January 2014

Reading Comprehension Assessment: A Comparison Of Academic And Clinical Measures

Emily Ann Stephens

READING COMPREHENSION ASSESSMENT:
A COMPARISON OF ACADEMIC AND CLINICAL MEASURES

by

Emily A. Stephens
Bachelor of Arts, University of North Dakota, 2012

A Thesis

Submitted to the Graduate Faculty

of the

University of North Dakota

in partial fulfillment of the requirements

for the degree of

Master of Science

Grand Forks, North Dakota
May
2014

This thesis, submitted by Emily Stephens in partial fulfillment of the requirements for the Degree of Master of Science from the University of North Dakota, has been read by the Faculty Advisory Committee under whom the work has been done and is hereby approved.

_____
Dr. Sarah Robinson, Chairperson

_____
Dr. Manish Rami, Committee Member

_____
Dr. John Madden, Committee Member

This thesis meets the standards for appearance, conforms to the style and format requirements of the Graduate School of the University of North Dakota, and is hereby approved.

_____
Dr. Wayne Swisher
Dean of the School of Graduate School

_____
Date

PERMISSION

Title	Reading Comprehension Assessment: A Comparison of Academic and Clinical Measures
Department	Speech Language Pathology
Degree	Master of Science

In presenting this thesis in partial fulfillment of the requirements for a graduate degree from the University of North Dakota, I agree that the library of this University shall make it freely available for inspection. I further agree that permission for extensive copying for scholarly purposes may be granted by the professor who supervised my thesis work or, in her absence, by the chairperson of the department or the dean of the Graduate School. It is understood that any copying or publication or other use of this thesis or part thereof for financial gain shall not be allowed without my written permission. It is also understood that due recognition shall be given to me and the University of North Dakota in any scholarly use which may be made of any material in my thesis.

Emily Stephens
5/01/2014

TABLE OF CONTENTS

LIST OF FIGURES

Figure Page

LIST OF TABLES

Table Page

ACKNOWLEDGEMENTS

I would like to thank my committee chairperson, Dr. Sarah Robinson, for inspiring me to become a researcher. Because of your guidance, I have become a better writer and a more responsible consumer of research. I would also like to thank my committee members Dr. Manish Rami and Dr. John Madden for their feedback in the research proposal and drafting process. A special thanks should be given to the teachers in the Grand Forks public school district who were instrumental in the recruitment process and to the families and children who participated in this study.

To my fellow classmates and thesis confidants, Lindsay Anderson and Ellie Lindsey; thank you for your endless support and words of encouragement when I thought I would never finish writing. You made this experience much more enjoyable. Thank you for encouraging me to present my research at the 2013 ASHA National Convention and for road-tripping 24 hours roundtrip to get there!

Thank you to my parents and brother for being (or acting) interested when I called to talk about my thesis and related things in the field of speech-language pathology. I could not have been blessed with a more fun and supportive family. I owe all of my successes to you.

Finally, thank you to my husband, Michael, for carrying my emotional burdens through high school, undergraduate, and graduate school. I appreciate you learning about my career and listening to my countless stories. You were, and continue to be, the best part about my day.

ABSTRACT

The Measures of Academic Progress (MAP) test is a computer-based academic assessment of reading comprehension that is used by school districts nationwide. The MAP has established reliability, but its validity is not widely available. The purpose of this study was to seek validity for the MAP by comparing it to the Gray Oral Reading Test—5th edition (GORT-5), a frequently used and valid clinical test of reading comprehension. Participants included twenty-eight fourth grade students ranging from 9;7 to 10;8 years of age, with a mean of 10;2. Each participant completed the GORT-5 and MAP test. The GORT-5 and MAP each yield two reading scores: one that exclusively measures reading comprehension and another that offers an overall score of reading ability, which considers skills such as vocabulary, phonological awareness, reading speed, etc. Results indicated a fair-to-moderate correlation (r=.39) between comprehension scores, and a strong correlation (r=.65) between overall scores of reading. Several theories are discussed as to why the comprehension scores were not strongly correlated.

CHAPTER I

INTRODUCTION

Reading comprehension is defined as the ability to extract meaning from written text. Skilled reading comprehension is dependent on the ability to read words quickly, accurately, and effortlessly so that an individual's processing capacity can attend to the task of assigning meaning to the message (Adams, 2004). Reading comprehension is a basic skill necessary for educational success. Therefore, emphasis is placed on developing students' reading comprehension skills throughout elementary and secondary education.

