository passages or narrative-literary passages as more interesting, or are both rated as equally interesting?

Generally, passage type was related to interest regardless of whether students were SPED or Non-SPED but the relationship was not consistent across the accessible and proxy state assessments. A more apparent finding was the importance of reading experience in relation to text genres. The older the students, and more years of exposure to the contrasting broad text genres, the more likely the students either explicitly or implicitly made a distinction between narrative and informational texts and tended to be more interested in informational-expository texts. Hence, in terms of the accessible assessment, the grade 8/9 group showed increased interest in informational-expository texts regardless of whether they were SPED or Non-SPED.

The proxy state assessment produced a slightly different pattern. Although there was a lack of relationship between passage type and interest for the proxy assessment, at the 8/9 grade level Non-SPED students showed more interest in informational-expository texts; however, regardless of passage type, SPED students showed more interest overall than Non-SPED students. The first result for the Non-SPED students is compatible with the general trend—the presence of a relationship between text type and interest in informational-expository texts. Mohr (2006) similarly found a relation between interest and informational texts in elementary children.

There are several possible reasons for these results. The first, already mentioned, is that awareness of, and interest in, text type or genre materializes to the point that it distinguishes reading preference as readers head into middle grades and have been reading more informational-expository texts. Generally, at about grade 4, students start to distinguish more between reading interests in general and reading interests as articulated in informational-expository texts organized in “subject area” texts. They develop interests and preferences based on domain and topical knowledge aligned with informational expository text. A second issue is the construct of interest itself, particularly its possible confounding with background knowledge. Clearly, from the earliest concerted research program on text comprehension, researchers universally assigned importance to the role of background knowledge, stored in schemata, on reading comprehension (Anderson & Pearson, 1984; Bransford & Johnson, 1972; Rummelhart, 1980). But the relation between so-called affective dimensions like interest and cognitive dimensions like comprehension is more complex (Hidi, 1990). Not only can interest be attributable to either the reader, who, for example, may or may not take an inquisitive stance toward a text, but it can also be attributable to a text topic or even to a given task. Assessing the impact of interest, with a focus on readers and either informational-expository text or narrative texts cannot be accurately achieved without considering the actual task in which reading passages is framed—in this case, taking a test. It is not unreasonable to assume that the task of taking a test would not typically engender interest in passages in relation to, say, reading for pleasure. Narrative texts, which are more typically associated with pleasure reading, are not viewed as compatible with the reading task of taking a test, as are informational-expository texts.

Secondary Research Question 4 (SRQ4): Does the text type in reading assessment passages (narrative-literary or informational-expository) impact student performance in general and by groups of students?

Generally, the answer is yes, text type affects student performance. As indicated by other analyses, students performed better on informational-expository passages. For the accessible assessment in grade 4/5, passage type did significantly impact student performance, and having a disability did not change this pattern. Students performed significantly better on the expository informational passages (see Table 25). Students who had an average score on the state reading test and were in the Non-SPED group scored on average 4.53 on the narrative passages of the reading assessment out of a possible 7 points. For informational passages, this score on average was 0.30 units higher. SPED students scored on average 0.76 units lower overall than the Non-SPED students. If they had scored higher than average on the state tests, their accessible assessment reading score was 0.09 units higher.

For the grade 8/9 accessible assessment group, passage type made no statistically significant difference in performance on the assessment for either the Non-SPED or SPED groups. For the grade 4 proxy state assessment, there was no significant difference in performance by passage type, but for grade 8/9, there was a significant difference, again with informational-expository texts yielding higher comprehension scores.

Although there was no clear impact of passage type on performance on the assessment, the pattern that did emerge is partly compatible with the findings for Secondary Research Question 3. Generally, regardless of whether students were in the Non-SPED or SPED group, they were more interested in expository texts and, in the case of younger grade 4/5 students on the accessible assessment, comprehended informational-expository texts better than narratives.

An anomalous aspect of performance in terms of text type is that grade 4/5 students, who generally have had more experience with narrative texts, performed better with informational-expository texts on the accessible assessment, and grade 8/9 students, who generally have had more experience with informational-expository texts, and showed more interest in them, performed the same on both narrative and expository texts on the accessible assessment. Grade 8/9 students performed in a more predictable way on the proxy assessment, comprehending the expository texts better than the narrative.

A possible explanation of the countervailing pattern—that is, better performance on the proxy than the accessible assessment, is that although the two assessments tapped the same reading construct, structural and coherence features of the accessible assessment although making it more accessible based on some subconstructs (e.g., increased engagement due to interest, pictures) may have been less comprehensible due to possibility that the segmentation may have reduced the text cohesion (Grimes, 1975; Halliday & Hasan, 1976). In some of the foundational work on the relation between text structure and comprehension, Goetz and Armbruster (1980) noted that the structure is not in the text but in the interaction between the reader and the text—specifically the text and the reader’s cognitive representation of it, as well as a reader’s expected structure. And variations in the reader’s expected structure, in this case segmentation of the expected structure, may have impacted comprehension. In an important synthesis of narrative text comprehension and the role of story grammar conceptions of text structure, Just and Carpenter (1987) noted that readers not only pay attention to the structure of elements in a story grammar, but essentially cognitively segment the story into episodes to help comprehend it. It is possible that our physical segmentation of passages in the accessible assessment was incompatible with the cognitive segmentations that readers employ to understand narratives, or that the segmentation created hybrid texts that did not behave in predictable ways. However, Just and Carpenter also noted when this episodic component, or other story structure, is not clearly defined or salient in comprehension, readers may rely on story content as embedded in various structure components mostly in terms of how it is related to the overall goal of the story in the causal chain (Just & Carpenter, 1987).

Secondary Research Question 5 (SRQ5): Do students score better on passages they rate as “kind of challenging” or “pretty challenging” in terms of how challenging (difficult) the passage was to read?

The data reported in this section further explains students’ situational motivation. Specifically, researchers (Bandura, 1997; Guthrie, & Wigfield, 2005) note that it is useful

to understand how difficult students perceive texts and tasks (reading various passages) *after* they complete them. That is because students' perceptions of difficulty are initially formed when they see the topic of the passage, but are fully shaped *as they read* and encounter ideas, vocabulary, and the way the ideas are presented by the author (Moje, Dillon, & O'Brien, 2000). These components of the text and the task often make reading more or less challenging in the mind of the student, and the answering of comprehension questions following the passage can also be perceived as a struggle. Findings from our study indicate that when students who took the accessible assessment rated passages as "kind of challenging" or "not too challenging" after they read them, they performed better than when they rated the passages as "very" or "pretty" challenging. This was true for students at both grade levels and for Non-SPED and SPED students. These same findings were generated for students who took the proxy state assessment.

These findings are similar to previous research: when students perceive a task as "doable," they score better on the task. This could be due to the interesting topics, the careful way the tests were laid out with columns and plenty of white space so that the reading passage looked like a real text students encounter in their personal reading. The pictures in both the accessible and the proxy state assessment also offered additional supports to students as they read, possibly making the task appear more "doable" (see Brookshier, Scharff, & Moses, 2002). In the accessible assessment, the passages were broken up into 3-4 segments with 2-3 questions after each segment. Again, the students may have perceived the passage as more "doable" (at the right level of challenge) because these supports were offered. This finding is similar to the work of researchers who found that to engage students, a text can be neither too easy nor too challenging, but needs to pull the reader in with certain textual moves that readers find appealing (Shallert, & Reed, 1997).

The only slight variation in the results occurred for students who typically perform well on state assessments. It appeared that the need for the passages to be at a level they perceived as "kind of challenging" or "not too challenging" was not as important to their performance. In fact, as the passages were perceived as being easier, their performance increased to a lesser degree. This makes sense when compared to the research that shows that for some students when a task is perceived as too easy, they do not find it as interesting or worth their effort.

Secondary Research Question 6 (SRQ6): Do students perform better on the *narrative-literary* passages when they indicated that they liked the passage because they knew something about it or when they indicated that they liked the passage because it was a story? How often did they mark both responses as reasons for liking the passage, and does doing so have an effect on performance?

On the grade 4/5 accessible assessment, Non- SPED students did best on passages they reported liking because they were stories. Likewise, SPED students did better on passages they liked because they were stories. Counter to the cognitively-based validation of the connection between prior knowledge and comprehension cited earlier, liking the passage

because students knew something about it, or perceived that they knew something about it, did not positively impact performance. Since having prior knowledge, and activating it in relation to a text topic or genre has continually proven key to understanding (Rumelhart, 1980), the latter finding is contradictory. For grade 8/9 students, the Non-SPED group evidenced a more predictable pattern, doing best on narrative passages they rated that they liked because they were stories but also because they knew something about them. That is, they did as well on passages that they rated highly because they were stories and also stories they believed they knew something about.

The SPED students, however, evidenced a different pattern: the SPED students' performance was lowest on passages they reported liking because they knew something about them. One explanation for this is that the SPED readers were struggling more with the basic processes, reading less fluently, and the relationship between that process and comprehension when reading segmented, possibly less coherent narrative passages, canceled out their positive stance toward the texts based on liking stories or presuming to know something about the topic. However, this explanation is not clearly supported by the performance on the proxy state assessment.

For grade 4/5 SPED and Non-SPED students taking the proxy state assessment, performance on the reading assessment did not depend on their reasons for liking the narrative-literary passages. Liking the passage because it was a story had no effect on performance. Likewise, for grade 8/9 students taking the proxy state assessment, performance on the reading assessment was not related to their reasons for liking the narrative-literary passages. Hence, there was not a clear enough relationship between liking a passage because readers believed they knew something about it or liking it because it was a story and performance on the passage to support our hypothesis.

Secondary Research Question 7 (SRQ7): Do students perform better on the *informational-expository* passages when they rated that they liked the passage because it provided them with information? Or did students perform better on the informational-expository passage when they rated that they liked the passage because they knew something about the passage? Or did they mark both responses as reasons for liking the passage?

We hypothesized that students might prefer passages because they know something about them or because the passages are expository, providing them with information. For the grade 4/5 accessible assessment for the informational-expository passages, performance on the reading assessment did not relate to the students' reasons for liking the informational-expository passages. For the grade 8/9 accessible assessment when Non-SPED students reported liking the passage (but not for the reasons provided), they scored on average 5.16 on the reading assessment (out of a possible 7). In another countervailing performance, the Non-SPED students scored significantly lower when they reported liking the passage for both reasons—they believed that it provided them with information and they knew something about it (see Table 51). SPED students had significantly lower reading assessment scores.

For the proxy state assessment for grade 4/5, there was no significant effect for liking the passage. For the grade 8/9 proxy assessment the same countervailing pattern as for the accessible assessment for 4/5 graders was found. If Non-SPED students reported liking the passage because they knew something about it, then their score was significantly lower—opposite of what we hypothesized. This was not the case for SPED students who scored highest when they liked the passage for both reasons (a trend, but not statistically significant). This result for SPED students supports the idea that their comprehension was supported for the typical research-supported reasons when the passages were informational-expository, but when we altered the text coherence and cohesion by segmenting the narrative passages in the accessible assessment, the measures that typically predict comprehension did not predict it.

For the grade 4/5 accessible assessment, the level of interest of SPED and Non-SPED students by the type of passage was similar (see Table 16). The minimal differences in interest levels by type of passages were not statistically significant (see Table 17) at the grade 4/5 level. At the grade 8/9 level for the accessible assessment for Non-SPED students, their interest increased when informational passages were presented. As indicated by the data in Table 18, the type of passage did have an influence on the interest level reported by grade 8/9 students who took the accessible assessment (see Table 19). This was true regardless of whether students were in the SPED or the Non-SPED group. First, older students are more likely to take note of differences in text structures or “patterns” of informational text (Singer & Donlon, 1989) and this positively impacts the understanding of good readers.

For the proxy state assessment, analyses indicated that there were no differences in the interest ratings for grade 4/5 SPED or Non-SPED students, regardless of the type of passage (see Table 21). For the grade 8/9 proxy state assessment, level of interest reported by students varied with the type of passage and whether the student group was SPED or Non-SPED (see Table 23). On average, the level of interest reported by Non-SPED students for narrative passages was 2.53 while for informational passages it was 0.25 units lower. Because a lower rating indicated higher interest, this means that Non-SPED students’ interest increased when informational passages were presented. For SPED students, interest was only 0.08 units higher for informational passages compared to narrative passages. *Regardless of the type of passage, SPED students reported more interest than Non-SPED students* (interest rating was 0.15 units lower for SPED students, indicating higher interest).

Secondary Research Question 8 (SRQ8): Do students who take the PARA accessible reading assessment more often use fewer accommodations than assigned to them than students who take the proxy state assessment?

Data on the use of accommodations indicated that greater numbers of accommodations were used by students in grade 8/9 compared to students in grade 4/5. The most frequently used accommodations were *Read Aloud Directions*, *Repeat/Paraphrase Directions*, and *Extended Time*. Analyses indicated that only one difference emerged

between the two assessments in the number of students using accommodations versus the number of students assigned to use accommodations. For each of the accommodations examined in grade 4/5 and all except the extended time accommodation in grade 8/9, there were no differences between the number of students assigned to the accommodation and the number of students using the accommodation.

For *Extended Time* in grade 8/9, the statistical difference indicated that extended time was used by more students in the proxy state assessment compared to the accessible assessment. This is consistent with what was hypothesized. Still, the finding is limited by the number of statistical tests that were conducted. Further, it is notable that for both the accessible assessment and the proxy state assessment, fewer students used each accommodation than were assigned to receive the accommodation. The later finding is consistent with previous research that highlighted inconsistencies between assigned and used accommodations (e.g., Shriner & DeStefano, 2003).

Secondary Research Question 9 (SRQ9): Does the number of omitted or not reached items differ for the accessible and proxy state assessments? Does the level of student motivation affect the number of omitted or not reached items?

We hypothesized that students taking the accessible assessment would omit or not reach fewer items than students taking the proxy assessment because of the accessibility features we included: more time on the assessment; interesting passages with pictures; passages divided into segments; and a test booklet format that was engaging. The results indicated that the number of not-reached items was very low. In fact, there was only one student with two not-reached items on the grade 4/5 accessible assessment, the grade 4/5 proxy assessment, and the grade 8/9 proxy assessment. Two students had not-reached item for the grade 8/9 accessible assessment. For several students it was the last item that was not reached—although there was not a time limit on the assessment. The students with the most number of omissions were those labeled as LD.

In examining the data across all of the other students we might speculate that because both the proxy state assessment and the accessible assessment used interesting passages (albeit the accessible assessment with more interesting passages with color pictures), and that test-takers were not given time limits to complete the assessment, these features alone may have supported students who normally do not typically complete the entire assessment to do so. Because we cannot disentangle the various accessibility features, we cannot pinpoint with accuracy whether passages themselves were the key component of helping students complete the tests, or whether the lack of time limits was a key reason. Regardless, the low omission rates are impressive.

References

- Abedi, J., Kao, J. C., Leon, S., Sullivan, L., Herman, J., Pope, R., & Nambiar, V. (2009). *Exploring factors that affect the accessibility of reading comprehension assessments for students with disabilities: A study of segmented text*. Minneapolis, MN: University of Minnesota, Partnership for Accessible Reading Assessments.
- Abedi, J., Leon, S., & Kao, J. (2007a). *Examining differential distractor functioning in reading assessments for students with disabilities*. Minneapolis, MN: University of Minnesota, Partnership for Accessible Reading Assessments.
- Abedi, J., Leon, S., & Kao, J. (2007b). *Examining differential item functioning in reading assessments for students with disabilities*. Minneapolis, MN: University of Minnesota, Partnership for Accessible Reading Assessments.
- Allen, M. J., & Yen, W. M. (1979). *Introduction to measurement theory*. Monterey, CA: Brooks/Cole.
- Anderson, T. H., & Armbruster, B. B. (1984). Studying. In P. D. Pearson, R. Barr, M. L. Kamil & P. Mosenthal (Eds.), *Handbook of Reading Research* (pp. 657-679). New York: Longman.
- Anderson, R. C., & Pearson, P. D. (1984). A schema-theoretic view of basic processes in reading comprehension. In P. D. Pearson, R. Barr, M. C. Kamil, & P. Mosenthal (Eds.), *Handbook of reading research (Vol. 1)*. White Plain, NY: Longman.
- Bandura, A. (1997). *Self-efficacy: The exercise of control*. New York: W. H. Freeman.
- Bandura, A. (1986). *Social foundations of thought and action*. Englewood Cliffs, NJ: Prentice-Hall.
- Bennett, R. E., Rock, D. A., & Kaplan, B. A. (1987). SAT differential item performance for nine handicapped groups. *Journal of Educational Measurement*, 24(1), 41-55.
- Bransford, J. D., & Johnson, M. K. (1972). Contextual prerequisites for understanding. Some investigations of comprehension and recall. *Journal of Verbal Learning and Verbal Behavior*, 11, 717-726.
- Brookshier, J., Scharff, L.F.V, & Moses, L.E. (2002). The influence of illustrations on children's book preferences and comprehension. *Reading Psychologist*, 23, 323-339.
- Cameron, A. C., Trivedi, P. K. (2009). *Microeconometrics using stata*. College Station, TX: Stata Press.

Cawthon, S. W., Winton, S. M., Garberoglio, C. L., & Gobble, M. E. (2011). The effects of American Sign Language as an assessment accommodation for students who are deaf or hard of hearing. *Journal of Deaf Students and Deaf Education*, 16(2), 198-211.

Centra, J. A. (1986). Handicapped student performance on the Scholastic Aptitude Test. *Journal of Learning Disabilities*, 19(6), 324-327.

Chall, J. & Squire, J. R. (1996). The publishing industry and textbooks. In R. Barr, M. L. Kamil, P. Mosenthal, & P. D. Pearson, *Handbook of reading research* (Vol. II, pp. 120-146). New York: Longman.

Chudowsky, N., & Chudowsky, V. (2010). *Rising scores on state tests and NAEP*. Washington, DC: Center on Education Policy.

Chudowsky, N., Chudowsky, V., & Kober, N. (2009). *State test score trends through 2007-08, Part 4: Has progress been made in raising achievement for students with disabilities?* Washington, DC: Center on Education Policy.

Cormier, D. C., Altman, J. R., Shyyan, V., & Thurlow, M. L. (2010). *A summary of the research on the effects of test accommodations: 2007-2008* (Technical Report 56). Minneapolis, MN: University of Minnesota, National Center on Educational Outcomes.

Cortina, J. M. (1993). What is coefficient alpha? An examination of theory and applications. *Journal of Applied Psychology*, 1, 98-104.

Cunningham, A. E., & Stanovich, K. E. (1997). Early reading acquisition and its relation to reading experience and ability 10 years later. *Developmental Psychology*, 33, 934-945.

Data Accountability Center. (2010). *Part B child count for 2008, students ages 6-21, 50 states, DC, PR, BIE schools*.

DeWalt, D. A., Berkman, N. D., Sheridan, S., Lohr, K. N., & Pignone, M. P. (2004). Literacy and health outcomes: A systematic review of the literature. *Journal of General Internal Medicine*, 19(12), 1228-1239.

Dillon, D. R., O'Brien, D. G., Kato, K., Scharber, C., Kelly, C., Beaton, A., & Biggs, B. (2009a). The design and validation of a motivating large-scale accessible reading comprehension assessment for students with disabilities *Fifty-eighth yearbook of the national reading conference* (pp. 277-293). Milwaukee, WI: The National Reading Conference.

Dillon, D. R., O'Brien, D. G., Kato, K., Scharber, C., Kelly, C., Beaton, A., & Biggs, B. (2009b). *The design and validation of a motivating large-scale accessible reading comprehension assessment for students with disabilities*. Minneapolis, MN: University of Minnesota, Partnership for Accessible Reading Assessments.

- Dillon, D. R., O'Brien, D. G., Thurlow, M., & Lekwa, A. (2011). *Report on the PARA motivation study: A technical report*. Minneapolis, MN: University of Minnesota, Partnership for Accessible Reading Assessments.
- Duke, N., & Carlisle, J. (2011). The development of comprehension. In M. L. Kamil, P. D. Pearson, E. B. Moje & P. P. Afflerbach (Eds.), *Handbook of Reading Research, Volume IV* (pp. 199-228). New York: Routledge.
- Entwisle, D. R., & Alexander, K. L. (1999). Early schooling and social stratification. In R.C. Pianta & M.J. Cox (Eds.), *The transition to kindergarten* (pp. 13-38). Baltimore: Paul H. Brookes.
- Fritschmann, N. S., Deshler, D. D., & Schumaker, J. B. (2007). The effects of instruction in a strategy on the reading comprehension skills of adolescents with disabilities. *Learning Quarterly*, 30(4), 1-18. From www.lexile.com/DesktopDefault.aspx?view=re&tabindex=2&tabid=31&tabpageid=358
- Fry, E. B. (1977). Fry's readability graph: Clarifications, validity, and extensions to level 17. *Journal of Reading*, 21, 242-252.
- Gersten, R., Fuchs, L. S., Williams, J. P., & Baker, S. (2001). Teaching reading comprehension strategies to students with learning disabilities: A review of research. *Review of Educational Research*, 71(2), 279-320.
- Goetz, E., & Armbruster, B. B. (1980). Psychological correlates of text structure. In R. J. Spiro, B. C. Bruce, & W. F. Brewer (Eds.), *Theoretical issues in reading comprehension* (pp. 201-220). Hillsdale, NJ: Lawrence Erlbaum.
- Grimes, J. (1975). *The tread of discourse*. The Hague: Mouton.
- Guthrie, J. T., & Wigfield, A. (2000). Engagement and motivation in reading. In M. L. Kamil, P. Mosenthal, P. D. Pearson, & R. Barr (Eds.), *Handbook of Reading Research, Volume III* (pp. 403-422). Mahwah, NJ: Lawrence Erlbaum.
- Guthrie, J.T., & Wigfield, A. (2005). Roles of motivation and engagement in reading comprehension assessment. In S.G. Paris, & S.A. Stahl (Eds.), *Children's reading comprehension and assessment* (pp. 187-213). Mahwah, NJ: Lawrence Erlbaum Associates.
- Guthrie, J. T., Wigfield, A., & Perencevich, K. (2004). *Motivating reading comprehension*. Mahwah, NJ: Lawrence Erlbaum.
- Halliday, M. A. K., & Hasan, R. (1976). *Cohesion in English*. New York: Longman.
- Hidi, S. (1990). Interest and its contribution as a mental resource for learning. *Review of Educational Research*, 60(4), 549-571.

- Johnson, E. S., Kimball, K., & Brown, S. O. (2001). American sign language as an accommodation during standards-based assessments. *Assessment for Effective Intervention, 26*(2), 39-47.
- Johnstone, C. J., Altman, J., Thurlow, M. L., & Thompson, S. J. (2006). *A summary of research on the effects of test accommodations: 2002 through 2004* (Technical Report 45). Minneapolis, MN: University of Minnesota, National Center on Educational Outcomes.
- Johnstone, C. J., Moen, R. E., Thurlow, M. L., Matchett, D., Hausmann, K. E., & Scullin, S. (2007). *What do state reading test specifications specify?* Minneapolis, MN: University of Minnesota, Partnership for Accessible Reading Assessments.
- Johnstone, C. J., Thurlow, M. L., Thompson, S. J. & Clapper, A. T. (2008). The potential for multi-modal approaches to reading for students with disabilities as found in state reading standards. *Journal of Disability Policy Studies, 18*(4), 219-229.
- Just, M. A., & Carpenter, P. A. (1987). *The psychology of reading and language comprehension*. Boston: Allyn and Bacon.
- Kato, K., Moen, R., & Thurlow, M. (2007). *Examining DIF, DDF, and Omit Rate by discrete disability categories*. Minneapolis, MN: University of Minnesota, Partnership for Accessible Reading Assessments.
- Kato, K., Moen, R., & Thurlow, M. (2009). Differentials of a state reading assessment: Item functioning, distractor functioning, and omission frequency for disability categories. *Educational Measurement: Issues and Practice, 28*(2), 28-40.
- Koretz, D. (1997). *The assessment of students with disabilities in Kentucky*. Los Angeles: CRESST/RAND Institute on Education and Training.
- Laitusis, C. C. (2007). Research designs and analysis for studying accommodations on assessments. In C. C. Laitusis & L. L. Cook (Eds.), *Large-scale assessment and accommodations: What works?* (pp. 67-79). Arlington, VA: Council for Exceptional Children.
- Levie, W. H. & Lentz, R. (1982). Effect of text illustrations: A review of research. *Educational Communication and Technology, 30*, pp. 195-232.
- Lexile Framework for Reading. Lexile Framework, MetaMetrics. Retrieved from lexile.com.
- McLoyd, V. (1979). The effects on extrinsic rewards of differential value on high and low intrinsic interest. *Child Development, 50*, 1010-1019.

- Meyer, B. J. F., Brandt, D., H. , & Bluth, G. J. (1980). Use of top-level structure in text. *Reading Research Quarterly, 16*(1), 72-103.
- Moen, R., Liu, K., Thurlow, M., Lekwa, A., Scullin, S., & Hausmann, K. (2009). Identifying less accurately measured students. *Journal of Applied Testing Technology, 10*(2).
- Moen, R. E., Liu, K. K., Thurlow, M. L., Lekwa, A. J., Scullin, S. B., & Hausmann, K. E. (2010). *Studying less accurately measured students*. Minneapolis, MN: University of Minnesota, Partnership for Accessible Reading Assessment.
- Mohr, K. (2006). Children's choices for recreational reading: A three-part investigation of selection preferences, rationales, and process. *Journal of Literacy Research, 38*, 81-104.
- Moje, E. B., Dillon, D. R., & O'Brien, D. G. (2000). Re-examining roles of learner, text, and context in secondary literacy. *Journal of Educational Research, 93*(3), 165-180.
- NARAP. (2006). *Defining reading proficiency for accessible large-scale assessments: Some guiding principles and issues*. Minneapolis, MN: National Accessible Reading Assessment Projects.
- National Assessment Governing Board. (2008). *Reading framework for the 2009 National Assessment of Educational Progress*. Washington, DC: Author. Available at: www.nagb.org/publications/frameworks/reading09.pdf.
- National Center for Educational Statistics. (2003). *The nation's report card: Reading highlights 2003*. Jessup, MD: United States Department of Education.
- National Center for Education Statistics. (2009). *The nation's report card: Reading 2009* (NCES 2010-458). Washington, DC: U.S. Department of Education, Institute of Education Sciences.
- O'Brien, D. G. & Dillon, D. R. (2008). The role of motivation in engaged reading of adolescents. In K. Hinchman & H. Sheridan-Thomas (Eds.), *Best practices in adolescent literacy instruction* (pp. 78-96). New York, NY: Guilford.
- Oakland, T., & Lane, H. B. (2004). Language, reading, and readability formulas: Implications for developing and adapting tests. *International Journal of Testing, 4*(3), 239-252.
- Pajares, F., & Schunk, D. H. (2002). Self and self-belief in psychology and education: A historical perspective. In J. Aronson (Ed.), *Improving academic achievement: Impact of psychological factors on education* (pp. 3-21). San Diego: Academic Press.

Paris, S. G., & Carpenter, R. D. (2004). Children's motivation to read. In J. V. Hoffman & D. L. Schallert (Eds.), *The texts in elementary classrooms* (pp. 61-82). Mahwah, NJ: Lawrence Erlbaum Associates.

Partnership for 21st Century Skills (2006). *Assessment of 21st century skills*. Tucson, AZ: Partnership for 21st Century Skills.

Pretorius, E.J. (2000). "What they can't read will hurt them": Reading and academic achievement. *Innovation, 21*, 33-41.

Purdie, N., Hattie, J., & Douglas, G. (1996). Student conceptions of learning and their use of self-regulated learning strategies: A cross cultural comparison. *Journal of educational psychology, 88*(1), 87-100.

Rose, T. L. (1986). Effects of illustrations on reading comprehension of learning disabled students. *Journal of Learning Disabilities, 19*, 542-544.

Rumelhart, D. E. (1975). Notes on a schema for stories. In D. G. Bobrow & A. M. Collins (Eds.), *Representation and understanding*. New York: Academic Press.

Rumelhart, D. E. (1980). Schemata: The building blocks of cognition. In R. J. Spiro, B. C. Bruce, & W. F. Brewer (Eds.), *Theoretical issues in reading comprehension* (pp. 33-58). Hillsdale, NJ: Erlbaum.

Russell, M., Kavanaugh, M., Masters, J., Higgins, J., & Hoffman, T. (2009). Computer-based signing accommodations: Comparing a recorded human with an avatar. *Journal of Applied Testing Technology, 10*(3).

Schiefele, U. (1999). Interest and learning from text. *Scientific Studies of Reading, 3*, 257-279.

Shallert, D. L., & Reed, J. H. (1997). The pull of text and the process of involvement in reading. In J. Guthrie & A. Wigfield (Eds.), *Reading engagement: Motivating readers through integrated instruction* (pp. 68-85). Newark, NJ: International Reading Association.

Shriner, J. G., & DeStefano, L. (2003). Participation and accommodation in state assessment: The role of Individualized Education Programs. *Exceptional Children, 69*(2), 147-161.

Singer, H., & Donlan, D. (1989). *Reading and learning from text*. Hillsdale, NJ: Lawrence Erlbaum.

Slavin, R. E., Lake, C., Chambers, B., Cheung, A., & Davis, S. (2009). Effective reading programs for elementary grades: A best-evidence synthesis. *Review of Educational Research, 79*, 1391-1466.

- Stanovich, K. E. (1980). Toward an interactive-compensatory model of individual differences in the development of reading fluency. *Reading Research Quarterly, 16*(1), 32-71.
- Stanovich, K. E. (1986). Matthew effects in reading: Some consequences of individual differences in the acquisition of literacy. *Reading Research Quarterly, 21*, 360-406.
- Stanovich, K. E. (1988). Explaining the differences between the dyslexic and the garden variety poor reader: The phonological core variable-difference model. *Journal of learning disabilities, 21*, 590-612.
- Thompson, S., Blount, A., & Thurlow, M. (2002). *A summary of research on the effects of test accommodations: 1999 through 2001* (Technical Report 34). Minneapolis, MN: University of Minnesota, National Center on Educational Outcomes.
- Thompson, S. J., Johnstone, C. J., Anderson, M. E., & Miller, N. A. (2005). *Considerations for the development and review of universally designed assessments* (Technical Report 42). Minneapolis, MN: University of Minnesota, National Center on Educational Outcomes.
- Thompson, S. J., Johnstone, C. J., & Thurlow, M. L. (2002). *Universal design applied to large scale assessments* (Synthesis Report 44). Minneapolis, MN: University of Minnesota, National Center on Educational Outcomes.
- Thompson, S. J., Thurlow, M. L., & Malouf, D. (2004, May). Creating better tests for everyone through universally designed assessments. *Journal of Applied Testing Technology, 10*(2). See <http://www.testpublishers.org/atp.journal.htm>.
- Thorndike, R. M. (2005). *Measurement and evaluation in psychology and education*. New Jersey: Pearson/Merrill.
- Thurlow, M. L., & Larson, J. (2011). *Accommodations for state reading assessments: Policies across the nation*. Minneapolis, MN: University of Minnesota, National Center on Educational Outcomes.
- Thurlow, M. L., Bremer, C., & Albus, D. (2011). *2008-09 publicly reported assessments results for students with disabilities and ELLs with disabilities*. Minneapolis, MN: University of Minnesota, National Center on Educational Outcomes.
- Thurlow, M. L., Laitusis, C.C., Dillon, D.R., Cook, L.L., Moen, R.E., Abedi, J., & O'Brien, D.G. (2009). *Accessibility principles for reading assessments*. Minneapolis, MN: University of Minnesota, National Accessible Reading Assessment Projects.
- Thurlow, M. L., McGrew, K. S., Tindal, G., Thompson, S. L., Ysseldyke, J. E., & Elliott, J. L. (2000). *Assessment accommodations research: Considerations for design and analysis* (Technical Report 26). Minneapolis, MN: University of Minnesota, National Center on Educational Outcomes.

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.

Thurlow, M. L., Moen, R. E., Liu, K. K., Scullin, S., Hausmann, K. E., & Shyyan, V. (2009). *Disabilities and reading: Understanding the effects of disabilities and their relationship to reading instruction and assessment*. Minneapolis, MN: University of Minnesota, Partnership for Accessible Reading Assessment.

Vellutino, F. R. (2003). Individual differences as sources of variability in reading comprehension in elementary school children. In A. P. Sweet & C. E. Snow (Eds.), *Rethinking Reading Comprehension*. New York: Guilford Press.

Waller, R. (1996). Typology and discourse. In R. Barr, M. L. Kamil, P. Mosenthal, & P. D. Pearson. (Eds.), *Handbook of reading research*, (Vol. II, pp. 341-380). New York: Longman.

Wainer, H., & Thissen, D. (1994). On examinee choice in educational testing. *Review of Educational Research*, 64, 159-195.

Wigfield, A., & Guthrie, J. T. (1997). Motivation for reading: Individual, home, textual, and classroom perspective. *Educational Psychologist*, 32(2), 57-135.

Wigfield, A., Guthrie, J. T., & McGough, K. (Spring, 1996). *A questionnaire measure of children's motivations for reading*. Instructional Resource #22, National Reading Research Center, Office of Educational Research and Improvement and ERIC Document ED 394 137.

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.

Appendices

A: PARA Reports and Presentations

B: Summary of Disability Categories

C: NAEP 2009 Text Types

D: Text Type and Difficulty Rating of Passages in the Accessible and Proxy State Reading Assessments

E: Visual Analysis Criteria

F: NAEP 2009 Cognitive Targets and Percentages Outlined in 2009 NAEP Reading Framework

G: Sample Multiple Choice Item from PARA Accessible Reading Assessment

H: Test Layout Criteria

I: Segmenting Criteria for PARA Accessible Assessment

J: Motivations to Read Questionnaire (MRQ) Research and Scoring Information

K: Assessor Training Manual

L: MRQ Administration and Logistical Instructions, Including Frequently Asked Questions and Answers

M. Equipercentile Test Equating Procedures Used to Equate Across Testing Times, Including Computational Approach and Conversion Tables

N: Analyses of Primary Research Questions for Scores Equated Separately for the Accessible and Proxy State Assessments

Appendix A

PARA Reports and Publications

2006

Johnstone, C. J., Thurlow, M. L., Thompson, S. J. & Clapper, A. T. (2008). The potential for multi-modal approaches to reading for students with disabilities as found in state reading standards. *Journal of Disability Policy Studies*, 18(4), 219-229.

National Accessible Reading Assessment Project (February 2006). *Defining reading proficiency for accessible large-scale assessments: Some guiding principles and issues*. Work as part of the Institute of Education Sciences, National Center for Special Education. Grant Numbers H324F040001 and H324F040002.

2007

Kato, K., Moen, R., & Thurlow, M. (2007, December). *Examining DIF, DDF, and omit rate by discrete disability categories*. Minneapolis, MN: University of Minnesota, Partnership for Accessible Reading Assessment.

2008

Thurlow, M. L., Liu, K. K., Lazarus, S. S., & Moen, R. E. (2008, January). *Questions to ask to determine how to move closer to universally designed assessments from the very beginning, by addressing the standards first*. Available at www.readingassessment.info.

2009

Abedi, J., Kao, J. C., Leon, S., Sullivan, L., Herman, J., Pope, R., Nambiar, V., & Mastergeorge, A. M. (2009, January). *Exploring factors that affect the accessibility of reading comprehension assessments for students with disabilities: A study of segmented text*. Minneapolis, MN: University of Minnesota, Partnership for Accessible Reading Assessment.

Dillon, D. R., O'Brien, D. G., Kato, K., Scharber, C., Kelly, C., Beaton, A., & Biggs, B. (2009). *The design and validation of a motivating large-scale accessible reading comprehension assessment for students with disabilities*. Minneapolis, MN: University of Minnesota, Partnership for Accessible Reading Assessment. Available at: http://www.readingassessment.info/resources/publications/motivatingassessments/motivating_assessments.htm.

Thurlow, M. L., Moen, R. E., Liu, K. K., Scullin, S., Hausmann, K. E., & Shyyan, V. (2009, March). *Disabilities and reading: Understanding the effects of disabilities and their relationship to reading instruction and assessment*. Minneapolis, MN: University of Minnesota, Partnership for Accessible Reading Assessment.

Dillon, D. R., O'Brien, D. G., Kato, K., Scharber, C., Kelly, C., Beaton, A., Biggs, B. (2009). The design and validation of a motivating large-scale accessible reading comprehension assessment for students with disabilities. *Fifty-eighth yearbook of the national reading conference* (pp. 277-293). Milwaukee, WI: The National Reading Conference.

Kato, K., Moen, R., & Thurlow, M. (2009). Differentials of a state reading assessment: Item functioning, distractor functioning, and omission frequency for disability categories. *Educational Measurement: Issues and Practice*, 28(2), 28-40.

Moen, R., Liu, K., Thurlow, M., Lekwa, A., Scullin, S., & Hausmann, K. (2009, May). Identifying less accurately measured students. *Journal of Applied Testing Technology*, 10(2).

2010

Thurlow, M. L. (2010). Steps toward creating fully accessible reading assessments. *Applied Measurement in Education* (Special Issue), 23(2), 121-131.

Abedi, J., Kao, J. C., Leon, S., Mastergeorge, A. M., Sullivan, L., Herman, J., & Pope, R. (2010). Accessibility of segmented reading comprehension passages for students with disabilities. *Applied Measurement in Education* (Special Issue), 23(2), 168-186.

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

McGrew, K. S., Moen, R. E., & Thurlow, M. L. (2010, May). *Cognitive and achievement differences between students with divergent reading and oral comprehension skills: Implications for accessible reading assessment research*. Minneapolis, MN: University of Minnesota, Partnership for Accessible Reading Assessment.

Moen, R., Liu, K., Thurlow, M., Lekwa, A., & Hausmann, K. (2010, August). *Studying less accurately measured students*. Minneapolis, MN: University of Minnesota, Partnership for Accessible Reading Assessment.

2011

Thurlow, M. L., & Larson, J. (2011, September). *Accommodations for state reading assessments: Policies across the nation*. Minneapolis, MN: University of Minnesota, Partnership for Accessible Reading Assessment.

2012

Johnstone, C. J., & Thurlow, M. (2012). Statewide testing of reading and possible implications for students with disabilities. *The Journal of Special Education*, 46(1), 17-25.

Appendix B

Summary of Disability Categories

(Excerpted from Thurlow, Moen, Liu, Scullin, Hausmann, & Shyyan, 2009; see www.readingassessment.info/resources/publications/DisabilitiesReadingReport/PARADisabilitiesReadingReport.html for all test and references)

Students with Specific Learning Disabilities

Learning disability is an umbrella term that refers to various groups of disabilities marked by significant difficulties in the acquisition or use of listening, speaking, reading, writing, reasoning, or mathematical skills (National Joint Committee on Learning Disabilities, 1998). Specific learning disabilities (SLDs) are “chronic conditions of presumed neurological origin which selectively interfere with the development, integration, and/or demonstration of verbal and/or nonverbal abilities. [It exists] as a distinct handicapping condition and varies in its manifestations and in degree of severity” (Learning Disabilities Association of America).

Learning disabilities historically were manifested by a discrepancy between achievement and intelligence. Often deficits occurred in particular academic areas and could not be explained by other factors (Mellard, Deshler, & Barth, 2004). Traditionally, a requirement in the diagnosis of learning disabilities was the use of the IQ-discrepancy formula. This discrepancy was measured by the difference in intelligence, or a child’s potential ability to perform, and the child’s actual performance. A student’s achievement had to be significantly lower than intellectual ability to be eligible for special education services for a learning disability. This approach has created significant controversy about over-identified or misidentified students with learning disabilities. Specifically, over-identification of students from minority groups (such as African Americans or Hispanic Americans) has emerged as a concern (Coffey & Obringer, 2000), although U.S. Department of Education (2002) data indicate that when compared with average percentages, the percentages of Hispanic students receiving services for learning disabilities are slightly higher and the percentages of African-American (non-Hispanic) students receiving services for learning disabilities are slightly lower. Recent federal policies permit approaches to the identification of students with learning disabilities that emphasize failure of students to respond to interventions rather than the discrepancy approach (IDEA, 2004; Vaughn & Fuchs, 2003). Subsequently, discussion has emerged about the validity of this approach for ELLs (Barrera, 2006; Klingner, Sorrells, & Barrera, 2007) and for comprehensive assessment of learning disabilities (Kavale, Holdnack, Mostert, & Schmied, 2003; Mastropieri, 2003).

Difficulties with basic print reading and reading comprehension are the most common problems associated with learning disabilities (Gersten, Fuchs, Williams, & Baker, 2001). Because of the strong connection between spoken and written language, reading problems often can be traced to early delays in receptive and expressive language

development (Catts, Fey, Zhang, & Tomblin, 1999; Catts & Kamhi, 2005; Scarborough, 2001). Among students who are diagnosed with learning disabilities, 80% are diagnosed because their reading skills lag behind; 90% of students with learning disabilities identify reading as their primary difficulty (President's Commission on Excellence in Special Education, 2002). Still, many children who receive special education services are able to close the achievement gap and read print on the same level as their peers. With the use of specialized techniques, these students can learn to generalize specific skills and strategies to a variety of reading situations.

Students with Speech or Language Impairments

This category encompasses a number of receptive and expressive impairments, including but not limited to, inability to understand or use language, stuttering, impaired articulation, or voice impairments. Hearing loss, neurological disorders, brain injury, mental retardation, drug abuse, physical impairments such as cleft lip or palate, and vocal abuse or misuse are all factors that can contribute to the severity of the impairment (National Dissemination Center for Children with Disabilities, 2008). There is evidence of a relationship between language impairments and reading disabilities, showing that a majority of native English speaking students who do not develop proficient reading skills in English had an early history of spoken-language deficits (Catts, Fey, Tomblin, & Zhang, 2002). Although the connection between language impairments and reading difficulties is not causal or inevitable, students with a history of language impairments are at risk for failure in reading achievement more than students without language impairments.

Recent reviews of literature indicated that approximately 50% to 60% of children with speech or language impairments outgrew the problem (Law, Boyle, Harris, Harkness, & Nye, 2000), while the remaining 40% to 50% had persistent speech or language impairments. Prathanee, Thinkhamrop, and Dechongkit (2007) examined factors associated with speech and language impairments, including biological factors (family history, prenatal care, gender, etc.) and environmental factors (number of siblings, maternal-paternal education, bilingual home, socioeconomic status, etc.). Studies invite further research to determine the strength of association and effect of these factors on student achievement.

Students with Mental Retardation

The mental retardation disability is “characterized by significant limitations both in intellectual functioning and in adaptive behavior as expressed in conceptual, social, and practical adaptive skills” (American Association on Mental Retardation, 2002). The causes of mental retardation in children vary widely, including fetal alcohol syndrome, genetic disorders like Down syndrome and fragile X syndrome, environmental factors like lead poisoning, or diseases such as meningitis. For many years, students with mental retardation were identified solely using intelligence testing. IQ levels among students labeled as having mental retardation can vary from 20-25 (profound mental retardation)

to 50-75 (mild mental retardation); according to DSM-IV-TR (American Psychiatric Association, 2000), 85% of individuals with mental retardation have *mild mental retardation*.