In the early elementary years, reading instruction focuses on developing decoding skills. In the later years, however, students are taught to develop their reading comprehension skills. In other words, students transition from "learning to read" to "reading to learn". This transition period occurs about fourth grade. During this time, many students are identified as poor readers due to the increased demands of reading. This phenomenon is known as the "fourth-grade slump." At this age, students are challenged to understand a more extensive vocabulary, a heavier content load, and to draw from background information to comprehend meaning. Additionally, students transition from reading story books to expository texts (Chall & Jacobs, 2003). Without skilled reading comprehension, all other educational tasks become difficult. Because of this, students' reading comprehension skills should be carefully monitored and assessed as the demands of reading increase.

Educators routinely assess students' reading comprehension skills. The goal of assessment is to make inferences about students' reading achievement. It is a systematic process of defining, selecting, designing, collecting, analyzing, interpreting, and using information to aid students' educational development (Erwin, 1991). The ongoing task of educators is to refine the assessment process to be more efficient, reliable, and valid.

Two primary categories of assessment tools exist for students: academic tools and clinical tools. Academic tools include assessments created or administered by general education faculty to all students, such as classroom tests and district-wide standardized tests. In Grand Forks, North Dakota, students' reading comprehension abilities are measured by the *Measures of Academic Progress* (MAP) test, as well as a state-wide assessment. Clinical tools are assessment measures typically created or administered by special education personnel to diagnosis disabilities and determine eligibility for special education programs. One commonly used clinical measure of reading comprehension is the *Gray Oral Reading Tests—5th Edition* (GORT-5). Only students referred for specialized testing complete clinical tests of reading comprehension.

The validity of MAP testing has primarily been studied by the Kingsbury Center, a research division of the parent company, Northwest Evaluation Association. Their data showed that the MAP scores for eleventh grade students predicted their pass/fail status on the reading portion of the North Dakota State-Wide Assessment with 77.33% accuracy (The Kingsbury Center, 2010). There are no data published as to the overall correlation between student performance on the two tests or data to illustrate validity at different age/grade levels. Therefore, MAP testing is routinely used, despite validity information not being widely

available. On the other hand, clinical assessment measures have strong, established validity but are rarely used.

On the basis of establishing validity for the MAP, the current study aims to compare it to a widely used and valid clinical test of reading comprehension, the GORT-5. A comparison of student performance on these tests would question or lend support to the validity of the MAP. It is hypothesized that there will be a strong correlation between MAP and GORT-5 results, because the GORT-5 has been strongly correlated to other measures of reading comprehension. A strong correlation would provide validity for the MAP and may potentially reduce the number of clinical tests administered by special educators by allowing special educators to refer to MAP reading scores, when considering qualifying fourth grade students for special education programs. This would be an optimal finding because administering fewer tests means reduced workload for educators and less time that students have to spend outside of the general education classroom.

CHAPTER II

REVIEW OF LITERATURE

Teaching students to comprehend written text is an ongoing goal of educators, particularly at the elementary level. Reading comprehension is important because nearly all other educational tasks rely on the ability to understand expository text. Reading deficits can cause generalized academic difficulties which are often the root of economic and social hardship. Therefore, it is critical that educators continually teach and monitor students' reading ability.

Reading comprehension occurs when meaning is extracted from written text. To do this, individuals need to read words quickly, accurately, and effortlessly so that enough processing capacity remains to assign meaning to the message (Adams, 2004). According to the *Simple View of Reading* (Gough & Tunmer, 1986; Hoover & Gough, 1990), reading = decoding x linguistic comprehension (i.e. R = D x C). Decoding skills, or the ability to recognize words in print, are dependent on code-related skills such as phonological awareness, knowledge of graphemes, phoneme-grapheme relationship, print conventions, and beginning writing skills. Some words have to be "sounded out" or decoded while other words break phonological rules and must be recognized by sight. Linguistic comprehension includes semantics, syntax, discourse, pragmatics, and concepts (Storch & Whitehurst, 2002). It has largely been interpreted as listening comprehension or the ability to understand language (Catts & Kamhi, 2005).

The Simple View of Reading theory suggests that the ability to decode words and to understand spoken language are both equally critical to reading comprehension. Therefore, the theory assumes a product model, rather than additive because one cannot read without either skill. To a degree, a strength in one area can compensate for the other and still result in reading comprehension, however, a significant deficit in either area will drastically reduce overall comprehension.