The term *mental retardation* is widely used and coded into federal law, but the term remains the subject of controversy. Some advocacy groups and professional associations argue that the negative stigma of the term mental retardation could be avoided by using other language. The ARC of the United States, one of the country's largest advocacy organizations for people with mental retardation, eschews the term *mental retardation* in its mission statement (The ARC, 2004) in favor of *cognitive, intellectual, and developmental disabilities*. In 2004, Special Olympics updated its terminology from *mental retardation* to *intellectual disabilities* (see the Language Guide under "About Us," then "Information about Intellectual Disabilities" at www.specialolympics.org). In this paper, we use the term *mental retardation* as a legal term defined by IDEA, while cognizant of this significant shift in terminology.

Characteristics of students with mental retardation vary widely. Students with mental retardation may have difficulty with expressive language, poor short-term memory, low level meta-cognition skills, and poor use of logic and organization. Some students who are labeled as having mental retardation also have motor difficulties that can affect their handwriting or their ability to hold reading material steadily (Rizopoulos & Wolpert, 2004). Students with mental retardation, similar to all students, demonstrate wide variation in strengths, weaknesses, interests, and motivation, all of which should be reflected in each student's Individualized Education Program (IEP).

Since school systems have begun to include students with moderate to severe mental retardation in assessments (IDEA, 1997, 2004) and accountability (NCLB, 2001), and thus also have included them in more academic instruction, these students have been achieving at much higher and more complex levels than was expected by researchers, practitioners, and advocates expected (see Moore-Lamminen & Olsen, 2005). This powerful evidence of achievement has forced educational professionals to revisit long-held assumptions about the benefits of academic instruction for all children, and is generating provocative reading research on new, rigorous approaches to reading and mathematics instruction for students with mental retardation (e.g., Browder, Ahlgrim-Delzell, Courtade, Gibbs, & Flowers, 2008; Browder, Spooner, Ahlgrim-Delzell, Harris, & Wakeman, 2008).

Students with Emotional or Behavioral Disabilities

This category covers a range of conditions, including affective disorders, anxiety disorders, schizophrenia, and conduct, attention, or adjustment disturbances. It does not include youngsters who are considered to be exhibiting social maladjustments, unless they also have emotional disturbance (National Association of School Psychologists, 2002). The official label for this group of students was formerly *serious emotional disturbance*; it was changed to *emotional disturbance* in the 2004 reauthorization of the

IDEA. We use the terminology commonly used in schools—emotional or behavioral disabilities—to refer to these students in this report. Students in this category are predominantly male, disproportionately African American, and may take medication such as stimulants, antidepressants, and anti-anxiety drugs (Bradley, Henderson, & Monfore, 2004).

Students with emotional or behavioral disabilities display both learning problems and behavioral deficits (Kauffman, 2005) that affect their academic progress. Nelson, Benner, Lane, and Smith (2004) reported that 83% of the sample of students with emotional or behavioral disabilities involved in their study scored below reading standards. Compared to general education students or students with learning disabilities, students with emotional or behavioral disabilities demonstrate much lower academic achievement (Anderson, Kutash, & Duchnowski, 2001). The number of students representing this disability category has increased. Greenbaum et al. (1996) observed that the percentage of children with emotional or behavioral disabilities increased by 31% (from 54% to 85%) over the seven-year span of their study.

Students with Autism Spectrum Disorders

Autism is a neurological disorder that affects a child's ability to communicate, understand language, play, and relate to others (National Dissemination Center for Children with Disabilities). Children with autism demonstrate communication deficits that may involve a complete absence of spoken language, along with an atypical focus or intensity of interests and repetitive patterns of behavior. Deficits in the development of certain language skills may occur, particularly in the area of comprehension of higher-level more complex discourse, which in turn, can affect reading comprehension ability (Wing, 1991). Although early reports suggested that close to half of children with autism never developed spoken language, current estimates, which reflect earlier diagnosis and intervention, suggest that 60-80% of children with autism do learn some spoken language (Rogers, 2006). Approximately 20% are considered "high functioning," with IQ scores within the normal range and fluent spoken language (Tager-Flusberg, Paul, & Lord, 2005). This group includes children with Asperger syndrome (Klin, Volkmar, & Sparrow, 2000).

While *autism* is the most familiar label for these children, the condition actually comprises a spectrum of disability. Children are given the label of autism when they demonstrate all three of the diagnostic criteria for autism: severe deficits in social interaction, deficits in communication, and the presence of stereotypical and repetitive patterns of movements, activities, and interests. The term *autism spectrum disorders*, however, refers to the complex group of related disorders that have similar autism-like characteristics: (1) *Pervasive Developmental Disorder-Not Otherwise Specified* (PDD-NOS) refers to a collection of features that look like autism but may not be as severe or extensive; (2) *Rett syndrome*, which primarily affects girls, is a genetic disorder characterized by speech and motor skills that regress with age, along with other neurological disorders; (3) *Asperger syndrome* refers to the existence of autistic

characteristics but relatively intact language abilities; and (4) *Childhood Disintegrative Disorder* refers to a condition in which development appears normal for the first few years of life, but then regresses with the loss of speech and other skills until the characteristics of autism are evident. In addition to this range of syndromes on the autism spectrum, children with autism spectrum disorder can show a wide range of levels of cognitive function, from superior levels of intellectual ability, to abilities in the range exhibited by children with mental retardation.

Some children with autism demonstrate *hyperlexia*, a precocious ability to decode words with relatively little ability to comprehend the meaning of what is read. The presence of hyperlexia can complicate the assessment of reading ability in these students because their strong word reading skills can mask deficits in the ability to understand what they read.

Children with autism spectrum disorders generally have problems in three core areas, with varying degrees of intensity: socialization, communication, and restricted patterns of behaviors and interests (Ruble & Gallagher, 2004). This may lead to difficulty understanding social rules such as taking turns and sharing, problems with understanding and reading the emotions of others, difficulty taking the perspective of other people, and problems initiating and maintaining interactions and conversation with other people (Barnhill, 2004).

Students Who Are Deaf or Hard of Hearing

Students who are deaf or hard of hearing face unique challenges when reading, particularly those youngsters who have been deaf since birth. Yet with targeted interventions and accommodations in reading instruction and assessment, students who are deaf or hard of hearing can become proficient readers. Understanding the characteristics of students who are deaf or hard of hearing, as well as the communities in which they live, is an important step toward developing effective instruction and appropriate assessment for these students.

For the purpose of this paper, students who are deaf or hard of hearing will be treated as one group, except when specific research makes a clear distinction. The terms *deaf* and *hard of hearing* both denote hearing loss; the term *deaf* signifies a more severe degree of hearing loss. Although the word is used inconsistently, *deafness* usually denotes “the inability to hear and understand any speech” (Gallaudet Research Institute, 2007).

Students who are deaf or hard of hearing vary widely in terms of the cause and the degree of hearing loss, the age of onset of hearing loss, educational background, language and communication methods, and how individual members within the community feel about their hearing loss. Some deaf or hard of hearing students prefer to identify themselves as members of a linguistic and cultural minority group, while others may identify themselves as students with a disability. The ways in which students identify

themselves reflect and shape their educational and communication experiences although for some, there are conflicts between the two. Across the United States, several languages and communication forms are used commonly by students who are deaf or hard of hearing, including American Sign Language (ASL), communication systems for visually encoding English such as Manually Coded English (MCE), Cued Speech, speech reading, total communication, and bilingual-bicultural approaches. In addition, some students use residual hearing and hearing devices, or may have surgically placed cochlear implants.

Cochlear implants receive signals from an external device which stimulates electrodes in the cochlea (Owens & Kessler, 1989). Geers (2002) concluded that the effectiveness of cochlear implants is related to the age of the individual at onset of deafness, the residual hearing before the implant, and a child's learning environment. Children with cochlear implants who were born with severe to profound hearing loss are not as likely to achieve the kind of proficiency in spoken language as their hearing peers (Mayberry, 2000), but can focus on developing skills that enable them to take full advantage of the sound they are able to access (Power & Leigh, 2000).

Students with Visual Impairments or Blindness

Reading is most often thought to involve viewing print on a page or other medium, such as computer screens, and then decoding that print. Because visual disabilities interfere with an individual's ability to see words on paper or other media, other modes of reading may need to be used. The degree of a student's visual impairment, along with the impact of potential additional disabilities, has varying implications for instruction, accommodations, and testing. Most students in this category have low vision, which includes tunnel vision and partial vision. Some students with visual impairments can read print efficiently without accommodations. Other students may use some tactile and auditory adaptations, but many can read print with magnification (Assistive Technology Strategies, Tools, Accommodations and Resources—AT STAR—Web site).

In recent decades, many educators in the United States have pushed for increased reliance on residual vision, and assistive technologies such as screen readers and magnifiers have become more widely available. These trends have led to a decline in the number of students learning to read proficient English braille (reflected in American Printing House for the Blind's annual distribution of federal quota, 2003). Considerable controversy has followed in the wake that has implications for reading. Some believe that proficiency in braille is essential for individuals with visual impairments or blindness to achieve satisfactory educational progress and argue that learning to read and write is necessary in order to become self-sufficient in adulthood (Johnson, 1996). On the other hand, other advocates suggest that braille may not always be appropriate for every student with a visual impairment, and that other avenues to accessing print are just as important as braille.

According to a study by Trent and Truan (1997), the age at the onset of blindness was the critical factor related to reading speed. Factors such as comprehension, degree of vision, methods of instruction, and attitudes toward reading or braille did not impact reading speed. The authors of the study also found that to increase braille reading speed, students should use braille daily and for a long period of time; early braille instruction was also an important contributor to speed. Rapid braille readers can decode as quickly as print readers (Erin, 2003). Still, on average, braille readers read at about half the speed of print readers, at about 150 words per minute (Pring, 1994).

Readers with visual impairments or blindness and readers without visual impairments show few differences in their linguistic and cognitive processes related to reading. Both types of readers use syntactic, semantic, and contextual clues to become proficient readers. Individuals who are blind or visually impaired simply face a barrier to accessing print (Koenig, 1992). Still, low vision or a diminished perceptual field has implications for developing reading fluency at both lower processing (efficient recognition of distinctive letter features, orthographic information, and sight word recognition) and higher processing levels (syntactic, semantic, and text discourse structure). Students with assistive visual technologies such as magnification and screen readers can often develop proficiency in processes like phonemic segmentation, blending, and decoding and gain enough automatic processing using these skills to develop proficiency in vocabulary and comprehension.

Appendix C

NAEP 2009 Text Types

Texts on the 2009 NAEP Reading Assessment

Stimulus Material: Literary

		Fiction	Literary Nonfiction	Poetry
LITERARY	GRADE 4	Adventure Stories Historical Fiction Contemporary Realistic Fiction Folktales Legends Fables Tall Tales Myths Fantasy	Personal Essay Autobiographical/ Biographical Sketches Speech	Narrative Poem Free Verse Lyrical Poem Humorous P
	GRADE 8	Science Fiction Plus Increasingly Complex Application of Grade 4	Character Sketch Memoir Plus Increasingly Complex Application of Grade 4	Ode Song (including ballad) Epic Plus Increasingly Complex Application of Grade 4
	GRADE 12	Satire Parody Allegory Monologue Plus Increasingly Complex Application of Grades 4 and 8	Classical Essay Plus Increasingly Complex Application of Grades 4 and 8	Sonnet Elegy Plus Increasingly Complex Application of Grades 4 and 8

**Texts on the 2009 NAEP Reading Assessment
Stimulus Material: Informational**

		Exposition	Argumentation and Persuasive Text	Procedural Text and Documents
INFORMATIONAL	GRADE 4	Informational Trade Book Textbook News Article Feature Article Encyclopedia Entry Book Review	Informational Trade Book Journal Speech Simple Persuasive Essay	Embedded in Text Directions Map Time Line Graph Table Chart
	GRADE 8	Historical Document Essay (e.g., informational, persuasive, analytical) Research Report Plus Increasingly Complex Application of Grade 4	Letter to the Editor Argumentative Essay More Complex Persuasive Essay Editorial Plus Increasingly Complex Application of Grade 4	Embedded in Text Recipe Schedules Plus Increasingly Complex Application of Grade 4
	GRADE 12	Essay (e.g., political, social, historical, scientific, natural history) Literary Analysis Plus Increasingly Complex Application of Grades 4 and 8	Essay (e.g., political, social) Historical Account Position Paper (e.g., persuasive brochure, campaign literature, advertisements) Plus Increasingly Complex Application of Grades 4 and 8	Stand-Alone Material Manual Contract Application Product Support Material Plus Increasingly Complex Application of Grades 4 and 8

Appendix D

Text Type and Difficulty Rating of Passages in the Accessible and Proxy State Reading Assessments

Passages for Grade 4 Accessible Reading Assessment

ID	Title	Text Type (E=Expository; L=Literary)	Difficulty Level (E=easy; M=medium; H=Hard)	Short Description of the Passage
13	What is a Bear?	E	E	Types of bears and characteristics
2	Bubble Power	E	H	How animals use bubbles to protect themselves
8	Susie, Fudge, and The Big Race	L	H	Training for and participating in a dog sled race
21	One-eyed Willie	L	E	A boy's mom and his classmates help him cope with an eye problem

Passages for Grade 8 Accessible Reading Assessment

ID	Title	Text Type (E=Expository; L=Literary)	Difficulty Level (E=easy; M=medium; H=Hard)	Short Description of the Passage
52	Boggs Bills	E	E	A man creates realistic drawings of money and what he does with them
30	Dragon Spit	E	M	A description of the spit of Komodo dragons
34	Snacks	L	E	A boy is challenged to create a snack that makes his teacher "think"
48	A Genius for Sauntering	L	M	A teenager's appreciation of a boy walking in a pair of jeans

Passages for Grade 4 Proxy Reading Assessment

ID	Title	Text Type (E=Expository; L=Literary)	Difficulty Level (E=easy; M=medium; H=Hard)	Short Description of the Passage
23	Giants of The North	E	M	Facts about walrus
20	Rice Balls	E	H	How to make rice balls from various ingredients
17	Grandma's Quilt	L	M	A grandma and her granddaughter help homeless people
28	Scaredy-Cat Abby	L	H	A babysitter overcomes her own fears to provide good care for younger kids

Passages for Grade 8 Proxy Reading Assessment

ID	Title	Text Type (E=Expository; L=Literary)	Difficulty Level (E=easy; M=medium; H=Hard)	Short Description of the Passage
4	The Most Gruesome Fish in the World pp. 35-39	E	E	Information about the hagfish
9	Body in the Bog	E	H	The mystery of the Tollund man found in a bog
10	Wolf Shadows, ch. 1 pg. 1 - 7	L	E	Two boys return from deer hunting and face danger on the trip home
6	Soledad/The Circuit; pp. 8-11	L	H	A story from the life of a migrant child

Appendix E

Visual Analysis Criteria (VIA)

(Developed by Dillon, O'Brien, Scharber, Kelly, and Beaton, 2008)

Criteria are articulated here for the inclusion/exclusion of visual items accompanying all passages on the reading assessment. The processes used to determine reliability include having pairs of reading researchers use the criteria with each passage to determine—through individual analysis followed by joint analysis/discussion—which visual items would be “retained” and those that would be “deleted.” Charts of analyses processes will be generated and then used by test developers.

Criteria: A photo, picture, map, sketch, illustration, or diagram that accompanies a passage will be retained for a test passage if the visual item meets criteria # 1 or #2 below and is potentially motivating (#3 below):

1. It is necessary for comprehending a passage because the text refers to the visual item and uses words to describe something in the item. It could then be assumed that the author intended reader to “read the visual item” and that, by doing so, the reader’s comprehension would be enhanced above and beyond merely reading the words in the passage. The text and picture together enhance comprehension. Another example of this would be a visual item that has a sentence underneath it that describes the content of the item.

a. *We will not retain* the visual item if it doesn’t enhance the passage merely because the passage merely states “see photo at right.” We will edit out the references to “photo at right” if the passage can stand on its own without the visual item.

b. *We will not retain* a visual item merely because it could be used as an assessment item (e.g., it is a “visually dependent item”-- maps, photos, illustrations, diagrams).

c. *We will not retain* a photo if it is of poor quality.

2. It is important to understanding the concepts discussed in the passage. If the author has included a picture or map to help readers understand a concept that is key to the passage or explained in the text, we will retain the visual item to support the reader’s understanding of the concept and thus their overall comprehension. Another example of this would be a visual item that helps support a reader’s prior knowledge about a concept described in the text by visually representing this concept/thing. Book covers could also fall into this category if they are important to understanding basic concepts described in the passage.

3. It appears to enhance a reader's motivation. Authors may have sought to secure photos or artistic drawings to
- a. help readers “see” and thus connect better with characters or settings in narrative texts;
 - b. draw the reader into the text or interest them; help create a positive stance toward wanting to engage with the text;
 - c. help break up the text for the reader, thus reducing some anxiety with reading long sections of text and also help the reader sustain her/his effort and persevere through the text; and
 - d. piques curiosity within the reader because the picture is unusual, appealing or bizarre.
 - e. *We will not retain* a visual item if it might possibly distract the reader from attending to the text. Specifically, this would be an item that is interesting but doesn't have anything to do with the core ideas in the text (e.g., a picture of a wolf within a passage about bears). We are concerned that having a lot of photos in a text could possibly deter some readers with disabilities from focusing on the print and comprehending the passage (readers might get off task studying the many photos vs. moving between the photos and the text). It is a balance of photos that will enhance the printed text and support comprehension but not overwhelm the passage.
 - f. *We will not retain* a visual item if there are already an abundance of visual items and more only creates a feeling of visual overload or over stimulation for the reader.
 - g. *We will not retain* a visual item if the story (narrative text) can be easily understood without the visual item and leaving the item out actually stimulates a reader's imagination so that they connect better with the pictures they conjure up on their own (e.g., if the picture of a character isn't one that the reader can relate to because of the way the person looks).
4. Cost and availability of the pictures. We will pay the royalties required for photos if at all possible. Paying for the visual items (usually pictures) is important to our goal of keeping texts intact and as true to the original source as possible. When the cost is prohibitive, and the visual item is not absolutely necessary (based on criteria 1-2 above), we will not retain it.

Appendix F

NAEP 2009 Cognitive Targets and Percentages Outlined in 2009 NAEP Reading Framework

Cognitive Targets for the 2009 NAEP Reading Assessment¹

	Locate/Recall	Integrate/Interpret	Critique/Evaluate
BOTH LITERARY AND INFORMATIONAL TEXT	Identify textually explicit information, such as: <ul style="list-style-type: none"> • Definitions • Facts • Supporting details Make simple inferences	Make complex inferences to: <ul style="list-style-type: none"> • Describe problem and solution, cause and effect • Compare or connect ideas, problems, or situations within or across texts • Determine unstated assumptions in an argument • Analyze how an author uses literary devices and text features 	Consider text critically to: <ul style="list-style-type: none"> • Judge author's craft and technique • Evaluate the author's perspective or point of view within or across texts • Take different perspectives in relation to a text
SPECIFIC TO LITERARY TEXT	Identify textually explicit information, such as: <ul style="list-style-type: none"> • Character traits • Sequence of events or actions • Setting Identify figurative language	Make complex inferences to: <ul style="list-style-type: none"> • Infer mood or tone • Integrate ideas to determine theme • Identify or interpret a character's motivations and decisions • Examine relations between theme and setting or characters Explain how rhythm, rhyme, or form in poetry contribute to meaning	Consider text critically to: <ul style="list-style-type: none"> • Evaluate the role of literary devices in conveying meaning • Determine the degree to which literary devices enhance a literary work • Evaluate a character's motivations and decisions • Analyze the point of view used by the author

¹ The Cognitive Targets matrix is for illustrative purposes only and should not be considered an exhaustive list. The cognitive targets will be elaborated further in the *Specifications for the 2009 NAEP Reading Assessment*.

Cognitive Targets for the 2009 NAEP Reading Assessment (continued)

SPECIFIC TO INFORMATIONAL TEXT	<p>Identify textually explicit information, such as:</p> <ul style="list-style-type: none"> • Topic sentence or main idea • Author’s purpose • Causal relations <p>Locate specific information in text or graphics</p>	<p>Make complex inferences to:</p> <ul style="list-style-type: none"> • Summarize major ideas • Draw conclusions and provide supporting information • Find evidence in support of an argument • Distinguish facts from opinions • Determine the importance of the information within and across texts 	<p>Consider text critically to:</p> <ul style="list-style-type: none"> • Analyze the presentation of information • Evaluate the way the author selects language to influence readers • Evaluate the strength and quality of evidence used by the author to support his or her position • Determine the quality of counterarguments within and across texts • Judge the coherence, logic, or credibility of an argument
---------------------------------------	---	--	---

The 2009 NAEP Reading Framework recommends cognitive targets, or behaviors and skills, for items from both literary and informational texts: Locate/Recall, Integrate/Interpret, and Critique/Evaluate. These cognitive targets illustrate the complex nature of the reading process, while the corresponding behaviors highlight the different behaviors elicited by different text types.

Distribution of Literary and Informational Passages

Grade	Literary	Informational
4	50%	50%
8	45%	55%
12	40%	60%

Distribution of Cognitive Targets by Grade

Grade	Locate/Recall	Integrate/Interpret	Critique/Evaluate
4	30%	50%	20%
8	20%	50%	30%
12	20%	45%	35%

Appendix G

Sample Multiple Choice Item from PARA Accessible Reading Assessment

Why is it important for scientists to know how quickly the hagfish reproduces?

- A They want to use it in experiments.
- B They want to prevent its extinction.
- C They want to collect its slime for use in foods.
- D They want to protect the fish that it consumes.

Appendix H

Test Layout Criteria

(Developed by Dillon, O'Brien, Scharber, Kelly, and Beaton, 2008)

Criteria were developed by PARA staff and then used to guide the reading assessment test layout development process. These criteria are as follows:

1. Illustrations should appear as close to the text as possible if the text (a) explicitly refers to a picture [related to Visual Item Inclusion Criterion 1]; or (b) the illustration supports a concept that is explained or inferred by the text [related to Visual Item Inclusion Criterion 2].
2. Illustrations that are included because they enhance engagement [Visual Item Inclusion criterion 3], should be placed at the same location in the test passage as they are in the original passage.
3. To maintain the ecological validity of the assessment as it relates to the original text (authentic text), all illustrations included in the assessment passages should be placed as close to the position they occupy in the original passages as possible.
4. The text font should be at least 12 point to maximize accessibility for the range of readers in the study. Text should be formatted in relation to the pictures so that the page layout from the original text is followed as closely as possible.

Appendix I

Segmenting Criteria for PARA Accessible Assessment

(Developed by Dillon, O'Brien, Abedi, 2009)

The following rank ordered list was developed for use in segmenting the passages for the PARA Field Study. This list is broken into four categories including overall required procedures for segmenting (these were used with all passages), preferred procedures (used with all passages if at all possible), and specific procedures for both narrative and expository texts.

1. Overall Required Procedures for Segmenting

- a.) The segmenting process should not change or alter the structure of the passage or the items that follow the passage. In other words, sentences and phrases are not to be re-ordered or re-written to allow segmenting to occur in particular places.
- b.) Pictures or visual items that go with particular segments of text should remain with the segment, even when this causes some page layout challenges.
- c.) Questions which are now located at the end of the passage, should be moved underneath the passage section where the answer can be found.
- d.) Questions referring to the entire passage should be placed to the end
- e.) Passages should be segmented into a reasonable number of segments. For example a passage for 4th graders (between 500 to 800 words) can be divided into three to four segments. Passages for 8th graders (700-1,200 words) can be divided into three to five segments.

2. Preferred Procedures for Segmenting

- f.) The process of segmenting each passage is based on the text structural features—episodic structure based on story grammar analysis for narrative texts, and propositional analysis for expository-informational texts
- g.) Segmenting should also be done in consideration of natural paragraph breaks (i.e., paragraphs were not split up).
- h.) In the PARA Field Assessment, 7 questions/items accompany each passage. When possible—and contingent first on what makes sense when breaking passages into segments—an equal number of questions should be placed after each segment (about 2-3) so that each segment is deemed worthy of careful reading by students. Inferential questions will often need to be placed at the end of the last segment.

- i.) When possible, the original 7 questions, placed in a particular order when the Motivation Test was calibrated, should remain in this same order when they are split apart and placed after segments of the passage. This may not be possible with inferential questions in particular. Therefore, it may be necessary to re-order some of the items from the way they were originally located.
- j.) Entire segments should fit neatly on one page to the extent possible, as long as the visual item/s that accompany a particular segment are not separated from that text.

3. Specific Guidelines for Narrative Texts

- k.) Use Rummelhart's (1975) concepts of syntactic structure of stories, organized around episodes (problem-solution series within a story) and events, to ensure that the episodic structure of the story is addressed and definitely not disrupted when segments are determined
- l.) Employ the idea of "cliff hangers" when appropriate as episodic break points, to promote motivation to read on. But do not use these break points if they appear to be an artificial stopping point.
- m.) Distribute the questions at the end of passage segments that match the content of the segment but note that with narrative texts, there will not be an even distribution of these questions by segment. More questions may be listed at the end of the last segment because the many questions require students to link information across the segments to summarize or infer more complex ideas. [Note. The items (multiple choice questions following each passage) were originally written for intact, unsegmented passages. The writers focused on the entire story and were asked to think about good questions that spanned the entire passage and use the NAEP 2009 cognitive targets to create questions, instead of thinking about the story in chunks and writing questions that worked with particular segments of stories. For this reason, similar length segments of text may be difficult and more questions may occur at the end of passages and more questions may be placed after the last segments because the questions require putting ideas together across the entire passage.]

4. Specific Guidelines for Expository Texts

- n.) Use the naturally occurring heading dividers within informational texts as the first consideration for segmenting but without violating the macropropositional and micropropositional structure and coherence.
- o.) In the case of passages without subheads, use the propositional structure to pick segmenting boundaries.

Appendix J

Motivations to Read Questionnaire (MRQ) Scoring Information and Research

The book *Assessment for Reading Instruction* provides information on the administration and scoring of the *Motivations for Reading Questionnaire* developed by Wigfield, Guthrie, and McGough (1996). This is a summary of information provided in that document.

Administration

Students are given the MRQ forms and told that the purpose of the questionnaire is to determine what they think and feel about reading. Students are assured that there are no right or wrong answers. Students generally read the questions and answer them, but for younger students (e.g., grades 3 and 4), it is suggested that the questions might be read to them.

Scoring

Items are scored on a 4-point scale, with higher values indicating stronger agreement. Other than these items, points are added together to obtain a total score. The authors suggest that scores should be obtained for the proposed dimensions of reading motivations. They argue that these scores give information on the pattern of responses and the ways in which different aspects of motivation to read are rated. The dimensions are:

- Reading Efficacy
- Reading Challenge
- Reading Curiosity
- Aesthetic Enjoyment of Reading
- Importance of Reading
- Compliance
- Reading Recognition
- Reading for Grades
- Social Reasons for Reading
- Reading Competition
- Reading Work Avoidance

Work Avoidance items indicate agreement with a negative statement, and thus are not added into the total score. It is also noted that the first two items in the Compliance dimension are reversed (a 4 is scored as a 1, a 2 is scored as a 3, etc.) before getting a score for this dimension and before adding into a total score.

Suggestions are also provided for handling cases in which students circle more than one response or add numbers to the scale. They also provide procedures for creating scale scores.

Appendix K

Assessor Training Manual

1. OVERVIEW OF THE PARTNERSHIP FOR ACCESSIBLE READING ASSESSMENT (PARA) STUDY

The PARA study will examine the accessibility, reliability, and validity of a reading assessment developed to measure reading proficiency for the widest range of students, especially students with disabilities. The study will also examine the results of improving the motivational characteristics of an assessment. Students will take either a proxy or accessible reading assessment, and also will complete a “Motivations for Reading Questionnaire (MRQ)”.

1.1 PARA Study Sample

Study participants will include students with disabilities from four categories and results will be compared with general education students. The disabilities will include one of four types: learning disabilities (LD), speech-language impairments (SLI), intellectual disabilities (ID), and deafness/hard of hearing (D/HH).

The PARA study sample consists of the participating states of Pennsylvania and Illinois. Districts and schools with sufficient number of students with disabilities in grades 4 and 8 will be recruited. There will be 3,000 student participants from both states, 1500 for each grade tested.

Student Group	Accessible Assessment	Proxy State Assessment
Grade 4		
Students with Disabilities		
Students with LD	150	150
Students with SLI	150	150
Students with ID	150	150
Students who are Deaf/HH	150	150
Students without Disabilities (ND)*	150	150
Grade 4 Total n = 1,500	750	750

Grade 8		
Students with Disabilities		
Students with LD	150	150
Students with SLI	150	150
Students with ID	150	150
Students who are Deaf/HH	150	150
Students without Disabilities (ND)*	150	150
Grade 8 Total n = 1,500	750	750

We will include students who normally take state assessments but will exclude those who did not (e.g. students with severe cognitive limitations). Students with disabilities will be provided the supports they are required to receive when participating in standardized assessments. All students with an IEP will be at a minimum administered the assessment in accordance with the requirements of their IEP. Students non-fluent in English (ELL) will be excluded because English fluency is required to read the passages on the test.

1.2 PARA Study Design

PARA is designed to address several key questions:

- Is the PARA accessible reading assessment more accessible than a proxy state reading assessment for LD, SLI, ID, and D/HH?
- Are scores from the PARA accessible reading assessment more reliable than scores from the proxy state reading assessments for LD, SLI, ID, D/HH?
- Is there evidence of the validity of scores from the accessible reading assessment?
- Does the PARA accessible reading assessment demonstrate the same characteristics (accessibility, reliability, validity) for students in grade 4 and 8?

1.3 Activities Prior to Data Collection

Before data collection could begin, the project team had to obtain approval for the study. This was done first at the state level, then at the school district level, and finally at the school level.

Each participating school was asked to identify a “School Contact” who would serve as the liaison between the school and study staff. Rosters were sent to schools to identify the students with disabilities and the general education students in both fourth and eighth grades. Using a random selection process, students with disabilities in each of the four disability categories as well as general education students were chosen to participate in the study.

The parents of selected students received a Parent Letter (see Exhibit 1-1. Parent Letter) outlining the study and the role of their child. Since participation is voluntary, parents could choose to opt out of the study.

You will receive a roster for each school in your caseload (with 4th and 8th graders and both students with disabilities and general education students). This roster will include only the children who will be participating in the study and will need to be tested (Exhibit 1-2. Roster). Rosters will include information on the child's disability, if applicable. All students who have had their parents refuse participation will be excluded from the rosters you receive. Please note this does not mean students cannot still refuse to participate when testing takes place. More information on how to handle child refusals can be found in Chapter 5.

Exhibit 1-1. Parent Letter



Partnership for Accessible Reading Assessment

<http://www.readingassessment.info>
readingassess@umn.edu

National Center for Educational Outcomes (NCEO)
350 Elliott Hall
75 E. River Road
Minneapolis, MN 55455
(612) 625-7241
Fax: (612) 624-0879
<http://www.nceo.info/>

National Center for Research On Evaluation, Standards, and Student Testing (CRESST)
300 Charles E. Young Drive North
GSE&IS Bldg, 3rd Flr.
Mailbox 951522
Los Angeles, CA 90095-1522
(310) 206-1532
Fax: (310) 825-3883
<http://www.cie.ucla.edu/>



1600 Research Boulevard
Rockville, MD 20850
(301) 251-1500
Fax: (301) 294-4475
<http://www.westat.com/>

Procedures:

For this study we will collect demographic information and reading test scores for each student. Students will take a questionnaire on their motivation to read (the Motivations for Reading Questionnaire) that should take about 15 to 30 minutes. Students also will take a reading comprehension test—either the experimental test developed by PARA, or a regular reading test. Both versions of the test will consist of 4 reading passages followed by 7 multiple choice questions each. The test should take about 45 to 60 minutes for students to complete.

Risks of participation:

There are no apparent risks associated with participation in this study. Students will miss some instructional time, though they will be actively engaged in reading and responding to comprehension questions.

Confidentiality:

All electronic data collected for this study will be stored in a secure location, accessible only to project researchers. Data that could potentially identify individual participants or schools will not be included in any reports, and will be physically removed from all test booklets.

Incentives:

Participating 4th graders will receive a \$15 gift card and 8th graders will receive a \$25 gift card to a local store.

Dear Parents/Guardian,

This letter is to confirm that your school has agreed to participate in the field trial of the reading test developed by the Partnership for Accessible Reading Assessment (PARA). We would like to provide you with information about the purpose of the study, potential risks and benefits, and to invite you to participate in the PARA Study.

Background & Purpose for this study:

PARA is a research project being conducted by the [National Center on Educational Outcomes \(NCEO\)](#), the [Department of Curriculum and Instruction](#) at the University of Minnesota, the [National Center on the Research on Evaluation, Standards, and Student Testing \(CRESST\)](#), and [Westat, Inc.](#)

PARA researchers have been working on ways to improve large scale reading tests so that scores reflect students' actual reading skills rather than unrelated student characteristics. Based on the results of several studies on this topic, PARA developed a reading test designed to have greater accessibility for students with disabilities. Westat will lead the field trial of the experimental reading test and collect the data.

Benefit to your child & school:

Students and teachers who participate in this study will directly inform the development of more accessible and therefore more valid tests of reading proficiency for all students, especially students who have disabilities. Although there are no immediate, direct consequences for students or teachers who participate, they will be making an important contribution to the improvement of the accuracy and usability of reading tests nationwide.

Exhibit 1-1. Parent Letter (continued)

Participation

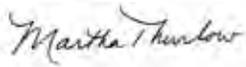
Your child's school has agreed to participate in this study and your child may be selected to participate. If you do not want your child to participate, please contact Bonnie Ho at Westat at our toll free telephone number 1-888-227-8133 or via email at parastudy@westat.com. If a voicemail is reached, please leave detailed information, including your first and last name, your child's first and last name, the grade, and the school name where your child attends. If it is okay for your child to participate, you do not need to respond to this letter.

We are thankful for your support and collaboration in this research; if you have questions, or would like to receive further information about this study or the PARA project, please contact Bonnie Ho toll-free at Westat.

Phone: 1-888-227-8133

Email: parastudy@westat.com

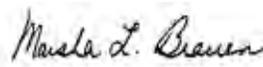
Sincerely,



Martha Thurlow
Principal Investigator
University of Minnesota, Twin Cities



Deborah R. Dillon, Ph.D.
Guy Bond Chair in Reading
& Professor of Literacy Education
University of Minnesota, Twin Cities



Marsha L. Brunen, Ph.D.
Vice President
Westat

Exhibit 1-2. Roster

Students Sampled For PARA Study

District Name: AURORA EAST USD 131				
School Name: C F SIMMONS MIDDLE SCHOOL			School State: IL	
City: AURORA				
STUDENTS WITH DISABILITIES				
FOURTH GRADE				
CHILDID	First Name	Last Name	Date of Birth	Assessment Type
10027	Rick	Hogan	12/12/2001	State Proxy
10028	John	Johnson	1/2/2004	Accessible
10029	Jane	Howard	3/3/2005	Accessible
10032	Robert	Hanson	4/4/2003	Accessible
EIGHTH GRADE				
CHILDID	First Name	Last Name	Date of Birth	Assessment Type
10045	Nancy	Gilbert	5/5/2006	State Proxy
10047	Paul	Johnson	7/4/2005	Accessible
10048	Howard	Richardson	6/6/2004	Accessible
10051	Jane	Martin	5/2/2004	Accessible
10054	Kedy	Nicholas	2/14/2003	Accessible
10058	John	Jo	6/7/2007	Accessible
10059	Gilbert	Queen	2/2/2003	Accessible
GENERAL EDUCATION STUDENTS				
FOURTH GRADE				
CHILDID	First Name	Last Name	Date of Birth	Assessment Type
10132	Patricia	Nicholas	12/12/2004	State Proxy
10133	Pat	Travis	12/1/2005	State Proxy
EIGHTH GRADE				
CHILDID	First Name	Last Name	Date of Birth	Assessment Type
10312	Rick	Shapiro	3/3/2006	Accessible
10313	Cathy	Howard	3/3/2007	State Proxy

1.4 Data Collection Components

You will be responsible for some, but not all of the data that will be collected for PARA. It is important for you to understand all of the data collection activities in case school staff asks you about them. The remainder of this manual will provide you with detailed information about the specific instruments and procedures that you will be using during the spring 2010 data collection.

1.4.1 *PARA Accessible Assessment and Proxy State Assessment*

The PARA study design is based on the random assignment of students to the treatment conditions (accessible and proxy assessment) at the school level. Both assessments will include informational-expository texts and fiction-literary texts and generic motivation questions following items for each passage. The assessments are both designed to conform to the National Assessment of Educational Progress (NAEP) 2009 Reading Framework guidelines. Both versions will be printed in booklets, and both will consist of 4 reading passages followed by 7 multiple choice items each. The test should take around 45 to 60 minutes for students to complete.

The experimental and control conditions vary in the following ways:

- a. **PARA Accessible Assessment (experimental)** –
Students in the experimental group will read accessible passages rated as “highly interesting,” and many of these include color pictures. These passages will also be divided or segmented with 2-3 stopping points where students pause, answer a few questions, and then read the rest of the passage.
- b. **Proxy State Assessment (control)** –
Students in the control group will read non-accessible passages rated as “less interesting” and these passages will not be segmented. If pictures are included they will be reproduced in black and white.

Students in grades 4 and 8 will be asked to read 4 passages (2 literary-fiction and 2 informational-exposition). The passages vary at each grade level in terms of difficulty, topics, and length. The passages will be 500-700 words in length for 4th graders and 700-1,000 words in length for 8th graders; these lengths match typical state assessments.

After students read portions/segments of passages (experimental group) or the entire passage (control group) they will answer 7 multiple choice questions (locate and recall, integrate and interpret, critique and evaluate) based on “cognitive targets identified on the new 2009 Reading Framework for the National Assessment of Educational Progress, and including questions about vocabulary in the passages.

After answering the 7 multiple choice comprehension questions, students will rate each passage based on the following: a) how interesting the passage was, b) how challenging the passage was to read, c) whether readers liked the passage because it was a story

(narrative texts) or they learned new facts (expository texts), and if they knew something about the passage (a rating of prior knowledge after reading the passage).

Assessing how students rate the passages in terms of interest and challenge, as well as how they see themselves as readers, helps educators understand the readers' comprehension abilities, and will promote the development of more valid reading assessments.

The passages are printed in a paper booklet, with a two-column format, considerable white space, and clearly specified answer choices. Students will mark answers in the test booklet.

1.4.2 Motivations for Reading Questionnaire (MRQ)

A survey will be given to all students to provide information about their perceptions of "self as reader." The MRQ will be administered to groups of students via a computer-based platform using an online survey tool. The MRQ is an explicit, general motivation assessment that measures aspects of student's motivation for reading such as self-efficacy, intrinsic and extrinsic motivation and reading attitudes that are correlated with large-scale assessment.

Therefore, researchers indicate that it is useful to measure students' general motivation and compare it to their performance on a large-scale reading assessment. The online survey should take between 15 to 30 minutes for students to complete.

The school contact will help students gain access to computers and the internet to take the MRQ.

1.5 PARA Organization and Staffing

This field trial study is part of a large, multi-year federally funded Partnership for Accessible Reading Assessment (PARA). PARA is a consortium of the National Center on Educational Outcomes (NCEO), the Department of Curriculum and Instruction at the University of Minnesota, the National Center on the Research on Evaluation, Standards, and Student Testing (CRESST), and Westat, Inc. The University of Minnesota (UMN) is the prime contractor and has designated Westat as the subcontractor.

Project Director: Marsha Brauen

Principal Investigators: Martha Thurlow and Deborah Dillon

Project Manager: Juanita Lucas-McLean

Field Director: Sylvia Segovia

Field Managers: Priscilla Ek and Cecilia Avison

Field supervisors for the assessment and MRQ are Robert Scheid and Debbie Singleton. Field supervisors will be the primary contact for coordinating with contacts from schools

for dates of testing and for receiving completed test materials. The Westat field staff report directly to the field supervisors.

1.6 PARA Contact Information

Field Supervisors:

Robert Scheid 1-888-229-4729
Cell: 925-200-6190
77 Chaucer Drive
Pleasant Hill, CA 94523
RobertScheid2@fosmail.westat.com

Michelle Singleton (Debbie) 1-888-231-1349
Cell: 202-387-1809
1724 First Street, NW
Washington, DC 20001
DebbieSingleton@fosmail.westat.com

Field Managers: 1-800-937-8281
Priscilla Ek 1-800-937-8281 Ext. 4872
RW 2632
1600 Research Blvd
Rockville, MD 20850

Cecilia Avison 1-800-937-8281 Ext. 3611
RW 2681
1600 Research Blvd
Rockville, MD 20850

2. DATA COLLECTION PROCEDURES

2.1 Roles of the Research Team

2.1.1 *Field Supervisor*

The PARA Field Supervisor is responsible for ensuring that the data collection process at each school within your assignment is completed accurately and within the time period allocated. The Field Supervisor is responsible for managing data collection in several districts that have agreed to participate in the PARA study.

Prior to data collection, the field supervisor will:

- Schedule testing with school contacts
- Coordinate travel arrangements for assessors

During data collection, the field supervisor will:

- Conduct conference calls with assessors three to four times per week
- Review/approve timesheets, expense forms and incentive logs

2.1.2 *Field Assessor*

As an assessor for the PARA study you will be responsible for conducting data collection at one or more school districts (and schools) participating in the study. You will be playing a very important role on this study because we are relying upon you to complete and return assessments as part of the data collection process. This means that you will be responsible for completing your assignment accurately and within the time period allocated. In many ways, the success of the study hinges upon your ability to understand the correct procedure for test administration and to return completed assessments back to the home office (Westat) each week. Your Field Supervisor will work closely with you during the data collection period to resolve any data collection challenges that may occur.

To assist you in managing your case assignment, each assessor will report directly to a Field Supervisor who in turn reports to a Westat Field Manager. The Field Managers for this study are Priscilla Ek and Cecilia Avison. The Field Supervisors are: Robert Scheid and Debbie Singleton. You will have multiple weekly scheduled phone calls with your Field Supervisor during which time you will report on the status of each case within your assignment. During the report call, you will:

- Review the status of each case
- Discuss any problems you may be having
- Discuss when you expect to complete your cases
- Receive additional cases, instructions, or clarifications on conducting the assessments
- Discuss/approve production costs (hours and expenses)

In addition to your scheduled calls, your Field Supervisor is available by phone should you have a question or issue that needs their immediate attention.

The Assessor is responsible for the following:

- Preparing assessment booklets and materials
- Reading directions to students
- Administering one-on-one and group assessments in schools (approx 45-60 min)
- Collecting the booklets and other assessment materials.
- Ensuring that quality control is achieved by reviewing the assessment forms
- Having the school contact complete the Incentive Request Form
- Packing and shipping assessment materials back to the home office weekly (or biweekly if necessary) using FedEx
- FedExing Timesheets, expense forms, and incentive logs to supervisor weekly (or biweekly if necessary)
- Calling your Field Supervisor at a minimum of three times a week for scheduled report calls to discuss current status of cases assigned to you and to receive new assignments.

2.1.3 School Contact

Each school will designate a staff person who will serve as the study's "School Contact." This person will be the liaison between the school and study staff and will be compensated by the study when data collection has finished.