Although the Simple View of Reading was not intended to serve as an instructional model, its clear-cut definition identifies two teachable skills, decoding and linguistic comprehension, worthy of classroom instruction. Pressley (2006) observed many classrooms around the United States and found that, unlike the emphasis placed on decoding instruction, comprehension skills were rarely, if ever, explicitly taught in schools. Additionally, too many children with good decoding skills and poor comprehension skills were placed in phonics programs when they actually needed formal instruction in comprehension. Therefore, educators need sensitive reading assessments to determine if a child's primary deficit is decoding or comprehension. Without reliable and valid assessment measures, educators are unable to appropriately address student's needs.

Some researchers have found that decoding skills are the greatest predictor of reading comprehension in the early years of literacy but that in later years, broader language skills (linguistic comprehension) gain predictive power in determining a child's reading comprehension (Storch & Whitehurst, 2002; Whitehurst & Lonigan, 1998). For students whose decoding ability supersedes linguistic comprehension, reading may not be difficult until about fourth grade, when they are challenged to understand a more extensive

vocabulary, a heavier content load, and utilize background information. At this time, expository text is introduced and story books become infrequent (Chall & Jacobs, 2003). In other words, text comprehension takes precedence over the ability to simply decode text. The term, "fourth-grade slump" describes the phenomenon for students to be identified as poor readers during this time. The importance of reliable and valid reading assessments heightens at this age so that struggling students can be identified and given the appropriate instructional supports. The ultimate goal of academic assessments is to accurately identify students' abilities so that their needs can be addressed, which is crucial to narrowing the achievement gap between students.

Before educators can formulate and implement appropriate reading instruction, they need access to sensitive assessments that accurately depict students' current reading abilities and perhaps, show progress over time. Two primary categories of assessment tools exist: academic tools and clinical tools. Academic tools include assessments created or administered by general education faculty to all students, such as classroom tests and district-wide standardized tests. Academic tests are often used to measure progress against state or federal benchmarks or school curriculum. In some cases, results from academic measures influence state funding that school's receive. Clinical tools, on the other hand, are assessment measures typically created or administered by special education personnel to diagnosis disabilities and determine eligibility for special education programs. Only students referred for specialized testing complete clinical tests of reading comprehension.

Reading comprehension provides the foundation for academic achievement across all content areas. As students progress through the education system, they are expected to

achieve greater levels of academic independence by synthesizing information from a variety of sources including textbooks, articles, etc. This skill must be performed at a high level to fulfill post-secondary education requirements. High-school seniors' performance on the Scholastic Aptitude Test (SAT), an assessment used by college admissions' boards to predict college readiness, has shown a steady decline in reading scores since 1972. Performance in 2013 marked the nation's lowest reading scores in forty years. According to the 2013 SAT Total Group Profile Report, the national reading average is down 34 points since 1972, whereas math scores have observed yearly fluctuations + or - 28 points and currently stand five points above the 1972 national average (College Board, 2013).

The 2013 SAT Report on College & Career Readiness states that only 43% of students in the class of 2013 met the benchmark for college readiness, which is noted by a composite score of 1550 for the areas of math, reading, and writing. This percentage has remained relatively unchanged in the last five years, which perpetuates the need for educators to adjust their teaching methods so that more students graduate high school with the skills necessary to achieve post-secondary education success (College Board, 2013). The SAT benchmark of 1550 is the score in which students have a 65% (higher or lower) probability of achieving a first year GPA of at least 2.67 (B-) (College Board, 2011). Notably, students who achieve this benchmark are more likely to enroll in a four-year university, more likely to complete their degree within four years, and more likely to earn a higher first-year GPA than their peers who did not meet the benchmark (College Board, 2010).

The SAT is one of many nation-wide standardized tests available to assess the academic skills of our students. Another test that provides longitudinal data is the National

Assessment of Educational Progress (NAEP), which provides a snapshot of the nation's academic achievement in a variety of content areas, including reading. Schools around the country are selected at random to take the assessment. There are two versions of the test: the main assessment, administered every two years, and the long-term trend assessment, administered every four years. The reading assessment results are reported at the following achievement levels: basic, proficient, and advanced. In 2013, 32% of fourth-grade students and 22% of eighth-grade students scored below the basic reading level for their grade (National Center for Education Statistics, 2014).