The School Contact will be responsible for the following:

- Completing the roster for students with disabilities and general education students for both 4th and 8th grades.
- Sending home parent letters with students selected to participate in this study.
- Coordinating with Field Supervisors to schedule a testing date or dates, if necessary.
- Being the contact person on the day(s) of testing. It is important for a testing location to be determined beforehand. The School Contact person will be responsible for pulling the children out of their classes and bringing them to the testing location.
- Arranging for students to gain access to computers in order to take the online motivation survey.

2.2 Handling Your Caseload

This section is designed to familiarize you with the steps involved in scheduling your site visits, using forms and reports, and following procedures during data collection. Because it is critical that we assess a high percentage of children in a short amount of time, we will monitor assessor productivity closely. We will reassign cases if an Assessor exhibits unexplained slow progress and low productivity.

2.2.1 Preparing Your Site Visit

You will receive your assigned cases and case materials from your Field Supervisor. Before data collection begins, your Field Supervisor (referred to as FS) will call the school to obtain all necessary information regarding your visit such as location, School Contact information, and number of kids to be tested. Before going to the school, you will call the School Contact to confirm the date and time of your arrival at the site. **It is important the school is aware of your arrival and is prepared to receive you.**

Before your visit, you should:

- Prepare testing materials by checking Child ID number on assessment booklets and scheduling sheets are the same
- Review student roster for any special accommodations and bring appropriate materials if necessary
- Plan route to school

2.2.2 Administering the Assessment

As the assessor in the field you are the person who has first-hand knowledge about what is going on in the sites where you are working. You will discuss the progress of each site with the Field Supervisor during your report calls. Should any difficult situations occur in a site, the FS will be responsible for assessing the situation and working with their Field Manager to reach a resolution. Knowing the name of the Westat Field Manager is important in the event that you have an emergency situation and are unable to speak with your FS after a reasonable time. Debbie Singleton's Field Manager is Cecilia Avison and Robert Scheid's Field Manager is Priscilla Ek. You can reach Field Managers at the following toll-free numbers **if you have an emergency situation and are unable to reach your FS:**

Cecilia Avison- 1-800-937-8281 ext 3611

Priscilla Ek- 1-800-937-8281 ext 4872

Upon arriving at the school:

- Sign in at front desk asking for school contact
- Locate testing classroom/space and get assessment materials ready
- Introduce study to student
- Obtain informed consent
- Conduct/Monitor Assessment
- Distribute incentive payment
- Have the school contact complete the Incentive Request Form
- Review test for accuracy and completeness
- If necessary, complete a NonResponse Report Form (NRR)
- In the rare event that the school provides their own assessor, you will still be present during the testing, but be sure to collect the completed test materials to send to the home office

2.3 Returning Assessments to the Home Office (Westat)

On a weekly basis, you should mail back to the home office the following (via Fedex):

- Your completed assessment booklets,
- Any NRR forms,
- Informed consent forms

On a weekly basis, you should send the following to your Field Supervisor (in a separate mailing).

- Timesheets
- Expense reports
- Incentive Log/Receipts,
- School Contact Incentive Request Forms

3. ADMINISTERING THE ASSESSMENT

3.1 Overview of the PARA Assessment

Students with disabilities and general education students will be randomly assigned to take one of two tests:

- **PARA Accessible Assessment or**
- **Proxy State Assessment**

Test booklets for each type of assessment will have different color covers. Information as to which test each student will be taking will be listed on the Roster. You do not need to worry about the differences between the two tests. Test booklets will come pre-labeled with the child's information on it (child full name, grade, school, test type), so the only thing you need to do is make sure each child gets the correct booklet. It is very important for you to write the date the child took the test as well as your name on the cover of the test booklet.

During the test (both types) students will be asked to read 4 passages (which vary in terms of difficulty, topics, and length), answer 7 multiple choice questions, and rate each passage. Students will mark their answers directly in the test booklet. The test should take around 45 to 60 minutes for students to complete. Students needing accommodations may take longer (depending on the accommodation). More information on accommodations can be found in Chapter 4.

3.2 Student Consent to Participate

Verbal Consent (4th graders)

The student assent form will be given to each fourth grader to read and they can participate only if they affirm verbally to their test administrator that they understand the purpose of the assessment.

Signed Consent (8th graders)

Eighth graders can participate only if they read and sign the consent form confirming they understand the purpose of the assessment.

3.3 Incentives

Student Incentives

Students will receive an incentive gift card to a store near their area for their participation in this study. 4th grade students will receive a \$15 gift card and 8th graders will receive a \$25 gift card. The higher amount reflects the challenges of recruiting an older population.

Assessors must take care to keep an accurate account of the distribution of gift cards. The best way to do so is to maintain a “weekly incentive log” and by keeping receipts. More information on the Incentive Log can be found in Chapter 5. The assessor will be held responsible for lost cards.

School Contact Incentives

The designated school contact person will be reimbursed for assisting with the coordination of data collection. This person will complete a “School Contact Check Request Form” provided by an assessor with their name, mailing address, and social security number. The incentive will vary from \$200 to \$500, depending on the number of students sampled to participate at their school. Check payments will be sent when data collection for the entire state has been completed.

3.4 Checklist of Items needed for the Assessment

Please make sure you have all of the following when going to a school to administer assessments:

- Roster
- Test booklets for the children to be tested
- **Directions for Students** (Exhibit 3-1)
- Sharpened pencils (collect after each session)
- Copies of “**Sample Test Items**” (Exhibit 3-2) (sample motivation test items) (collect after each session)
- Student Consent forms
 - **4th Grade Assent Form** (Exhibit 3-3) (collect after each session)
 - **8th Grade Consent Form** (Exhibit 3-4)

- Gift cards
 - \$15 for 4th graders
 - \$25 for 8th graders
 - **Gift card receipt** (Exhibit 3-5)
- **School Contact Check Request Form** (Exhibit 3-6)

3.5 Administering the Assessment

These are the following steps to administer the assessment:

1. Hand out consent forms to 4th and 8th graders
2. Read “CONSENT” section from “Directions for Students” handout
3. Collect back the consent forms
 - a. 8th graders need to have signed the form
4. Hand out the “Sample Items” Pages
5. Read “SAMPLE ITEMS” section from “Directions for Students” handout
6. Collect back the Sample Items Pages
7. Hand out test booklets and pencils
8. Read “TEST DIRECTIONS” section from “Directions for Students” handout
9. Begin testing
10. When testing is complete, children should turn in their tests (and pencils)
11. Give child a gift card and have them sign the receipt (give them a copy of the receipt and make sure to keep a copy for yourself)
 - a. 4th graders: \$15
 - b. 8th graders: \$25
12. Please staple the consent form to the inside front cover of the booklet (this can be done after you have left the school)
13. Have the School Contact fill out the Check Request Form

Frequently Asked Questions relating to General testing can be found in the Appendix.

Exhibit 3-1. Directions for Students



Partnership for Accessible Reading Assessment

DIRECTIONS FOR STUDENTS

CONSENT:

Good morning/Good afternoon students. I am here as part of a research study that looks at ways to make reading tests better for students. Each year you take a reading test that all of the other students in grade 4/8 in this state take. Sometimes those tests are hard for kids. A group called the Partnership for Accessible Reading Assessments would like some information from kids about ways these tests could be done differently. We are studying what would help some students "show what they know" more easily.

Today I am going to be asking you to work on some reading passages like the ones on state reading tests. These are not real tests-the score will not count for anything-but please read the passages carefully and do your best on the questions that follow each reading passage. When you complete the test you will receive a gift card.

I have handed you each a consent form. [If 4th graders] Please read it and when I walk around to pick it up you can tell me if you agree to participate. [If 8th graders] Please read it and sign it and I will walk around and pick them up.

SAMPLE ITEMS:

I have handed you each a piece of paper that shows what some special questions look like that follow the passages and the regular questions. These special questions will have a box around them (point out the box). You will mark your choice for each question by filling in the circle next to the best answer.

Let's look at the first question. It says "How interesting was the passage?" There are four choices for you to choose how interesting you thought the passage was: the passage was very interesting, pretty interesting, kind of interesting, or not at all interesting. When you have made your selection use your pencil to bubble in A, B, C or D.

Now let's look at question number 3. It says "Do you like this passage?" Here you have to decide YES or NO and bubble in the circle. Next it asks WHY you liked the passage, either because it provided you with information or because you knew something about the passage. Here you can select one or both reasons as to why you liked the passage.

Any questions?

DIRECTIONS TO STUDENTS:

The booklets that I have passed out to you will not necessarily have the same passages as the person next to you. You will notice that the booklets are different colors. Because booklets may be different from person to person, your classmates may finish at different times.

You have plenty of time to complete the test. This is not a timed assessment. I do ask that you try to work at a steady pace. Read the directions carefully and when I say you can begin, please start working. If you have any questions about what you need to do please raise your hand.

You can open your test booklet and begin!

Exhibit 3-2. Sample Test Items

We want to know what you thought of the passage you just read. For each item, mark the circle that shows how you felt about the passage.

1. How interesting was the passage?

- A This passage was very interesting to read.
- B This passage was pretty interesting to read.
- C This passage was kind of interesting to read.
- D This passage was not at all interesting to read.

2. How hard was the passage to read?

- A This passage was very challenging to read.
- B This passage was pretty challenging to read.
- C This passage was kind of challenging to read.
- D This passage was not at all challenging to read.

3. Did you like this passage?

- Yes No

If you liked this passage, why did you like it? (You can select one or both of the answers below.)

I liked it because it provided me with information

- Yes No

I liked it because I knew something about the passage

- Yes No

Exhibit 3-3. 4th Grade Assent Form



Partnership for Accessible Reading Assessment

Dear Student,

You are invited to be in a research study. You are a student in 4th grade and you will take a state reading test this year. We have some ideas about how to change the tests to make them show what kids know how to do. We made a sample test to try out some of these ideas with students. We think you can give us good information about making tests better. The people in charge of this study are from the University of Minnesota. We are from the Partnership for Accessible Reading Assessment project (PARA) at a company called Westat.

If you agree to be in this study there are two parts:

Part 1: You will fill out a survey about your reading habits on the computer.

Part 2: You will also read 4 passages and answer some questions after each one.

The study takes about 2 or 2 ½ hours. It may take place over two days. Your school will decide the best days and times. You may miss some class time.

If you finish all of the activities in this study you will receive a \$15 gift card to a store in your area.

We will write some reports about the things we learn in this study. We want to write about how groups of students do on the reading passages and the questions. When we write we will not use your name. We will not use any information that will identify you. We will not identify your school.

Please ask any questions that you have about this study. You can ask your teacher or you can ask the researchers.

If you want to participate, just answer “Yes” when asked if you understand what you read on this paper. Remember, being in this study is up to you. You can change your mind at any time and decide you no longer want to participate. Being in this study is your choice and no one will be mad at you if you don’t want to do it.

Exhibit 3-4. 8th Grade Consent Form



Partnership for Accessible Reading Assessment

Dear Student,

You are invited to be in a research study that looks at ways to make state reading tests better at showing what kids know. We are inviting you because you are in 8th grade. Most students take state reading tests in 8th grade. We think you can give us good information. Please read this paper and ask questions before you agree to be in the study. The people in charge of this study are from the University of Minnesota. We are from the Partnership for Accessible Reading Assessment project (PARA) at a research company called Westat.

Background Information:

We are studying state reading tests like the ones that many students take each year. Some students read better than the state reading test scores show. We have some ideas about how to change the tests to make them show what kids know how to do. In this study, we will use a sample reading test that is designed to try out some of these ideas. We would like to know what students think about it.

Procedures:

If you agree to be in this study there are two parts:

Part 1: You will fill out a questionnaire about reading on the computer.

Part 2: You will also read 4 passages and complete some multiple choice questions about the passages. If you use accommodations on tests you can use those accommodations for the reading activities.

The study will most likely happen across two days. It should take about 2 or 2 1/2 hours total to complete. Your school will decide the exact time of the study.

Compensation:

If you finish all of the activities in this study you will receive a \$25 gift card to a store in your area.

Exhibit 3-4. 8th Grade Consent Form, (continued)

Risks and Benefits of Being in the Study:

You may miss some class time if you participate in this study. However, you will get practice in reading passages that are like the ones on state tests and answering the multiple choice questions that go with them. This practice can help you when you take the actual state reading test.

We are using reading activities that look like reading tests you take each year. The scores will not count. Your grades will not be affected by taking the test. We will not give your scores to anyone. This information will only be used to find out if changes in the test make a difference in the way you read and answer questions.

Confidentiality

We will write some reports about the things we learn in this study. We want to write about how groups of students do on the reading passages and the questions. When we write we will not use your name. We will not use any information that will identify you. We will not identify your school.

We will keep the information we collect in a locked file cabinet or on a password protected computer. Only the people who work on this project will be able to see the information that you give us. All your responses will be confidential.

Participation is Voluntary

If you want to participate, please sign at the bottom of this page. No one will be upset if you decide that you do not want to be in this study. If you decide to be in the study, you may stop at any time.

Contacts and Questions:

The researchers conducting this study are from the University of Minnesota and Westat. You may ask any questions you have right now. If you have questions later, please contact Bonnie Ho at Westat.

Phone:

Email: parastudy@westat.com

Exhibit 3-4. 8th Grade Consent Form, (continued)

Statement of Consent:

I have read the above information or an adult has read it to me. I have asked questions, and they were answered. I agree to participate in this study.

Student Name (Print): _____

Student Signature: _____

Date: _____

Assessor Name (Print): _____

The researchers will give you a copy of this form to keep.

Exhibit 3-5. Gift Card Receipt

PARA Partnership for Accessible Reading Assessment
Incentive Receipt - Data Collection 2010

Amount of Gift Card Received for Child Assessment

Date: _____ \$ _____

Assessor Name: _____ Received by: _____ Print Name

Child ID: _____ Signature _____

White - Site Coordinator/Hersey Office Yellow - Intervention/Assessor Pink - Respondent

Exhibit 3-6. School Contact Check Request Form

 Partnership for Accessible Reading Assessment	
Check Request Form, Data Collection 2010	
Please send the incentive check for assisting with the coordination of data collection to:	
Date _____	
Assessor Name _____	School Contact Name (check will be made out to this person) _____
School Name _____	Street Address _____
	City _____ State _____ ZIP _____
	Social Security Number _____

4. TESTING ACCOMMODATIONS

4.1 Accommodations for the PARA Assessment

The assessments that students are taking are designed to be more accessible than typical state assessments. In addition, these assessments are shorter than typical state assessments. These characteristics may reduce the need for accommodations typically provided for the state assessments.

If a student will need an accommodation when taking the PARA Assessment, it will be noted on the roster. Your field supervisor will have discussed this beforehand with the school contact, when scheduling for testing took place. The school contact will be responsible for making any special arrangements to allow for any special accommodations needed.

These are the following accommodations that a student taking the PARA assessment could need:

1. LP: Large print: Blank booklets with large print will be provided for you. You should write all the information included on the label of the original booklet onto the large print booklet and send back along with the original booklet.
2. AD: Read aloud directions: Assessors will have to read the directions aloud to the students. This may require one-on-one testing. The student(s) needing this accommodation will probably need a separate space so as not to interfere with other students. The school contact should have this arranged ahead of time.
3. SD: Sign interpret directions: In situations such as this we will use the sign interpreters the school uses. The school contact should have this arranged ahead of time. Make sure the sign interpreter signs the confidentiality form.
4. RD: Repeat/Paraphrase Directions: See instructions for #2. The same applies. Paraphrasing means restating using other words to help students better understand the directions.
5. ME: Magnification equipment: We will use the magnification equipment the school uses. The school contact should have this arranged ahead of time.
6. PS: Proctor/scribe: We will use the proctor/scribe the school uses. The school contact should have this arranged ahead of time.
7. ET: Extended time: We can allow up to 2 hours for this test

8. WB: With breaks: Students can take breaks throughout the test. A five minute break every 30 minutes should suffice.
9. IT: Individual testing: The student or students needing this accommodation will probably need a separate space for one-on-one testing so as not to interfere with other students. The school contact should have this arranged ahead of time.
10. There is a possibility that students may require “other” accommodations than those listed above. The school contact will be responsible for making any special arrangements to allow for any special accommodations needed.

Please make any notes about accommodations used in the inside back cover of the test booklet, when applicable.

Frequently Asked Questions relating to testing with Accommodations can be found in the Appendix.

5. OVERVIEW OF DATA COLLECTION MATERIALS

5.1 Using This Manual

This manual briefly explains the background of the study. It also explains the protocol for obtaining parent and student consent and provides an overview of data collection materials and procedures. In addition, it includes instructions, examples, and copies of all related study forms/letters. The manual has been organized to guide you throughout the field period. Assessors and Field Supervisors should use the manual frequently as a reference to answer procedural questions. Keep your copy in a place where it can be easily obtained when needed, such as in your car.

You will receive a variety of materials to assist you in carrying out your work. This chapter provides an overview of those materials and their uses.

5.2 Field Materials

This manual provides examples and descriptions of materials to use to enlist cooperation of children. It also includes examples and descriptions of materials to collect and record data.

The field materials can be grouped into two broad categories: case specific and bulk. Materials that are case specific are prepared for a specific child. These are labeled assessment booklets and rosters.

Bulk supplies are materials and items that are not case or project specific. They include the Time and Expense Report (T&E) and Trip Expense Report (TER) forms, used by all field staff, as well as office items such as pens, pencils, highlighters, regular envelopes, FedEx labels (addressed to Therese Koraganie or to Field Supervisor), and boxes. Bulk supplies also include Parent Consent Letters (in case you need them unexpectedly), Assent forms for 4th grade students, Consent forms for 8th grade students, test sample pages, test directions for students, Transmittal forms, Incentive Receipts, Incentive Logs, Check Request forms, NonResponse Report Forms (NRR), Scheduling Matrix forms, blank large print assessments, and Data Collector Code of Conduct and Assurance of Confidentiality forms.

It is your responsibility to make sure that you always have a sufficient supply of all materials necessary to complete your assignments. Each day, before starting work, you should check over your materials. If you do not ensure that you have all needed materials, you could waste valuable time returning to get them at your residence or hotel, which could also delay your team, interfere with study activities, cause cancelled appointments, and increase the risk of losing data.

For your safety, and to secure any data collection materials in your car, make sure that your car is locked. Lock your car whenever you are in the field, whether you are in your car or not. All data collection materials must be secured at all times. You will be responsible for accounting for all materials that are entrusted to you, especially completed assessment booklets.

The contents of each shipment of supplies should always be checked against the transmittal list enclosed. If you are missing any items, or have a shortage, you need to contact your Field Supervisor, and he/she will contact Westat immediately.

5.3 Forms and Materials Used in the Study

This section lists materials that are to be used in the study. It is divided into two sections: Collecting the Data, and Wrapping Up. Additional details on specifically how to use these materials are contained in other chapters throughout this manual.

5.3.1 Collecting the Data

Field Supervisors Scheduling Matrix (Exhibit 5-1)—This form allows the Field Supervisor to record the weekly schedule for up to six Assessors.

Weekly Individual Scheduler Form (Exhibit 5-2)—This form allows the Assessor or the Field Supervisor to plan a 1-week schedule for one person.

Exhibit 5-1. Field Supervisors Scheduling Matrix

PARA Partnership for Accessible Reading Assessment
 2010 Data Collection

Field Supervisor Weekly Scheduling Matrix - Week of Monday, _____, 2010

Mon	Tue	Wed	Thu	Fri	Sat	Sun
Name: AM						
PM						
Name: AM						
PM						
Name: AM						
PM						
Name: AM						
PM						

Exhibit 5-2. Weekly Individual Scheduler Form


Partnership for Accessible Reading Assessment

Individual Weekly Scheduling Matrix
 Data Collection 2010
 Week of Monday, _____, 2010

	Monday	Tuesday	Wednesday	Thursday	Friday	Saturday	Sunday
8 AM							
9 AM							
10 AM							
11 AM							
12 PM							
1 PM							
2 PM							

Consent Letters for Parents or Primary Caregivers—Schools will send to the parents/primary caregivers a passive informed consent letter, asking them to call the study’s toll-free numbers only if they DO not agree for their child to participate.

Informed Consent Forms for 8th Grade Students—You will need to make sure that 8th grade participating students sign a Consent form before the assessment.

Informed Assent Forms for 4th Grade Students—You will need to make sure that 4th grade participating students verbally agree before the assessment.

The Child Assessment Booklet—You will be provided with labeled child assessments and also large print versions (not labeled) in case they are needed.

Sample Test Items—You will hand out the “Sample Test Items” page, read from your directions handout, and collect the sample pages back before starting the assessment.

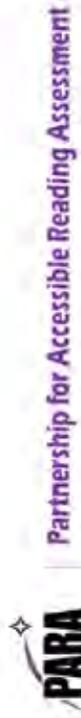
Test Directions for Students—Follow the instructions verbatim and precisely given to you in the handout to administer the assessment.

Gift Card Incentives—You will be provided with gift cards of \$15.00 for the 4th graders and \$25.00 for 8th graders. Please distribute the gift cards only after the completion of the assessment.

Gift Card Incentive Receipts—Following completion of Child Assessment, give the child the corresponding gift card, fill in the child’s complete name, ID number, your name, the date, and have the child sign the form. The Receipt is printed on 3-ply paper so make sure to give the child a copy and keep a copy for yourself before sending to your supervisor.

Incentive Log (Exhibit 5-3)—The Incentive Log is your most detailed record of all of the gift cards given to students. It will be used to record information on the amount of each card. The Incentive Log is printed on 3-ply paper and should be finalized on a weekly basis to document the total number of incentives cards that have been given out. By maintaining accurate records you will not be charged personally for gift cards that cannot be accounted for.