Contrary to the SAT reading trends, NAEP long-term reading scores have gradually increased since the first reading test was administered in 1971. Consistent with the SAT, however, students continue to perform better in the area of mathematics than they do with reading. Since the first NAEP mathematics test was administered in 1973, the nine-year-old national average has increased by 25 points, thirteen-year-olds have increased 19 points and there is no statistical difference for the scores of 17-year-olds. Compare that to the reading test; students have made half the progress, noting a 13-point increase for nine-year-olds, an 8-point increase for 13-year-olds, and no statistical difference for 17-year-olds (National Center for Education Statistics, 2013).

In Grand Forks, North Dakota, students' reading comprehension abilities are measured by the *Measures of Academic Progress* (MAP) test, an academic assessment tool created by the Northwest Evaluation Association (NWEA). The MAP test is a state-aligned computerized assessment that compares student achievement to expected grade level

outcomes and measures progress over time. The MAP is designed to identify student's weaker skills and content areas so that teachers can reinforce these items in the classroom.

Students in kindergarten through eleventh grade can complete MAP testing four times per year in the areas of math, science, reading, and language. Students in the Grand Forks, ND public school system take the MAP test twice per year, in September and May. A unique feature of the MAP test is its scoring system, which allows student progress to be tracked on a single scale as they progress through the education system. For instance, the mean reading score for students at the beginning of their kindergarten year was 142.5. This number increases until the mean reading score reached 223.7 at the end of students' 11th grade year (NWEA, 2011). This scoring system allows parents and educators to compare students' scores over time, and to compare students' performance to that of their peers.

Raw scores are converted to a standard score (RIT score, based on an equal interval scale) and displayed in a chart that denotes what scores are considered to be at, below, or above grade-level for all content areas at each testing period throughout the year. An advantage of the RIT system is that the scores are tied directly into the curriculum, rather than being based on the performance of specific groups of students. Since test information is anchored in the curriculum, it is possible to track student progress accurately from year to year, to change the tests to keep pace with the curriculum, and to maintain consistent norms.

One commonly used clinical measure of reading comprehension is the *Gray Oral Reading Tests—5th Edition* (GORT-5). The GORT-5 measures oral reading fluency and comprehension. It is a popular among researchers (Coulter, 2004; Grunau, Whitfield & Davis, 2002; Pham, Fine & Semrud-Clikeman, 2011; Srivastava & Gray, 2012;) and special

education personnel because of its impressive reliability, validity, and ease of use. Designed for students ages 6-23, the GORT-5 consists of 16 developmentally sequenced stories. Each story contains five open-ended comprehension questions which requires students to recall main ideas, supporting ideas, and make inferences about the text (Weiderholt & Bryant, 2012). The GORT-5 has two equivalent forms that have an alternate-form reliability of greater than .90. Test-retest reliability, using same and alternate form, is greater than .85. Internal consistently reliability is greater than .90, sensitivity = .82, and specificity = .86.

Reading ability can be assessed silently or orally, both of which have their advantages. Silent reading assessments, such as those conducted in the MAP, are used to assess multiple individuals at once. Because of this, silent reading assessments are typically more time and cost-effective. Oral reading assessments, however, allow the clinician to determine specific deviations made by the student during reading. This gives way to understanding the student's decoding skills and whole-word recognition of irregular words. It also allows the clinician to assess oral fluency, a combined measure of reading speed and accuracy, and prosodic features (Wiederholt & Bryant, 2012).

In response to increasing technological expansion, many academic assessments have become computer-based rather than paper-based. Computer-based assessments allow for easier and quicker scoring and can be adapted to student needs (e.g. MAP questions become more or less challenging throughout the test, based on student performance). Little research has been conducted on how this shift might impact student performance in reading comprehension measures. Liu (2005) found that when reading on screen, people have poorer

sustained attention and spend less time reading text with in-depth concentration. Instead, there is higher incidence of scanning, reading selectively, and looking for keywords.

Modern textbooks and computer resources often follow a non-linear pattern, meaning readers have the ability to control the sequence in which they acquire information. In modern textbooks, there is a main body of text with accompanying dialogue boxes for definitions, key points, charts, graphs, pictures/captions, etc. In non-linear computer resources, such additional information is embedded throughout the text via hyperlinks, tabs, expanding and retracting text boxes, etc. The difference in computer text is that the additional information can appear and disappear by viewing alternate screens, pop-up boxes, and expanding and retracting text boxes whereas non-linear paper text remains a permanent display.