Exhibit 5-3. Incentive Log



WEEKLY GIFT CARD INCENTIVE LOG
Data Collection 2010

Assessor Name: _____

	Date	School	Child ID	\$15.00 card	\$25.00 card
1					
2					
3					
4					
5					
6					
7					
8					
9					
10					
11					
12					
13					
14					
15					
16					
17					
18					
19					
20					
21					
22					
23					
				TOTAL	

White - Home Office Copy Yellow - Site Coordinator Copy Pink - Assessor Copy

School Contact Check Request Form—Have the school contact complete this form at the end of data collection activities. Payment will be sent after data collection has concluded.

Nonresponse Report Form (Exhibit 5-4)—This form is filled out if a complete assessment cannot be obtained or if the case cannot be resolved. The reasons for nonresponse must be documented. The two types of nonresponse are:

- Refusal/Break-off (the child refuses or breaks off during any part of the Child Assessment);
- Child is absent from the school on the day of the assessment.

Frequently Asked Question about General Testing Issues (Appendix A)—This document outlines testing situations that may arise and how to handle them. If a situation arises that is not covered by these FAQs contact your supervisor.

Frequently Asked Questions for Special Accommodations (Appendix B)—This document outlines specific situations that may arise in relation to accommodations for students with disabilities. If a situation arises that is not covered by these FAQs contact your supervisor.

5.3.2 Wrapping Up

Transmittal Form (Exhibit 5-5)—This is a 2-ply form that allows you to check off the materials that you are returning so that upon receipt, a Westat staff member or the Field Supervisor (depending who is the recipient) can make sure that all of the materials that you have sent have been received. The white copy should be enclosed with the package, the yellow copy is for the sender (i.e., Field Supervisor or Assessor) to keep. You must always include a Transmittal Form in the package when sending instruments and/or forms. The Transmittal Form that you keep is a valuable tool in tracking down a package should it get lost. It is also your record that you sent the package and the date you shipped it.

Exhibit 5-4. Nonresponse Report Form



Partnership for Accessible Reading Assessment

NonResponse Report Form (NRR)
Data Collection 2010

Date: _____

School Name: _____

Child ID: _____

Assessor Name: _____

Child Name: _____

A. WHICH INSTRUMENT?

Assessment _____

MIRQ _____

B. WHAT TYPE OF NON-RESPONSE?

Refusal → → → → Go to #1

Break-off → → → → Go to #1

Absent → → → → Go to #5

Other → → → → Go to #6

REFUSAL:

1. At what point did the refusal/break-off occur? _____

2. Reason for Refusal/break-off (Check all that apply)

Confidentiality

Not interested

Other (specify) _____

Negative reaction to test

No reason given

3. Stated reason for Refusal (record verbatim) _____

4. Nonverbal reasons you feel caused the Refusal _____

OTHER NONRESPONSE:

5. Note situation and whether you feel assessment can be completed _____

Disabled Language Problem Unintelligible

White: Home Office Copy

Yellow: Site Coordinator Copy

Pink: Assessor Copy

Exhibit 5-5. Transmittal Form



From (sender): _____ Date: _____
 Sent to (receiver): _____
 Page: _____ of _____ FedEx Airbill # (found at top of label) : _____

Check each item enclosed and total:

	Child ID/Name	IC	AS	IR	IRF	NRR	TOTAL
1							
2							
3							
4							
5							
6							
7							
8							
9							
10							
11							
12							
13							
14							
15							
	Total						

Other materials included in shipment (i.e., T&E, TER, etc.): _____

Key	
IC = Informed Consent	IRF = Incentive Request Form
AS = Assessment Booklet	NRR = Non-Response Report Form
IR = Incentive receipt	

White Copy: Receiver

Yellow Copy: Sender

Mailing Supplies—You will be given boxes, FedEx labels, and other supplies as needed to return material to the Field Supervisor and/or Westat.

Time and Expense Forms (T&E)—In order to receive a paycheck, you must fill out a T&E each week. It is used to record the hours you have worked, your productivity, and the mileage incurred on local travel. In addition, you will complete a weekly production report and attach it to your T&E.

Trip Expense Report (TER) Form—To receive payment for expenses incurred during travel outside a local area requiring an overnight stay, you must fill out a TER and return it with your timesheet for that week. It includes a place to record hotel expenses, meals, and mileage while on an overnight trip. All travel expenses must have prior approval from the Field Supervisor and the Field Manager.

6. USING ADMINISTRATIVE FORMS AND TRAVEL PROCEDURES

We have established administrative procedures to help ensure that data collection activities and travel reimbursement run smoothly for Assessors. This chapter will explain in detail the procedures that you and your Field Supervisor must follow throughout the data collection process. *All forms discussed in this chapter must be sent to your Field Supervisor via FedEx separately from your shipments of case materials to the home office.*

6.1 Administrative Duties

6.1.1 Reporting Time and Expenses and Travel Expenses

The Time and Expense Report (T&E) and Travel Expense Report (TER) are to be used by Assessors to record time worked and any expenses incurred during travel. Field Supervisors review and approve all T&Es and TERs before sending them to Westat.

It is the Assessor's responsibility to make sure:

- The hours and expenses are recorded completely and accurately;
- All receipts are stapled to each T&E and TER; and
- They submit T&Es and TERs each week on Wednesday or Thursday to their Site Coordinator.

A ballpoint pen should be used when filling out a T&E or TER; a pencil should never be used. If a mistake has been made, it should be crossed out with a pen; **do not use whiteout**. Put your initials beside the mistake, and enter the correct information. All changes on the T&E and TER must be initialed so that it is clear who made the changes.

Westat's accounting department issues paychecks, travel advances, and travel reimbursements via direct deposit only once a week on Fridays. If a T&E or TER is not received by the accounting department by Tuesday at 1:00 pm EST, it will be held for payment for the next week's payroll.

6.1.2 Field Time and Expense Reporting

Be sure that your **Time and Expense (T&E) Report Forms** (see Exhibit 6-1) is submitted on a weekly basis. **Westat will not pay for hours worked that are submitted more than 30 days after the work is performed or for expenses that are submitted more than 45 days after the costs are incurred.**

In addition to recording miles traveled and hours worked, the T&E also provides space for recording other allowable, out-of-pocket expenses that may be incurred while working for Westat, such as travel by mass transportation, tolls, parking, telephone calls, and postage.

Expenses that are **not allowed** include car insurance, traffic tickets, towing, taxi fares, car repairs, car upkeep, or cellular phones. Calling cards used for long distance calls are acceptable, and copies of bills should be submitted for reimbursement of Westat business calls. However, pre-paid phone cards are not permitted since calls are not itemized. Westat can only reimburse itemized work-related phone calls.

- Supervisor copy (Pink); and the
 - Employee copy (Gold).
- Each copy is a different color and is reproduced on pressure-sensitive paper that requires no carbon paper. For this reason it is important that you be careful and only write on one set of sheets at a time. Use a black or blue ballpoint pen and press firmly. Keep the last copy for your records. If you make a mistake, cross out the mistake with your pen, do not use whiteout. Put your initials next to the mistake and enter the correct information. All changes to your T&E must be initialed so that it is clear who made the changes. This is for your safety and benefit.
 - Each pay period begins on Monday and ends the following Sunday at close of business. As part of your daily routine, you should record hours worked, miles traveled (if applicable) and any related expenses for that day. You must record all hours worked in the week they occur: never hold over hours worked from one week to the next. You must submit your T&E reports in accordance with this schedule; that is, each report must show the Monday starting date indicated on the report, and must include only time worked between that date and the field work week ending date. After finishing your work on Sunday, all columns should be tabulated, and the summary section on hours, mileage, and expenses completed.
 - It is Westat policy that all overtime be approved in advance. If you anticipate that your assignment may require working more than 40 hours in a week (or more than 8 hours a day in some states), you must discuss your schedule with your supervisor **before** working the extra hours. However, there will be times when unanticipated events result in additional hours worked without prior approval. Notify your supervisor as soon as possible if you incur unapproved overtime hours. Westat will pay for all hours worked but a reasonable explanation must be provided or you may be subject to disciplinary action.
 - The first four copies of the T&E should be mailed to your supervisor Sunday or Monday. Westat's accounting system issues time and expense checks only once a week on Friday. In order for payroll to issue a payroll check on Friday, Supervisors need to send T&E forms to the home office by no later than Thursday, so that home office managers have time to review the timesheets, ask questions to the Field Supervisors, if needed, and be able to submit them to the payroll office by 1 p.m. the following Tuesday.. Therefore, it will take approximately 2 weeks from the time you submit your T&E to your supervisor, until you receive your check. If you have not received your check in 3 weeks let your supervisor know. Interviewer checks are mailed every Friday from Westat's main office in Rockville, Maryland. You will receive two separate checks or notification of deposits: one for hours worked and one for mileage and expenses. The graphic below illustrates this process.



- **Submit your T&E on a weekly basis.** Westat policy requires that time and expense reports be submitted no later than 30 days after the work is performed or 45 days after the expenses are incurred.
- **Never record names or notes anywhere on a time sheet.** If you have a comment you would like to send to your supervisor or the home office, paper clip it to your timesheet.

Expenses

The T&E form is used for submitting allowable out-of-pocket work expenses. For any expense over \$1, a receipt is required. All receipts should be stapled to the back of the right hand side edge of the **green copy** of the time sheet. **Allowable expenses** are:

- **Mileage.** Mileage is paid at the rate of 50 cents for each mile driven while on Westat business.
- **Telephone Calls.** If you do make some toll or long distance project related calls, total the amount of the expenses and enter the amount on the T&E along with a copy of your telephone bill. Your statement should have all calls highlighted for which you are requesting reimbursement. If a pay phone was used, indicate that on

your T&E as “Pay Phone” and no receipt is needed. Pre-paid phone cards are not permitted since calls are not itemized.

- **Postage.** Expenses for purchasing stamps for mailings to your supervisor or Westat are reimbursable. Receipts are easily obtained at the post office. However, in this project it is not likely that you will need to pay for postage since we are using FedEx as the main means of correspondence.
- **Other.** Miscellaneous expenses such as:
 - **Tolls.** Tolls or road fees on bridges, expressways, tunnels, etc. are acceptable and should be shown on the T&E report with attached original receipts.
 - **Parking.** In congested urban areas where free parking is not available, Westat will pay for parking. To obtain reimbursements for parking, enter charges on your T&E and attach the original receipt. If metered parking, indicate on your T&E as “Parking Meter” and no receipt is needed.
 - **Travel by Mass Transportation.** Fares for bus, subway, and other forms of mass transportation used on official business are allowable expenses. Indicate the number of trips on a single day (original receipts should be attached).
 - **Faxes.** Expenses for sending faxes to your supervisor or the project field room are reimbursable. Submit your receipts with your T&E.

Be sure that you have not submitted any **unallowable expenses** such as:

- car insurance,
- traffic tickets,
- towing,
- taxi fares (unless supervisor approved),
- car repairs or car upkeep,
- meals (unless supervisor approved),
- pager fees, and
- cellular phones.

Completing the Field Time and Expense Report

To complete the **Non-Exempt Fieldworker I Time and Expense Report**, follow the specific instructions on the following pages. For instructional purposes, the report will be divided into the following sections:

TOP WINS#, Name, City/State, Period Ending Date

MIDDLE LEFT Time and Activities–Field work (FW) and local travel (T), and Production Report

MIDDLE (Project Name, Project Charge Code, Type of rate)

MIDDLE RIGHT (Summary of expenses by project code, miles driven, and Production Report)

BOTTOM LEFT (Adjustment(s) to Pay)

BOTTOM RIGHT (Home Office Review, Certification and Approval)

The image shows a 'NON-EXEMPT* FIELDWORKER I TIME AND EXPENSE REPORT' form from WESTAT. The form is divided into several sections, each labeled with a large black box containing white text: 'TOP', 'MIDDLE LEFT', 'MIDDLE', 'MIDDLE RIGHT', 'BOTTOM LEFT', and 'BOTTOM RIGHT'. The form includes various tables, checkboxes, and input fields for recording work hours, expenses, and production data. The 'TOP' section includes fields for WINS#, Name, City/State, and Period Ending Date. The 'MIDDLE LEFT' section is a grid for recording time and activities. The 'MIDDLE' section is for project details. The 'MIDDLE RIGHT' section is a 'SUMMARY OF EXPENSES BY PROJECT CODE' table. The 'BOTTOM LEFT' section is for 'ADJUSTMENTS TO PAY'. The 'BOTTOM RIGHT' section is for 'CERTIFICATION AND APPROVAL'.

TOP

Be sure to print the following information across the top of your T&E Report:

WINS NUMBER: (Westat Information Number System), be very careful to enter the correct number. If by error you enter a wrong number (say by one digit) and another

Westat staff member has that number, your paycheck may end up being sent to that person.

WESTAT AN EMPLOYEE-OWNED RESEARCH CORPORATION	3550 Research Blvd. Rackville, MD 20850 301-251-1500	NC
WINS # [W][][][] - [][][][]		

NAME: Print your legal name in the correct order: **Last Name, First Name, Middle Initial**

CITY/STATE: Record the city and state of your residence. This is for identification purposes only and is not your mailing address. If you have a new address make sure you notify your supervisor immediately to insure that the proper paperwork is submitted to the home office. The address on the T&E will not update the address your check is mailed to.

NON-EXEMPT* FIELDWORKER | TIME AND EXPENSE RE

Please Print

Name _____
Last First MI

City/State _____
City State

PERIOD END DATE: Enter the Sunday date that ends the T&E week.

PERIOD ENDING DATE: Sunday

[][] - [][] - [][]
Month Day Year

MIDDLE LEFT

Enter Month and Date. Enter time to the nearest .25 hour.								
	MON	TUE	WED	THU	FRI	SAT	SUN	
	M/D							
	/	/	/	/	/	/	/	Total
FW								
T								
FW								
T								
FW								
T								
FW								
T								
HOURS								
MILES								(A)
Production Report								

TIME: Time is to be reported for each day worked by the type of field activity (fieldwork or travel). Record the month and date for each day of the week. Time is recorded to the nearest quarter hour. When recording partial hours, you use decimals instead of minutes (i.e., .25 = 15 minutes, .50 = 30 minutes, .75 = 45 minutes. 1.0 = 1 hour). Do not use fractions.

ACTIVITY: Two types of field activities are reported under each charge code:

- **FW (FIELDWORK):** Include all time spent working on the PARA study:
- **T (TRAVEL):** Include the time spent traveling between your home and your work assignment.

There is room to record Field Work and Travel time for 4 different charge codes on the page. If you work on more than 4 different projects, fill out and attach a second T&E.

TOTAL WORKED HOURS: At the end of each day, total the number of hours worked for the day. At the end of the week enter the total number of hours worked on each project in the “Total” column. Then total the hours worked on each project and enter this at the bottom of the total column in “**TOTAL HOURS.**” The sum of the hours entered in the total column should match the sum of the hours entered in the daily row.

MILES DRIVEN: Record the miles driven each day in the bottom box in the daily column. At the end of the week total the daily miles driven and enter this amount in the right hand box designated “**MILES DRIVEN**” (A).

PRODUCTION REPORT: Each day on the first line write the number of assessments completed. At the end of the week, total the daily assessments and enter this number in the “Total” column.

MIDDLE

This area designates the project(s) to which you are charging your time and expenses. Each T&E allows you to bill time and expenses to four different charge codes. If you bill time or expenses to more than 4 charge codes during a single week, fill out a second T&E and label the bottom of Page 1, 1 of 2, and the bottom of Page 2, 2 of 2. The check boxes in the middle of the page designate your pay rate.

- Project Name: Enter the name of the project: PARA Study;
- Project charge code: Fill in from left to right. **Do not zero fill if your charge code contains less than 10 digits.** Charge code: 8110.01.04.
- Type of Rate: Assessor Circle “P” Special Rate.

**MIDDLE
RIGHT**

TOTAL EXPENSES (per row) including mileage.

NOT FOR OUT-OF-TOWN TRAVEL

SUMMARY OF EXPENSES BY PROJECT CODE							
Respondent Incentives (DRESPI)	Copying (DCOPY)	Telephone (DTELE)	Supplies (DSUPP)	Postage (DPOST)	Other Specify: (FLDOTH)	Mileage Allowance (FLDLOC) per Mile	Total Expenses Including Mileage
						Miles	
						\$	
						Miles	
						Miles	
						Miles	
						Miles	
						Miles (A)	
						\$	

Rate per mile (\$.50)

MILEAGE ALLOWANCE (number of miles and dollar amount at \$.50 per mile)

TOTAL EXPENSES (per column)

SUMMARY OF EXPENSES BY PROJECT CODE: You can bill expenses to four different (project) charge codes on each T&E. Enter allowable expenses for travel, telephone, postage, and other expenses for each project you worked on during the week on the corresponding row. Use one staple to attach all receipts to the **back** of the lower left hand side of the green copy only.

- **Respondent Incentives:** This column should remain blank since you will receive incentives in the form of gift cards from Westat.
- **Copying (DCopy):** List any copying expense you may have incurred.
- **Telephone (DTELE):** List any telephone expense you may have incurred for each charge code.
- **Supplies (DSUPP):** List any pre-approved supply expenses you may have incurred.
- **Postage (DPOST):** List any postage expenses you may have incurred such as purchasing stamps to mail documents to Westat. Be sure to obtain receipts.

- **Mileage Allowance (FLDLOC):** Enter the mileage allowance (50 cents) at the top of the column. Enter the total miles driven. Multiply the total miles driven by 50 cents to arrive at the reimbursement amount. Round all ½ cents up. Total the number of miles and reimbursement amount at the bottom in the “Totals” row. Total miles in the expense total (B) must match total miles in the Miles driven box (B).
- **Other (FLDOTH):** List the name and amount of each miscellaneous expense you may have incurred. If you incurred more than one miscellaneous expense, list and record the amount in the next box in the column. Total the miscellaneous expenses and circle the total for each change code. Be sure to attach receipts.

Allowable “other” expenses include:

- **Tolls.** Tolls or road fees on bridges, expressways, tunnels, etc. are acceptable and should be shown on the T&E report with attached original receipts.
- **Parking.** In congested urban areas where free parking is not available, Westat will pay for parking. To obtain reimbursements for parking, enter charges on your T&E and attach the original receipt. If metered parking, indicate on your T&E as “Parking Meter” and no receipt is needed.
- **Travel by Mass Transportation.** Fares for bus, subway, and other forms of mass transportation used on official business are allowable expenses. Indicate the number of trips on a single day.
- **Totals:** In the row labeled ‘Totals’, sum the total of each column sum of Copying, sum of Telephone, etc.)
- **Total Expenses:** In the final column total all expenses, including mileage, by charge code. Sum all project totals at the bottom of the column.

BOTTOM LEFT

Enter the number of hours to be adjusted

Enter the adjustment rate per hour

Enter the total adjustment

Enter reason for adjustment

Enter charge code for adjustment

ADJUSTMENT TO PAY													Rate	Charge Code
Pay amount	Hour/ unit	\$												
* For Fieldworker Use *														
* For Supervisor Use *														
													GRAND TOTAL	

Page of COPIES: WHITE: PAYROLL GREEN: A/P CANARY: FIELD DIRECTOR GOLD: SUPERVISOR

If you bill time or expenses to more than 4 charge codes during a single week, fill out a second T&E and label the bottom of T&E 1, 1 of 2, and the bottom of page 2, 2 of 2.

ADJUSTMENT TO PAY: In the PARA study there is no adjustment to pay to the special rate.

PAGE of : Number of pages; this will be blank unless you have worked on more than 4 different charge numbers in a given week. If you work on more than four different charge numbers you need to complete two time sheets for the same week. The first time sheet would contain information about the first 4 charge codes and be labeled 1 of 2. The second T&E would contain information about the next 4 charge codes and would be labeled 2 of 2.

BOTTOM RIGHT

Enter the name of the supervisor who authorized overtime hours here.

CERTIFICATION AND APPROVAL

This report will not be processed unless this section is completed. This form is complete and accurate, if overtime hours reported.

Authorized by _____

(Employee's signature) (Date)

For Supervisor(s): I have reviewed and approved this T & E report.

(Supervisor's signature) (Date) _____
(Supervisor's signature) (Date)

HOME OFFICE REVIEW

Adjustment(s) to Pay Overtime, Paid at: Base
 Special
 Supervisor

Signature _____ Date _____

SIGN HERE AND DATE

HOME OFFICE REVIEW: This area is for office use only.

CERTIFICATION & APPROVAL: This box must be completed and signed before the T&E can be approved and paid. Westat policy requires prior authorization for more than 40 hours a week (or 8 hours a day in some states). If you have worked more than 40 hours that week, provide the name of the person who authorized the overtime. If no prior authorization given, a reasonable explanation must be provided to your supervisor. In all cases you must sign and date this box. Your T&E will not be processed unless this section is completed and signed in ink. Use only dark blue or black ink. DO NOT use pencil or red ink.

COPIES: WHITE: PAYROLL GREEN: A/P CANARY: FIELD DIRECTOR PINK: SUPERVISOR GOLD: EMPLOYEE
non-exempt employee is paid by the hour for all hours worked and all hours worked are to be reported on this time sheet. Pay for hours worked is

COPIES: Send the first four copies to your supervisor. Keep the final gold copy for your records.

6.2 Travel Procedures

Westat travel practices must reflect both the highest professional standards and our obligation to stay within our contract budget.

6.2.1 Trip and Expense Reporting

Westat uses the **Travel Expense Report Form (TER)** (see Exhibit 6-2) for travelers (overnight stay required) to report the time spent and expenses incurred while conducting project business. It is essential that assessors report their expenses each day.

This report should be used for reporting all cash and personally charged expenditures related to lodging, per diem, and transportation. **Attach receipts using a paperclip for all expenditures except meals to the form.**

Exhibit 6-2. Trip Expense Report Form (TER)

TRIP EXPENSE REPORT – FIELD

Complete for **EMPLOYEE PAID EXPENSES** for **out of town trips ONLY**

WINS #: W11221-13141516 Leave Residence/Office at 8:00 a.m./p.m. on 2/18/10 (Date)
 Name: Reynolds Burt A Return Residence/Office at 4:00 a.m./p.m. on 2/10/10 (Date)
 (Last) (First) (MI) Project Number 8110104 Project Name PARA Study
 Address: 12 Main St, Pittsburg, PA 12345 Destination: City Eric City State PA
 Purpose of Trip: Assessments in Eric City Employee's Signature: Burt Reynolds Date 2/11/2010
 Approver's Signature: [Signature] Date 2/15/2010

Day of Week MM/DD/YY	2/18/10	2/19/10	2/10/10	1/1	1/1	1/1	1/1	TOTAL
Meals & Incidentals Per Diem	\$34.50	\$46.00	\$34.50					\$115.00 (DMeals)
Employee Paid Lodging	\$85.00	\$85.00						\$170.00 (DLodg)
Employee Paid Airfare								\$ (DAir)
Employee Paid Auto Rental								\$ (DAuto)
Mileage Allowance: (280 miles @ .50 per mile)	\$70.00		\$70.00					\$140.00 (DMiles)
Ground Transportation (i.e., taxi, metro, bus)								\$ (DGround)
Telephone - Business								\$ (DPhone)
Other: (Specify) <u>Gas</u>	\$40.00		\$20.00					\$60.00 (DOther)
								\$
Total Claimed	\$229.50	\$131.00	\$124.50					\$485.00

FOR OFFICE USE ONLY

Outstanding Advance: _____
 Reimbursement Amount: _____

Charge Code: _____ Amount: _____
 _____ Amount: _____

COPIES: WHITE: A/P BLUE: HOME OFFICE YELLOW: SUPERVISOR PINK: EMPLOYEE

6.2.2 *Completing the Trip Expense Report Form*

Instructions for completing the TER are as follows:

WINS#, NAME, ADDRESS:	In the upper left corner fill in your WINS number, your name, your <u>permanent address</u> .
PURPOSE OF TRIP:	Write "Assessments for PARA study"
PROJECT:	The project number is 8110.01.04. Record "PARA" as the Project Name.
ITINERARY:	In the upper right corner circle "Residence" and enter the time and date of your departure at the beginning of your trip, circling a.m. or p.m., and the time and date of your return at the end of your trip. If you are traveling between sites, "Residence" would also mean the hotel where you are staying. If your stay in one place is longer than a week, then you would not enter the return information.
DESTINATION:	Enter the city and state of your travel destination.
EMPLOYEE'S SIGNATURE:	You must sign every TER. Westat will not process unsigned TERs.
APPROVER'S SIGNATURE	Leave this line blank.
DAY OF WEEK: MM/DD/YY:	Record the days of the week that you are traveling and the date under the corresponding day. You need to enter the month, day, and year. The days should correspond to your trip rather than your T&E.
	Record costs incurred for each day of travel and for days that you are in the field. If an expense is charged on your personal credit card, it should be treated as a cash expenditure and therefore shown on your report.
MEALS & INCIDENTALS PER DIEM:	Enter the per diem amount for each full day of travel (This amount will depend on the city and state). Do not exceed the daily per diem.
EMPLOYEE PAID LODGING:	Enter the total cost of lodging, including tax. If Westat has arranged for direct billing, write "Paid by Westat" on the line.
EMPLOYEE PAID AIRFARE:	(not applicable for this project)
EMPLOYEE PAID AUTO RENTAL:	(not applicable for this project)
MILEAGE ALLOWANCE:	If you are using your own automobile, Westat will reimburse you at a flat rate of .50 cents per mile for business-related travel. Multiply the number of miles traveled each day by .50 cents per mile and enter that number in the grid for the appropriate day. If you are using a rental car, leave this row blank. Remember that mileage here refers only to the number of miles traveled between your home and the hotel, or from one hotel to another when you are changing sites.
GROUND TRANSPORTATION (I.E., TAXI, METRO, BUS):	Enter the total amount, including tip.
TELEPHONE-BUSINESS:	Include costs incurred and receipts.

OTHER (SPECIFY): Enter any other expenses in this section. This would include parking, tolls, and gas for rental cars. Charge Westat only for the gas used while driving the rental car for business purposes (and ordinary living travel, such as to dinner or shopping). If you make any side or tourist trips, the gas for this should be paid directly by you.

If additional space is needed for miscellaneous expenses, attach a separate sheet of paper and itemize, making certain that the expenses are reported by day and date incurred.

TOTAL CLAIMED: Add up each day's total expenses and enter the totals in this row.

TOTAL: On the far right is a total column by expense. Add across each row and enter the weekly total for each type of expenditure on the appropriate line. Add the Total Claimed for each day of the week as well, verifying that the totals added vertically (by type) and horizontally (by day) are equal.

FOR OFFICE USE ONLY: Leave this section blank.

6.2.3 Travel Specifics

Usually the need to stay overnight is obvious. However, if you feel that the distance from your home to a site would require excessive time or distance, you or your Field Supervisor must coordinate travel with the Westat Travel office (Field Supervisor must request approval of Field Manager).

For assignments requiring rental cars, and lodging, arrangements will be made using the Westat Travel Office. After assessors receive their travel assignment and have discussed the travel requirements (e.g. the origin and destination of their trip), the Field Supervisor will contact the Westat Travel Office to make arrangements, if the assessor prefers this option..

Westat Travel Office: 1-800-544-7755 (leave a message)
Regular Hours: Monday-Friday 8:30am-5:30pm EST
After hour emergencies call Omega Travel 1-800-685-6342

Assessors have the choice of making their own travel arrangements by calling the Westat Travel Office. **However, Assessors that choose to do this are REQUIRED to ask the Westat Travel Agent to email a copy of their itinerary to their Field Supervisor.** Field Supervisors need to have any travel information for their assessors at hand at any point in time.

Personal Auto Usage

The use of a personal automobile is usually used under one of two circumstances: for travel to and from airports/train terminals/bus terminals, or for travel to and from the site where they will be working. **The current reimbursement rate for the business use**

of personal automobiles is .50 cents per mile, which includes gasoline, insurance, and upkeep. **No reimbursement will be made for traffic tickets, car repairs, car upkeep, or towing.**

Using a personal automobile to and from the embarkation point is authorized if the total cost, including parking, is less than or equal to the cost of taxis or buses. A personal automobile can also be used for travel to and from the site where they will be working. Once assessors arrive at their destination and begin working, the usual costs associated with the use of a personal automobile during assignments (such as parking, mileage to their assigned schools from the hotel and back, tolls, etc.) should be entered on their T&E. If driving far enough to merit an overnight stay, once the assessor reaches the hotel, they should think of it as home and list personal automobile expenses incurred locally on the T&E. The trip from their residence to the hotel and the trip from the hotel back to their residence should be reported on the TER. **Receipts must be attached to the TER for all expenses associated with travel to and from their destination.**

Rental Cars

Generally, car rentals will be used when assessors are required to go on an out-of-town assignment where a fair amount of travel within a city or between nearby cities will be necessary. **Rental car arrangements must always be made through a Westat Travel agent.** This will insure insurance is covered on the rental car via the corporate agreement between Westat and the rental agency.

Gasoline purchased for rental cars is reimbursed by Westat if the assessor attaches the receipts to their TER. The gasoline cost should be listed under the column for the day it was purchased in the “Other” section of the TER. Westat will pay for the gasoline used in conducting project business, going to dinner, and doing other necessary errands in and around the site.

Lodging

The Assessor’s hotel accommodations will be coordinated with Westat’s Travel Office. **No changes in accommodations can be made without Westat authorization.** If Westat can arrange direct billing with a hotel, the assessor may bill only the room and tax charges directly to the home office. Any charges for meals, telephone calls, or incidental charges billed to the assessor’s room should be paid by the assessor by cash or personal credit card and submitted on the assessor’s TER. These charges should not be added to the hotel bill paid by Westat.

The assessor must submit hotel receipts and record the amount on the TER in order to receive reimbursement. If the assessor stays at a specific hotel for more than 1 week, a copy of the receipt must be attached to every T&E on which that stay is billed.

Per Diem: Meal Allowance and Incidentals

The assessor will be reimbursed for the approved per diem allowance for the area in which the assessor is on travel for each full (24-hour) day the assessor is on travel status. The first and the last day of travel, Westat will pay 75% of the per diem allowance. Certain high-cost cities allow for a higher meal and incidental allowance. The Field Supervisor will inform the assessor of the per diem amount for the city the assessor is traveling to. The assessor can also look up the government per diem rates online at the US General Services Administration web site:http://www.gsa.gov/Portal/gsa/ep/contentView.do?contentType=GSA_BASIC&contentId=17943&noc=T

The per diem allowances include “incidentals” such as tips, personal phone calls, and laundry. These expenses included in the per diem will be paid for each day of travel, as well as the assessor non-working days while assigned to the assessor site. **Westat will not pay for any lodging or per diem while the assessor is at home or on a vacation break.**

6.2.4 Lodging and Per Diem in Special Situations

Lodging and per diem will be paid for every day that the assessor is assigned to a site away from their hometown. Special provisions are made for the following situations:

- No per diem or lodging can be paid while the assessor is in their hometown.
- When the assessor is traveling between study sites and staying with friends or relatives, the assessor may charge the project \$25.00 per night for lodging plus the per diem.

6.2.5 Miscellaneous Expenses

Westat will not pay for entertainment or liquor expenses. If incurred, the assessor will need to pay for those expenses. Other incidental expenses will not be reimbursed without specific prior approval from the Field Supervisor or Field Manager. The fee charged to the purchase of traveler’s checks is an allowable expense. Also, any fee for cashing a check is an authorized expense. Be sure to submit a receipt.

6.2.6 Travel Advances

The assessor may request a travel advance to pay for allowable expenses incurred while traveling on Westat business. However, **the assessor must account for the full amount of the travel advance through use of a TER and supporting receipts.** This amount is not payment to the assessor but only a loan to cover the expenses for per diem, gasoline, and hotels while the assessor is on travel status until the assessor submits their TER and has been reimbursed. Upon completion of the assignment, any excess funds must be returned to Westat in the form of either a check or money order made payable to Westat.

Travel advance requests must be submitted to the Field Supervisor by Mondays of each week to receive funds on the following Friday if the assessor has direct deposit or Monday if the check is mailed to the assessor.

7. PROFESSIONAL CONDUCT

Professional conduct is multifaceted and extremely important. You will be going to different schools, and the types of personnel and people associated with each school will vary. It is important that your arrival on the first day and all succeeding days at any of these settings runs smoothly. This can best be achieved through professional, friendly, and respectful interactions with all staff. We have all heard the expression that “first impressions last,” but in reality, the impressions that you and your team create on an on-going basis are crucial to the success of this study.

7.1 Professional Conduct in Schools

Members of the study staff must conduct themselves in a professional manner at all times. Our visit should leave a good impression and make the participants feel comfortable. We want to be welcomed back next visit!

Your appearance is the first thing that will be noticed. In general, you should aim for an appearance that is neat, suitable, and inconspicuous. Business-type attire is encouraged, but you should not be overdressed. Do not wear jeans; dress pants or slacks are acceptable, and men are not required to wear ties. Do not wear perfume or cologne of any kind as some participants may have allergic reactions or be distracted. Avoid loud colors (e.g., hot pink, bright green) as well as large or noisy jewelry that may distract the children or staff. Particularly during data collection, avoid wearing black and white checks. Some research has shown that if a tester is wearing them, the child is much more likely not to look at the assessor or anything the assessor is holding because the checks are visually overpowering. Do not eat, drink, or chew gum during assessments. Remember to wear your PARA ID badge at all times.

7.2 Cultural Sensitivity

We live in a world that encompasses many different cultures. As professionals we need to be culturally competent, that is, able to honor and respect those beliefs, interpersonal styles, attitudes, and behaviors both of the families and the multicultural staff who are providing services. For example, depending on cultural beliefs a person may or may not shake hands when meeting or may make eye contact or avoid it. It is difficult to anticipate every situation; therefore, to accommodate the wide range of cultural beliefs that you will encounter, at least four traits are important:

- **Flexibility:** You may be used to doing things in a particular manner. However, follow rules of etiquette that you observe from either project staff, teachers, or parents that you feel will put them at ease with your presence. However, do not violate the protocol of the study.

- **Sensitivity:** Please keep in mind that most of the children that you will be testing have especial needs. Be patient and be caring, but be very careful not to come across as patronizing. Act natural and treat them with dignity and respect. Project staff must be careful not to pass judgment either through facial expressions or verbal comments on any information gathered or observed behavior.
- **Trust:** Time, respect, consistency, and follow-through are important in developing a relationship that includes trust.
- **Recognition of Priorities:** Although our priority is to complete this study in an efficient and timely fashion, we need to realize that we are guests at the schools and participation is voluntary.

7.3 Interacting with School Staff and Teachers

Research in schools requires a masterful and flexible approach in dealing with administrative staff, teachers, parents, and children. The relationships established will define our success. Close and frequent communication and an effective working relationship between PARA team members and project staff are crucial for the smooth flow of information. It is very important that all staff involved with PARA be fully informed.

School contacts (liaisons) may need to feel that they are active participants in the study and that their suggestions and comments are valued. However, remember that data collected cannot be shared with them. It will be the school contact's responsibility to arrange with the classroom teachers to excuse participant students to take the assessment. However, if a conflict between PARA assessment and classroom procedures arises, (e.g., a teacher or aide wishes to be present during the assessment), the school contact should discuss privately with the teacher. As an assessor, it is not your job to get involved with the teachers. The school contact should explain that our study procedures call for children to be assessed by our trained staff in a quiet setting free of distractions. If school staff requests to be present during the assessment, ask them to sit behind the child and not to coach.

We need to maintain as similar conditions as possible for all study participants. Neither children, parents, nor school staff should hear conflicting information. If they ask you a question that you are not sure of the answer, it is important that you tell them that you will get back to them. Either find the information in your manual and printed material or ask your Field Supervisor, but always make sure you get back to them with an answer as promptly as you can. If, while at the school you learn or notice any problems that arise between PARA staff and school personnel, please discuss the situation with your Field Supervisor. In general, Field Supervisors should handle matters that are sensitive. Some discussions may necessitate the assistance of the school contact.

Office staff, where applicable, can also be helpful in easing our task. We appreciate what everyone is doing to help us, and we must continually inform staff of our gratitude. Be mindful that PARA is not the staff's only concern or responsibility during the time we spend on site.

When you arrive to the school sign in at the front office to let them know of your arrival. Introduce yourself and notify the personnel that you are there to conduct PARA study activities. You may be asked to wait for an escort or be given a pass and shown directions to the office or class. In either case, remember that you are a guest in the building and, as such, must familiarize yourself with the school's rules. Be polite, unobtrusive, and maintain a low profile. Remember to ask about restrooms, a storage area, or other information that may be useful to you before you leave the office (if this has not already been determined by your Field Supervisor while setting up the appointment). While at the site, refrain from discussing personal plans and experiences with fellow school staff, or children.

At lunchtime, project staff may indicate that they made arrangements for you to join them. If this is the case, you may choose to accept their offer, but please bring your own lunch unless they insist on providing it.

Please be sure to keep the following protocols in mind when carrying out assigned tasks.

- If you are working at a school as part of a team you will need to arrive at the same time. Sign in with the front office to let them know you are there. The school contact should have arranged a place for you to conduct your work.
- Introduce yourself to the school contact and other school staff with whom you come in contact. Ask preliminary questions about the schedule, children, or other matters that are appropriate to your tasks at this time. Discuss any concerns that you have or that they may have, as well as protocols.
- Do not hug, kiss, or touch the children unless their health or well-being is in jeopardy. As cute as they may be, it is unprofessional and may cause discomfort among the staff or parents.
- Do not in any way intrude into the school staff's personal space. In short, follow the common rules of courtesy and respect while in the building. We want to be respected for our professionalism and demeanor, reduce as much of the burden on the schools we can, and be welcomed back for any necessary follow-up activities.
- Discuss with the school contact the assessment schedule for the day and the most effective way to carry it out. Your Field Supervisor will help you to coordinate this transition smoothly.
- It is best to have the school contact introduce you to the student prior to starting the assessment. Also, try to have a short informal conversation with the students, for

example, ask them what subject they like best, before starting with the assessment.. This gives the student a chance to become used to you and to feel comfortable with you.

7.4 Confidentiality and Security

We pledge to the participants that we will preserve and protect the standards of confidentiality. Participants must not only be convinced of the legitimacy and value of our studies, but they must trust that their responses will be treated in the strictest confidence. We want the student participants to answer freely with the confidence that no one outside the project will see or hear about their responses. Therefore, the security of the assessment booklets must be protected at all times. Proper handling and storage of these materials are critical to ensure against loss, breach of security or participant confidentiality, and other hazards. Be careful not to discuss any aspects of the data gathered while on school the grounds or in any public location such as restaurants, hotels, or stores where you might be overheard.

If either the participants or school personnel have questions about who will see the data, tell them that only authorized project staff will see the data and that no names or any other identifying information will be used in reports. While you are onsite, do not discuss any specific details of the data collected with any of the school staff, teachers, or administrators.

While conducting the assessments, the unused materials and booklets must be secured and kept with you. Take all materials with you wherever you go. Never leave materials lying around for any child or staff member to pick up or see. This includes blank booklets and forms.

7.5 Policy on Suspected Child Abuse

Though it is highly unlikely and not anticipated, field staff may encounter a case of suspected child abuse on the part of a family member or a staff member at the school. Assessors should discuss it only with their Field Supervisor. Be sure to have this discussion off school grounds and in complete privacy. It is, as you know, an extremely sensitive issue. Be careful in any case not to jump to conclusions. Do not discuss your suspicions with any project staff, other PARA team members, or anyone else. If necessary, your Field Supervisor will contact the Field Director, Juanita Lucas-McLean, at Westat, who may in turn contact you, and a decision will be made about how to proceed.

Of course, if you *witness* any physical child abuse, you are obligated to get help and notify the authorities. That is, depending on the situation, you may want to enlist the aid of other adults in stopping the abuse. Then notify the school principal and possibly the police. Again, you should contact your Field Supervisor.

Appendix A

Frequently Asked Questions

General Testing Issues

1. Prior to administering the test, is it important to go over the special “situated motivation questions” that are in boxed text after each passage and set of comprehension questions?

Yes. The questions in the boxed text are worded differently than the comprehension questions following each story and the set of questions also differs between narrative texts and expository texts.

2. Should we remind students that the test is not timed?

Yes. This is important because part of our study seeks to put students at ease so that the pressure of time doesn't cause anxiety.

3. What about students who need more time than their peers?

Students should be allowed the time they need to complete the assessment. Generally, only a few students need extra time. In other situations like this, proctors have sometimes taken these students to a predetermined quiet area and allowed them to finish. Students should finish in 1.5 hours or 2.0 total hours.

4. What criteria should test administrators use when making the determination that a student can not complete the assessment?

In past situations, test administrators have noted that students finished when given time and encouragement to finish (but no prompts to “keep working” because these would be additional accommodations that are not built into our assessment). If encouragement is given several times and the student clearly has given up or indicates that they do not plan to finish, then the test administrator should write notes to this effect on the form of the test booklet, indicating that the student did not finish and why.

Appendix B Frequently Asked Questions Accommodations

1. Why do I have to make sure that students with disabilities receive the accommodations designated for them on the student roster?

We want students to be taking the PARA assessment with the supports that they typically receive when taking standardized assessments. This will make the results of our assessment more valid.

2. What should I do if a student does not want to use an accommodation that the student roster says he or she should get?

You should ask the student why he or she does not want to use the accommodation, then note that the accommodation was not provided in the booklet (specify which accommodation) and the reason the student gave for not wanting to use the accommodation.

3. What should happen if a student needs a sign language interpreter?

We should know this information prior to testing and an interpreter should be available. Make sure to have the interpreter sign the confidentiality form.

4. Do I need to record the accommodations that a student uses during the assessment?

Yes, please record this information in the test booklet, indicating the accommodation and how it was used by the student.

5. What should I do if students want (and need) more than 2 hours to complete the assessment?

Generally, only a few students need extra time. In other situations like this, proctors have sometimes taken these students to a predetermined quiet area and allowed them finish. Students should finish in 1.5 hours or 2.0 total hours (2 hours is the longest amount of time a student should require).

6. What should I do if students request (and should receive re: IEP) a read aloud accommodation, and this is listed in the IEP as an accommodation?

You should provide this and then please record this information in the test booklet, indicating this particular accommodation.

7. What happens if students request a break between reading several of the passages, and this is listed in the IEP as an accommodation?

You should provide this and then please record this information in the test booklet, indicating this particular accommodation.

8. What happens if students request to take the test in a separate room/quiet place, and this is listed in the IEP as an accommodation?

You should provide this and then please record this information in the test booklet, indicating this particular accommodation.

9. Should I prompt students to keep working through the assessment if they appear to be pausing or at risk of not working through the passages?

You may offer general encouragement to finish the assessment, but no specific prompts because these, in and of themselves, are a form of an accommodation. A specific prompt would include something like “You are working so hard and you only have 3 questions to go; keep at it you can get these questions completed.” Instead, a general prompt is “Thank you for working on the assessment. Please keep working to finish it.” If you offer several general prompts to a student, please record this information in the test booklet.

10. What should I do if I notice that students with disabilities have (accidentally) skipped items or an entire passages and items in the assessment booklet?

Please point this out to the student and redirect their efforts to completing these missing items or passages.

Appendix L

MRQ Administration and Logistical Instructions, Including Frequently Asked Questions and Answers



Partnership for Accessible Reading Assessment

MOTIVATIONAL READING QUESTIONNAIRE (MRQ)

The Motivations for Reading Questionnaire (MRQ) measures different dimensions or aspects of elementary school-age children's reading motivations. It can be used to discover the nature of children's motivations for reading, as well as some of the things about reading children do not find motivating. It consists of 54 items, and can be administered in 15-30 minutes.