Non-linear presentation allows readers more independence which helps them to acquire information in a manner that fits their needs. Having more freedom however, can impose a greater cognitive load on readers from having to sustain a mental representation of their reading sequence while they navigate from one text area to another (Lee & Tedder, 2003; Pazzaglia, Toso, & Cacciamani, 2008), particularly for computer text because the reader has to toggle between screens. Some researchers have found reading comprehension to be adversely affected by the increased cognitive demands associated with reading non-linear text (Barab, Young, & Wang, 1999; Beishuizen, Stoutjesdijk, & Zanting, 1996).

Srivastava and Gray (2012) analyzed non-linear reading comprehension in eighth-grade students; 14 with a language-learning disability (LLD) and 25 with typical language development (TLD). Reading passages, and coordinating comprehension questions, were adopted from grade-level state assessments from Florida, Texas, and Washington. Each

participant read one paper passage and one computer passage. Reading passages were embedded with topic-specific vocabulary words. Within paper passages, vocabulary words were bolded and italicized in the body of the text and defined in nearby text boxes. Within computer passages, vocabulary words were made into hyperlinks throughout the body of the text which, when clicking, would open a pop-up box that contained the definition.

Participants were instructed to read the comprehension questions before reading the passage. They were allowed to refer back to the questions as they were reading, but could not use the passage after beginning to answer the questions.

To determine how various skilled contributed to reading comprehension, tests of oral language (Clinical Evaluation of Language Fundamentals-4th Edition), decoding (GORT-4), and working memory (Visual Spatial Span Board task from the Wechsler Memory Scale, 3rd edition) were administered to the students prior to the reading comprehension task. Srivastava and Gray (2012) found that these factors contributed differently to reading comprehension scores per group and condition. For the TLD group, vocabulary and phonological working memory predicted reading comprehension in computer-based task, but not visuospatial working memory. In the paper-based task, vocabulary was the largest predicting factor. Decoding was not a predicting factor in either case, perhaps because the passages were written at grade-level and the participants with TLD did not have reading deficits.

Comparatively, for the LLD group, decoding was a predictor for reading comprehension in the computer-based task, most likely due to the fact that the decoding abilities of the participants in this group were lower than the level at which the passages were

written. For the paper-based task, phonological working memory, visuospatial working memory, and semantic knowledge predicted reading comprehension.

Overall, Srivastava and Gray (2012) found that students with LLD scored significantly lower in both tasks than did students with TLD, but there were no significant differences in the amount of time that the groups spent reading and answering questions. Despite a greater cognitive load required for the hypertext in computer-based assessments, neither group seemed to be affected by this difference. Therefore, individuals scored the same on computer-based measures as they did on paper-based measures.

Ackerman and Lauterman (2012) studied the effect of time-pressure on 80 undergraduate's reading comprehension performance on paper-based and computer-based tasks. Participants were randomly assigned to a text medium group (i.e. paper or computer) and further assigned to a group that would take the test under time pressure or free regulation (i.e. no time limit). After reading the passage and before answering comprehension questions, participants made predictions of their performance (POP), based on how well they believed to have understood the passage. Results showed that participants tended to be more overconfident in their comprehension of text during time-pressured computer tasks. The researchers also collected data about participant preference for learning and found that participants who generally preferred learning via printed materials, scored significantly higher on paper-based tests than computer-based tests. For participants who generally preferred to learn text on screen, their performance was not significantly changed based on test medium. Regardless of medium preference however, scores were generally lower on time-based computer tasks than on time-based paper tasks.

To date, current research on text medium's effect on comprehension is lacking and available research has proved inconclusive. Some studies found that reading comprehension in students with TLD is higher when reading computer-based text (Blohm, 1982; L-Allier, 1980; Reinking & Schreiner, 1985) whereas other studies found that reading comprehension suffers as a result of computer-based text and hypertext (Heppner, Anderson, Farstrup, & Weiderman,1985; Macedo-Rouet, Rouet, Epstein, & Fayard (2003). Additionally, several studies determined no difference between text mediums (Srivastava & Gray, 2012; Fish & Feldman, 1987; Gambrell, Bradley, & McLaughlin, 1985). It is important to consider how text medium could influence student performance. At this time, researchers have no reason to believe that text medium would be the sole cause for differing outcomes in student performance on these assessments.

The current study aims to compare results from two reading comprehension measures: the MAP, a computer-based academic assessment, and the GORT-5, a paper-based clinical assessment. Clinical tests undergo rigorous research to establish high reliability and validity (see previous data for GORT-5). Clinical tests are used to assess students who are referred for special education testing. Generally, the most reliable and valid tests are administered the most often, however, even the most reliable and valid clinical tests are administered far less frequently than academic tests such as the MAP, SAT, NAEP, etc., which are administered to large bodies of students (e.g. every student in a district, every graduating senior, etc.)

While academic tests are frequently used, reliability and validity information is not well known. Per host company websites and public documentation issued by the host

companies, descriptive information is given for how the tests are developed and analyzed for reliability and validity, however, quantitative data (i.e. coefficients) are often not provided, at least as public information. According to an alignment study by the Kingsbury Center (2010), the reading portion of the MAP predicted eleventh grade students' pass-fail status on the reading portion of the North Dakota State Assessment (NDSA) with 77.33% accuracy. The Kingsbury Center is a research division of the MAP's parent company, Northwest Evaluation Association and to-date, is the only person/organization to have researched MAP validity. Their research failed to illustrate validity at different grade levels, despite the fact that the NDSA is administered to students in grades 3-12 or provide an overall correlation between student performance on the two tests.

The question exists: Do students score similarly on academic and clinical tests of reading comprehension? Specifically, for fourth-grade students in the Grand Forks, North Dakota public school system:

1) What is the relationship between scores obtained on the GORT-5 and the reading portion of the MAP?

A strong correlation between the GORT-5 and the MAP would provide validity for the MAP and may potentially allow special educators to reduce the number of clinical tests administered, by referring to MAP scores, when considering qualifying fourth grade students for special education programs.

CHAPTER III

METHOD

Participants

Twenty-eight participants (18 males, 10 females) were recruited from fourth grade classrooms in the Grand Forks, North Dakota Public School district. Seven of 13 schools were represented. Participants ranged from 9;7 to 10;8 years of age, with a mean of 10;2. All participants were native English speakers. Individuals were neither included nor excluded based on socio-economic status. Because the focus of the study was on reading comprehension abilities in typically developing children or those with a language-based impairment, individuals who have been diagnosed with Autism Spectrum Disorder, cognitive deficits, or hearing impairments were excluded from the study. Individuals with language disorders, however, were included. All participants received $20 cash for their participation.

Materials

Measures of Academic Progress

The Measures of Academic Progress (MAP) is a computer-based assessment that compares student achievement to expected grade level outcomes and measures progress over time. Students in the Grand Forks, ND public school district take the math and reading portions of the MAP in September and May of each school year. For the purpose of this study, the reading portion of the MAP will be examined. The reading score includes subtests which assess comprehension in a variety of fictional and nonfictional texts, including poetry,

short stories, recipes, Venn diagrams, and maps with written directions (S. Robinson, personal communication, March 20, 2013).

The MAP is a dynamic assessment that adjusts to student responses throughout the test, therefore presenting each student with a unique set of questions. This adaptive method enabled students to be assessed at their exact skill level (NWEA, n.d.). Each question was assigned a specific point value. A student's raw score was calculated by adding these points. To compare students' scores over time, and to compare students' performance to that of their peers, raw scores were converted to a standard score, or RIT score. The RIT score is based on an equal interval scale. The 2011 MAP Normative Data reports the national mean RIT reading score to be 206.7 at the end of students' fourth grade year, therefore, if a second-grader had a mean reading RIT score of 207, he/she would be performing at a fourth-grade level (NWEA, 2011). For the purpose of this study, RIT reading scores were collected.

The MAP was administered in a quiet environment to one classroom of students at a time. Students were assigned to specific seats with their own Apple Notebook computer. They were instructed to remain quiet during the testing period and were allowed to read silently in the chairs upon completing the test. Throughout the estimated 40-minute assessment, their classroom teacher monitored the students in the room.

MAP scores were organized in performance charts and made available to the students' schools for distribution to classroom teachers and parents. The research team obtained the participants' MAP scores from school administration after providing the administration with the names of participants and parental consent forms.

Gray Oral Reading Tests—Fifth Edition

To assess participants' reading ability, the *Gray Oral Reading Tests—Fifth Edition* (GORT-5) was administered. The GORT-5 is an individually administered test of oral reading skills and comprehension. It is norm-referenced for individuals between the ages of 6-23 years. Materials include the Student Test Book, Examiner Manual, Examiner Record Booklet, and a stop watch. The GORT-5 consists of 16 developmentally sequenced stories. Each story contains five open-ended comprehension questions which required the participant to recall main ideas, supporting ideas, and make inferences about the text (Weiderholt & Bryant, 2012).