ADMINISTRATION INSTRUCTIONS

1. Before bringing the study children to the computer lab to complete the MRQ online survey, you should have been provided with the following materials:
 - a. A copy of the Student List: these are the students who will do the MRQ online survey.
 - b. The Username/Password cards: Each study child will have need one of these (with their name on it) to gain access to the online survey.
 - c. The MRQ Administration Instructions: These are instructions for you on what to say to the children before starting the survey.
 - d. The MRQs Not Completed List: If any children are unable to complete the MRQ during the month of February, we ask that you include their name on this list along with the reason. This list should be sent back to Westat at the end of data collection.
2. A computer lab should have been secured for 15-30 minutes on the scheduled day.
3. You will be responsible for pulling out the study children, taking them to the computer lab, and proctoring the survey.
4. When children are in the computer lab, say the following:

Good morning/Good afternoon students! You are being asked to be part of a research study that looks at ways to make reading tests better for students. Each year you take a reading

test that all of the other students in grades 4 or 8 in this state take. Sometimes those tests are hard for kids. A group called the Partnership for Accessible Reading Assessments would like some information from kids about ways these tests could be done differently. We are studying what would help some students “show what they know” more easily. Today we will ask you to complete a survey on the computer. The statements on the survey tell how some students feel about reading. You will silently read each statement and decide whether it talks about a person who is like you or different from you. Remember there are no right or wrong answers. We only want to know how you feel about reading.

[IF STUDENTS HAVE NOT TAKEN READING TEST YET] After today, someone will visit our school and ask you to take a reading test where you will read four short reading passages and answer some questions. These are not real tests that you’ll receive a grade on, but we do want you to do your very best! For having completed the survey we’re going to do today and the reading test after today, you will receive a gift card to a store in your area.

[IF STUDENTS HAVE TAKEN READING TEST ALREADY] Before today, someone visited our school and you took a reading test where you read four short reading passages and answered some questions. For having completed the survey we’re going to do today and the reading test you already took, you received a gift card to a store in your area.

5. Instruct the children to access the following web link: www.parasurvey.org
6. Pass out the Username/Password cards to students.
7. Say: *Before you begin the survey, let’s look at the examples together:*
 - a. *Here are three examples (have students look at these on the survey)*
 - i. I like ice cream
 - ii. I like to swim
 - iii. I like spinach
 - b. *If the statement is **very different from** you, check the radio button next to # 1*
 - c. *If the statement is **a little different from** you, check the radio button next to # 2.*
 - d. *If the statement is **a little like you**, check the radio button next to # 3.*
 - e. *If the statement is **a lot like you**, check the radio button next to # 4.*
8. Say: *This is not a timed survey, but please try to work at a steady pace. Read the directions throughout the survey and when I say you can begin, please start working.*

If you have any questions about what you need to do, or if you need help reading a word, or if you don't understand a statement, please raise your hand. You may begin.

9. When children have completed the MRQ please collect the Username/Password cards. These should be destroyed.
10. Please complete the MRQs Not Completed List if any of the students on the list did not complete the MRQ during the February data collection period.

MOTIVATIONAL READING QUESTIONNAIRE (MRQ)

The Motivations for Reading Questionnaire (MRQ) measures different dimensions or aspects of elementary school-age children’s reading motivations. It can be used to discover the nature of children’s motivations for reading, as well as some of the things about reading children do not find motivating. It consists of 54 items, and can be administered in 15-30 minutes.

LOGISTICAL INSTRUCTIONS

11. Review List of Students to see who needs to complete the online survey.
12. Gather the Username/Password cards for the students who will take the MRQ
13. Secure a computer lab for the students
14. Decide the best day and time for the MRQ
 - a. Allow 15-30 minutes for the students to complete the online survey
 - b. The MRQ can be done before or after the reading test, during the month of February
15. Decide who will be responsible for pulling out the study children, taking them to the computer lab, and proctoring the survey.
 - a. Provide this person with:
 - i. A copy of the Student List
 - ii. The Username/Password cards
 - iii. The MRQ Administration Instructions
 - iv. The MRQs Not Completed List
16. When students take the survey, the students must use enter their name exactly as it appears on the class roster since name is a key field used to associate all students’ information across different databases.
17. Students should enter in their teacher’s last name only, as directed by the teacher (e.g., Smith).
18. Students should enter in the name of the school as directed by the teacher. The name can be a short version (e.g., Valley View) without the word “Elementary” or “Middle” or “School.



MOTIVATIONAL READING QUESTIONNAIRE (MRQ)

Frequently Asked Questions

1. How do the students get into the MRQ survey tool?

You will be provided with Username/Password cards for sampled children. The URL for the MRQ survey is www.parasurvey.org

2. Should survey proctors do the sample items with the group as a whole before students progress through the actual items on their own?

Yes, proctors should do the sample items with the group as a whole before students progress through the actual items on their own. It is not necessary, but it may be helpful to the students and reduce questions later on. It also will help them understand how to discriminate between the answer choices and this makes the survey more valid.

3. Should survey proctors explain how the MRQ survey works in terms of students needing to complete all items on a computer page before the survey tool allows them to progress forward?

Yes—there are restrictions built into the survey that require students to fill in all three items on the computer “page” before moving on with the survey (or the survey won’t let you progress forward). It would be wise for proctors to explain this to students when they give the survey directions and prepare proctors that this may confuse some of our special education students at first.

4. Are any of the fields in the MRQ survey required?

Yes, all of the questions are required to obtain a quality subscale score and so that the analyses and reports can be complete and accurate.

5. Will the students have a sense of how much of the MRQ survey they have completed and how much remains?

Yes. The survey displays a completion thermometer bar, to show the child how much of the survey they’ve completed.

6. Should proctors help students with any of the MRQ survey questions or help them understand words in the answers choices?

Yes, you may do this. The MRQ is very understandable and straightforward but if students have questions about words in the sentences or if they want a statement read aloud, that would be fine—it won't change the construct being measured. However, proctors should not suggest a response or indicate how they would respond because this might influence a student's choices.

7. Can students do the MRQ survey in two time blocks?

Yes, but it is not preferable. A child can stop in the middle of a "page" in the computerized survey, press "Save & Exit" and close their browser. When they enter the survey login again, they will be taken to the last page they were on, with any previous answers they entered on that page displayed. If they do not press "Save & Exit," they will still resume on the last page they saw, but their answers on that page will not have been saved.

Appendix M

Equipercentile Test Equating Procedures Used to Equate Across Testing Times, Including Computational Approach and Conversion Tables

Before conducting equating, we examined the dimensionality of the reading assessments. Exploratory factor analysis was used to compare the underlying factor structure of both the accessible and state proxy assessment forms for grades 4 (5) and 8 (9). Scree plots, fit statistics, and item loadings were all used as criteria for identifying the number of factors underlying each assessment. In all cases, the first eigenvalue was much larger than the remaining eigenvalues indicating a dominant factor. The first eigenvalue ranged from 9.12 to 10.69, whereas the second eigenvalue ranged from 1.29 to 2.00. The magnitude of the difference between the eigenvalues of the first and second factor was large and point to a single factor underlying each form. In addition, the eigenvalues fell on a relatively straight line after the first factor (scree) supporting the notion that one factor underlies the item responses.

Table M1 presents the fit statistics for a one factor versus two factor model for each test form by grade. The fit for the one and two factor models were all equally acceptable. The values for the root mean square error of approximation (RMSEA) were all less than .05 and the CFI/TLI indices were above .96 for a one and two factor model. While values of the selected fit indices improved from a one to two factor solution, the differences in fit were small enough to suggest a single factor model for each test form. Finally, the pattern of factor loadings was examined to understand the item-level associations to the one and two factor model. In all two-factor solutions, the vast majority of items loaded strongly on the first factor and there were no items with cross-loadings. Together, the scree plots, fit indices, and item loadings provided a clear indication that a single factor is more likely to fit the data than two factors for both the accessible and the proxy test forms.

After obtaining data supporting uni-dimensionality of items in both accessible and proxy tests, we conducted equipercentile equating separately for students in grade 4 (grade 5 in fall) and students in grade 8 (grade 9 in fall). Below is a short description of the analyses.

Equipercentile Equating for Grade 4 (grade 5 in fall)

The fourth grade proxy state and accessible versions of the assessment were equated using equipercentile equating with pre-smoothing of both distributions. Descriptive statistics for the two distributions are contained in Table M2. The proxy state assessment has a slightly greater mean and standard deviation compared with the accessible version of the test. Both the state proxy and accessible versions of the assessment show slight negative skew. The cumulative distributions are similar at the lower end of the scale indicating similar percentiles; however, the cumulative distributions diverge at the higher end of the scale suggesting a lack of correspondence of percentiles. Based on the cumulative distributions, the state proxy assessment appears to be slightly easier than the accessible version of the assessment especially for students in the upper end of the

distribution.

Equipercntile equating was carried out with the RAGE-RGEQUATE program with the goal of adjusting scores on the state proxy assessment to be comparable with scores on the accessible version of the test. RAGE-RGEQUATE conducts random groups equipercntile equating using cubic spline post-smoothing, log-linear pre-smoothing, and beta-4 pre-smoothing.

Observed distributions were pre-smoothed via log-linear smoothing with $C = 4$ based on fit and changes in fit. After pre-smoothing, equipercntile was conducted on the smoothed distributions and scores on the state proxy assessment were re-scaled to comparative scores on the accessible assessment. This conversion is described in Table M3. Descriptive statistics for the accessible and re-scaled proxy state assessments are presented in Table M4.

Equipercntile Equating for Grade 8 (Grade 9 in fall)

The 8th grade proxy state and accessible assessments were equated using equipercntile equating with pre-smoothing of both distributions. Descriptive statistics for the two distributions are shown in Table M5. The accessible assessment has a slightly greater mean compared with the proxy state assessment. Both the proxy state and accessible assessments are platykurtic. The cumulative distributions are offset, that is, the accessible assessment appears easier for all of the points along the scale.

Equipercntile equating was carried out with the RAGE-RGEQUATE program with the goal of adjusting scores on the state proxy version of the test to be comparable with scores on the accessible version of the test. RAGE-RGEQUATE conducts random groups equipercntile equating using cubic spline post-smoothing, log-linear pre-smoothing, and beta-4 pre-smoothing.

Observed distributions were pre-smoothed via log-linear smoothing with $C = 4$ based on fit and changes in fit. After pre-smoothing, equipercntile was conducted on the smoothed distributions and scores on the state proxy version of the test were re-scaled to comparative scores on the accessible test. This conversion is contained in Table M6. Descriptive statistics for the accessible and re-scaled state proxy are contained in Table M7.

Appendix M: Tables

Table M1. Fit indices from EFA for 1 and 2 factor solutions by grade

	1-FACTOR			2-FACTOR		
	CFI	TLI	RMSEA	CFI	TLI	RMSEA
Grade 4						
State Proxy	0.99	0.99	0.03	1.00	0.99	0.02
Accessible	0.98	0.98	0.03	1.00	1.00	0.01
Grade 8						
State Proxy	0.97	0.96	0.04	0.99	0.99	0.02
Accessible	0.98	0.99	0.02	0.99	0.99	0.02

Table M2. Descriptive Statistics for Grade 4 State Proxy and Accessible Assessment Scores

	Grade 4 State Proxy	Grade 4 Accessible
N	563	564
Mean	18.84	17.70
Standard Deviation	5.96	5.61
Skewness	-.49	-.46
Kurtosis	-.75	-.64

Table M3. Conversion Table for Grade 4 State Proxy

Grade 4 State Proxy Observed Score	Grade 4 State Proxy Equated Score
0	-0.39
1	0.51
2	1.44
3	2.42
4	3.43
5	4.45
6	5.47
7	6.49
8	7.51
9	8.52
10	9.52
11	10.51
12	11.48
13	12.42
14	13.33
15	14.22
16	15.10
17	15.97
18	16.84
19	17.72
20	18.63
21	19.55
22	20.49
23	21.45
24	22.43
25	23.44
26	24.48
27	25.62
28	27.06

Table M4. Descriptive statistics for Grade 4 State Proxy Rescaled and Accessible Assessment Scores

	Grade 4 State Proxy Rescaled	Grade 4 Accessible
N	563	564
Mean	17.70	17.70
Standard Deviation	5.59	5.61
Skewness	-.46	-.46
Kurtosis	-.66	-.64

Table M5. Descriptive Statistics for Grade 8 State Proxy and Accessible Assessment Scores

	Grade 8 State Proxy	Grade 8 Accessible
N	919	892
Mean	14.75	16.40
Standard Deviation	6.06	6.15
Skewness	.00	-.12
Kurtosis	-1.08	-1.12

Table M6. Conversion Table for Grade 8 State Proxy

Grade 8 State Proxy Observed Score	Grade 8 State Proxy Equated Score
0	1.12
1	2.22
2	3.21
3	4.21
4	5.21
5	6.23
6	7.25
7	8.30
8	9.36
9	10.44
10	11.55
11	12.67
12	13.80
13	14.92
14	16.01
15	17.05
16	18.05
17	19.01
18	19.95
19	20.86
20	21.77
21	22.67
22	23.56
23	24.46
24	24.46
25	26.28
26	27.16
27	27.92
28	28.39

Table M7. Descriptive Statistics for Grade 8 State Proxy Rescaled and Accessible Assessment Scores

	Grade 8 State Proxy	Grade 8 Accessible
N	919	892
Mean	16.40	16.40
Standard Deviation	6.14	6.15
Skewness	-.11	-.12
Kurtosis	-1.12	-1.12

Appendix N

Analyses of Primary Research Questions for Scores Equated Separately for the Accessible and Proxy State Assessments

(Note: Results are Similar to Those for Results When Scores were Equated Together)

In previous analyses, the fourth grade and the eighth grade state proxy and accessible versions of the assessment were equated using equipercntile equating with pre-smoothing of both distributions. Analyses were conducted based on the equated scores across times of administration, that is, spring and fall. To make sure the equating process did not have any unanticipated impact on the results of the study when comparing students' performance on the accessible and proxy assessments, we conducted equating separately for accessible and proxy assessments and re-analyzed the data based on the new equated outcomes.

Table N1 presents a summary of descriptive analyses for the proxy assessment results (see Table N2 for the conversion data. For the proxy assessment) Similarly, Table N3 presents the descriptive statistics for the accessible test for grade 4 (see Table N4 for the conversion data for the accessible assessment). As data in Table N1 show, the mean proxy assessment score for the spring administration ($M=18.66$, $SD=5.66$) is almost identical to the mean for the fall administration ($M=18.98$, $SD=6.16$), indicating that the two administrations provided comparable outcomes. For the accessible assessment, however, the students assessed in the spring had a slightly lower mean ($M=16.91$, $SD=5.69$) than students in the fall ($M=18.30$, $SD=5.46$).

Table N1. Descriptive Statistics for Fall and Spring 4th Grade Proxy Assessment Scores Before and After Equating

	5th Grade Proxy Fall	4th Grade Proxy Spring	4th Grade Proxy Spring Equated Score
N	317	246	246
Mean	18.98	18.66	18.98
Standard Deviation	6.16	5.66	6.15
Skewness	-0.50	-0.49	-0.50
Kurtosis	2.20	2.29	2.20

Note: Observed distributions were pre-smoothed via log-linear smoothing with $C = 4$ based on fit (e.g., χ^2) and changes in fit.

Table N2. Conversion Table for 4th Grade Proxy across Seasons

4th Grade Proxy Fall Observed Score	4th Grade Proxy Spring Equated Score	4th Grade Proxy Fall Observed Score	4th Grade Proxy Spring Equated Score
0	-0.42	15	14.84
1	0.44	16	15.97
2	1.40	17	17.12
3	2.42	18	18.27
4	3.44	19	19.43
5	4.46	20	20.56
6	5.47	21	21.68
7	6.49	22	22.77
8	7.50	23	23.85
9	8.51	24	24.89
10	9.53	25	25.90
11	10.56	26	26.84
12	11.60	27	27.65
13	12.66	28	28.29
14	13.74		

Table N3. Descriptive Statistics for Fall and Spring 4th Grade Accessible Assessment Scores Before and After Equating

	5th Grade Accessible Fall	4th Grade Accessible Spring	4th Grade Accessible Spring Equated Score
N	320	244	244
Mean	18.30	16.91	18.3
Standard Deviation	5.46	5.69	5.45
Skewness	-0.48	-0.41	-0.48
Kurtosis	2.49	2.16	2.47

Note: Observed distributions were pre-smoothed via log-linear smoothing with $C = 9$ for Spring and $C = 4$ for Fall based on fit (e.g., χ^2) and changes in fit.

Table N4. Conversion Table for 4th Grade Accessible across Seasons

4th Grade Accessible Fall Observed Score	4th Grade Accessible Spring Equated Score	4th Grade Accessible Fall Observed Score	4th Grade Accessible Spring Equated Score
0	0.49	15	16.73
1	0.73	16	17.54
2	0.82	17	18.32
3	1.73	18	19.12
4	3.65	19	20.01
5	5.63	20	20.99
6	7.40	21	22.06
7	8.86	22	23.14
8	10.04	23	24.17
9	11.05	24	25.15
10	11.98	25	26.13
11	12.94	26	27.19
12	13.92	27	28.20
13	14.91	28	28.49
14	15.85		

The results for grade 8 students were similar to those reported for grade 4 students with some minor differences. Table N5 presents summary descriptive data for grade 8 students taking the proxy state assessment (see Table N6 for conversion data). Similarly, Table N7 presents summary statistics for grade 8 accessible assessments (see N8 for conversion data). As the data in Table N5 show, students in the spring performed slightly higher on proxy assessment ($M=15.28$, $SD=5.40$) than students in the fall ($M=14.27$, $SD=6.57$). For the accessible assessment, students in grade 8 performed higher ($M=17.80$, $SD=5.79$) than students in the fall ($M=15.06$, $SD=6.18$). Although these differences were relatively high in some cases, they did not systematically affect the scores in any direction.

Table N5. Descriptive Statistics for Fall and Spring 8th Grade Proxy Assessment Scores Before and After Equating

	8th Grade Proxy Fall	8th Grade Proxy Spring	8th Grade Proxy Spring Equated Score
N	481	438	438
Mean	14.27	15.28	14.27
Standard Deviation	6.57	5.40	6.55
Skewness	0.18	-0.21	0.18
Kurtosis	1.87	1.96	1.86

Note: Observed distributions were pre-smoothed via log-linear smoothing with $C = 4$ for Spring and $C = 5$ for Fall based on fit (e.g., χ^2) and changes in fit.

Table N6. Conversion Table for 8th Grade Proxy across Seasons

8th Grade Proxy Fall Observed Score	8th Grade Proxy Spring Equated Score	8th Grade Proxy Fall Observed Score	8th Grade Proxy Spring Equated Score
0	-0.49	15	12.47
1	-0.35	16	13.69
2	0.55	17	15.14
3	1.56	18	16.80
4	2.56	19	18.57
5	3.56	20	20.31
6	4.53	21	21.93
7	5.46	22	23.40
8	6.34184	23	24.79
9	7.20	24	26.07
10	8.04	25	27.21
11	8.86	26	28.11
12	9.68	27	28.44
13	10.52	28	28.49
14	11.44		

Table N7. Descriptive Statistics for Fall and Spring 8th Grade Accessible Assessment Scores Before and After Equating

	8th Grade Accessible Fall	8th Grade Accessible Spring	8th Grade Accessible Spring Equated Score
N	454	438	438
Mean	15.06	17.80	15.05
Standard Deviation	6.18	5.79	6.17
Skewness	0.19	-0.42	0.18
Kurtosis	1.94	2.12	1.95

Note: Observed distributions were pre-smoothed via log-linear smoothing with $C = 4$ based on fit (e.g., χ^2) and changes in fit.

Table N8. Conversion Table for 8th Grade Accessible across Seasons

8th Grade Accessible Fall Observed Score	8th Grade Accessible Spring Equated Score	8th Grade Accessible Fall Observed Score	8th Grade Accessible Spring Equated Score
0	-0.46	15	11.01
1	0.13	16	11.69
2	1.04	17	12.46
3	2.00	18	13.38
4	2.94	19	14.50
5	3.87	20	15.87
6	4.76	21	17.51
7	5.63	22	19.30
8	6.42	23	21.09
9	7.15	24	22.77
10	7.85	25	24.30
11	8.52	26	25.70
12	9.15	27	26.99
13	9.76	28	28.08
14	10.37		

The results from each separate analysis were consistent with the analyses presented earlier based on the first equating plan. The results did not support the research question/hypothesis number 1. That is, students with LD and students with disabilities combined (LD, SLI, MR, and D/HH) performed the same on the accessible assessment and the proxy state assessment. The results of analyses based on the new equated scores for research question number 2 were also consistent with the results presented earlier based on the first equating plan. That is, the reliability of scores was not higher for the accessible reading assessment when compared with the proxy state assessment for students with LD, and for the combined group of students with disabilities.

For research question number 3, results from the first and second equating plans were partially consistent. These results indicated that non-special education students performed better on the proxy state assessment than on the accessible assessment. No differences were observed in the performance of LD students across assessments. The analysis outcomes based on the first and second equating plans confirmed that incorporating accessibility features into the design did not compromise the validity of assessments.

In addressing research question 4 of this study, the outcomes of the analyses based on the second equating plan were also consistent with the first equating plan and indicated in both grades 4 and 8 alike: (a) students with LD and students with disabilities combined (LD, SLI, MR, and D/HH) performed the same on the accessible assessment and the proxy state assessment; (b) reliability of accessible reading assessments for students with LD and students with disabilities combined on the PARA accessible reading assessment and the proxy state assessment were the same; and (c) non-special education students performed the same on the PARA accessible assessment. Once again, conducting equipercentile equating separately for the accessible and proxy assessments did not have much effect on the outcomes of the analyses for three of the four primary research questions.