During administration, the participant was asked to read a short story aloud while the examiner timed the participant's reading speed and marked any reading errors in the Examiner Record Booklet. Errors included disfluencies, mispronunciations, additions, and omissions of words. After the participant was finished reading, the examiner removed the story from the participant's view and verbally asked him/her five open-ended comprehension questions. Each answer received a score of 1 for correct or 0 for incorrect. These scores were summed to obtain a comprehension score out of five possible points. Per the Examiner Manual, the entry point for fourth grade students is story number two.

After each story, using a provided conversion chart, the examiner converted the participant's reading speed to a Rate score, on a scale of 1-5. Likewise, an Accuracy score was calculated by summing the number of reading errors and converting it to a scale of 1-5. The Rate and Accuracy scores were then summed to obtain a total Fluency score out of 10 possible points. The Fluency score was used to establish the basal and ceiling levels for the

test. The basal was marked by a score of 9 or 10 on two consecutive stories. If the participant failed to achieve a basal on the first two stories, the researcher administered lower-level stories until a basal was achieved. The participant progressed through the test by reading and answering questions until a ceiling was obtained by achieving a Fluency score of 2 or less for two consecutive stories.

The following raw scores were summed from each story: Comprehension Score, Rate Score, Accuracy Score, and Fluency Score. These scores were converted to scaled scores and percentile ranks for each category. The student's Fluency and Comprehension scaled scores were summed to determine the student's Oral Reading Index, which has a mean of 100 and a standard deviation of 15. Of all the scores obtained on the GORT-5, the ORI has the greatest reliability, therefore making it the most clinically useful assessment measure. Students who have an ORI score of 90 or above have reached an oral reading level expected for their age.

Procedure

Fourth grade students were recruited through flyers distributed to classrooms and from an advertisement published in *Kids Connections*, a monthly newsletter sent to all parents/guardians of students in the GFPS district. The advertisement contained the purpose of study, methodology, compensation details, and instructions on who to contact if interested (see Appendix A).

Interested parents/guardians were instructed to contact the primary investigator via phone or email. The study was explained further and the parents/guardians were given an opportunity to ask questions. This initial contact also served as a screening to determine participant eligibility (e.g. age, native language, any existing medical or educational

diagnosis). A member of the research team later contacted the parents/guardians to schedule a time for participant testing. Research was conducted at one of two locations, at the participant's school, either before or after school hours or during weekend or evening hours on the University of North Dakota campus.

A research team consisting of three graduate assistants administered the research protocol to all participants. Prior to the participant's arrival on site, the researcher set up the materials necessary to carry out the assessment by arranging the tests, manuals, informed consent form, writing utensils, and a video recorder in a quiet room with minimal distractions.

At the beginning of each testing session, the researcher obtained the parent/guardian's signature on a consent form (see Appendix B) and the participant's signature on an assent form (see Appendix C). Through the assent form, the purpose of the study was explained and the participant was assured that he/she did not need to participate in the study and could cease participation at any time. The participants were encouraged to do their best and to expect that some questions would be easy and some would be difficult. As needed, the participant could take breaks. Parents/guardians were given the option to stay in the testing room, or a nearby waiting area, whatever the participant was most comfortable with. Each testing session was video recorded for the purpose of obtaining inter-rater reliability.

The testing protocol was a part of a larger research study and consisted of the Gray Oral Reading Tests—5th ed. (GORT-5), the Peabody Picture Vocabulary Test—4th ed. (PPVT-4), the Expressive One-Word Vocabulary Test—2 ed. (EVT-2), and a researcher-designed vocabulary assessment based on the student's current weekly spelling list. The

participants completed MAP testing at their school, as part of a district-wide requirement. The testing protocol was administered within one month of the completion of the MAP test. The order of test administration was counterbalanced to control for any order effects according to a pre-determined schedule. The testing session took about one hour to complete. For a complete description of GORT-5 test administration and scoring, refer to the materials section.

The administered tests were scored online according to the procedures in their test manuals. The data was entered into a password-protected spreadsheet, kept on the primary investigator's computer, and later transferred into the Statistical Package for the Social Sciences (SPSS) program for analysis.

Videos and test protocols were coded with a subject number to ensure participant privacy. All hand-written data sheets, test protocols, and videos (on a flash drive) were stored in a locked file cabinet and kept separate from the consent forms. All research materials will be kept for a period of three years before being destroyed according to University of North Dakota policy. The primary investigator and the members of the IRB audit team will be the only individuals with access to the filing cabinet.

Data Analysis

Two types of statistical procedures were used to analyze the data. Descriptive statistics, including mean, median, standard deviation, and range were collected for the MAP and GORT-5 scores. A correlation coefficient analysis (Pearson-r) was conducted to determine the degree of relationship between the MAP and GORT-5. The data are presented in tables and graphs as well as narrative form in chapter four.

CHAPTER IV

RESULTS

Twenty-eight fourth grade students participated in the study. They were recruited from seven of 13 public schools in Grand Forks, ND. Two reading tests were administered to each participant, the MAP and the GORT-5.

Descriptive statistics (range, mean, and standard deviation) of test scores are presented in Table 1. Each test yields two reading scores. The GORT-C (comprehension) and MAP-C (comprehension) are exclusive measures of reading comprehension, whereas the GORT-ORI (Oral Reading Index) and MAP-O (overall) are overall scores of reading ability. The GORT-ORI includes comprehension, rate, and accuracy of oral reading. The MAP-O includes comprehension, vocabulary, phonological awareness, concepts of print, and word structure.

Table 1. Range, mean, and standard deviation of scores of fourth graders' performance on the GORT-5 and MAP tests (n = 28).

	Range	M (SD)
GORT-C	4-14	8.71 (2.75)
MAP-C	196-250	217.75 (11.67)
GORT-ORI	81-118	96.75 (11.60)
MAP-O	196-245	216.00 (10.82)

The GORT-C score was calculated by adding the number of comprehension questions that were answered correctly. This number was converted to a scaled score with a possible

range of 1 to 20. Scores falling within 8 to 12 are considered to be in the normal range for fourth grade students. Ten participants fell within the normal limits. Four participants scored above average. The participants' GORT-C scores ranged from 4-14.

The GORT-ORI raw score is calculated by adding the sub scores of comprehension, rate, and accuracy then converting to a scaled score with a possible range of 52 to 150. Scores falling within 90 to 110 are considered to be in the normal range for fourth grade students. 15 participants fell within the normal limits. Four participants scored above average. The participants' GORT-ORI scores ranged from 81-118.

The raw scores of the MAP-O and MAP-C were converted to RIT scores, which lie on an equal interval scale. These scores are independent of grade level, therefore allowing for the ability to track progress from year to year. According to the 2011 MAP Normative Data, the mean RIT reading score for students finishing their fourth grade year is 206.7. Scores of 198-207 are considered to be typical for fourth grade students (NWEA, 2011). The range of scores for participants on the MAP-O was 196 to 245 with a mean of 216. The range of MAP-C scores was 196 to 250 with a mean of 217.75.

Descriptive statistics (range, mean, and standard deviation) for the amount of time that the participants took to complete the reading portion of the MAP are displayed in Table 2. Participants spent between 20-54 minutes completing the reading portion of the MAP with a mean of 32.43 minutes.

Table 2. Range, mean, and standard deviation of the amount of time it took fourth graders to complete the MAP (n=28).

	Range	M (SD)
MAP Duration	20-54	32.43 (8.95)

Four correlational analyses were performed using Pearson r to determine the strength of the relationship between test scores. The first analysis compared the reading comprehension scores of students on the two tests. For this analysis, scores on the GORT-C were compared to the MAP-C. This analysis yielded a fair to moderate correlation of $r = .39$, $r^2 = .15$ at the $p < 0.05$ level. See Figure 1.

The second analysis compared students' overall reading abilities. For this analysis, standardized scores on the GORT-ORI were compared to the MAP-O. This analysis yielded a strong correlation of $r = .65$, $r^2 = .42$ at the $p < 0.01$ level. See Figure 2.

The third and fourth analyses compared participant performance on the MAP, to the length of time it took to complete the test (MAP Duration). Duration of GORT-5 testing was incorporated into performance outcomes via basal and ceiling levels so further analysis was not needed. See materials section for more details. The correlation between MAP Duration and MAP-C was $r = .075$, $r^2 = .006$ and the correlation between MAP Duration and MAP-O was $r = .014$, $r^2 = .0002$ demonstrating no relationship between duration and outcomes. See Figure 3 and Figure 4, respectively.

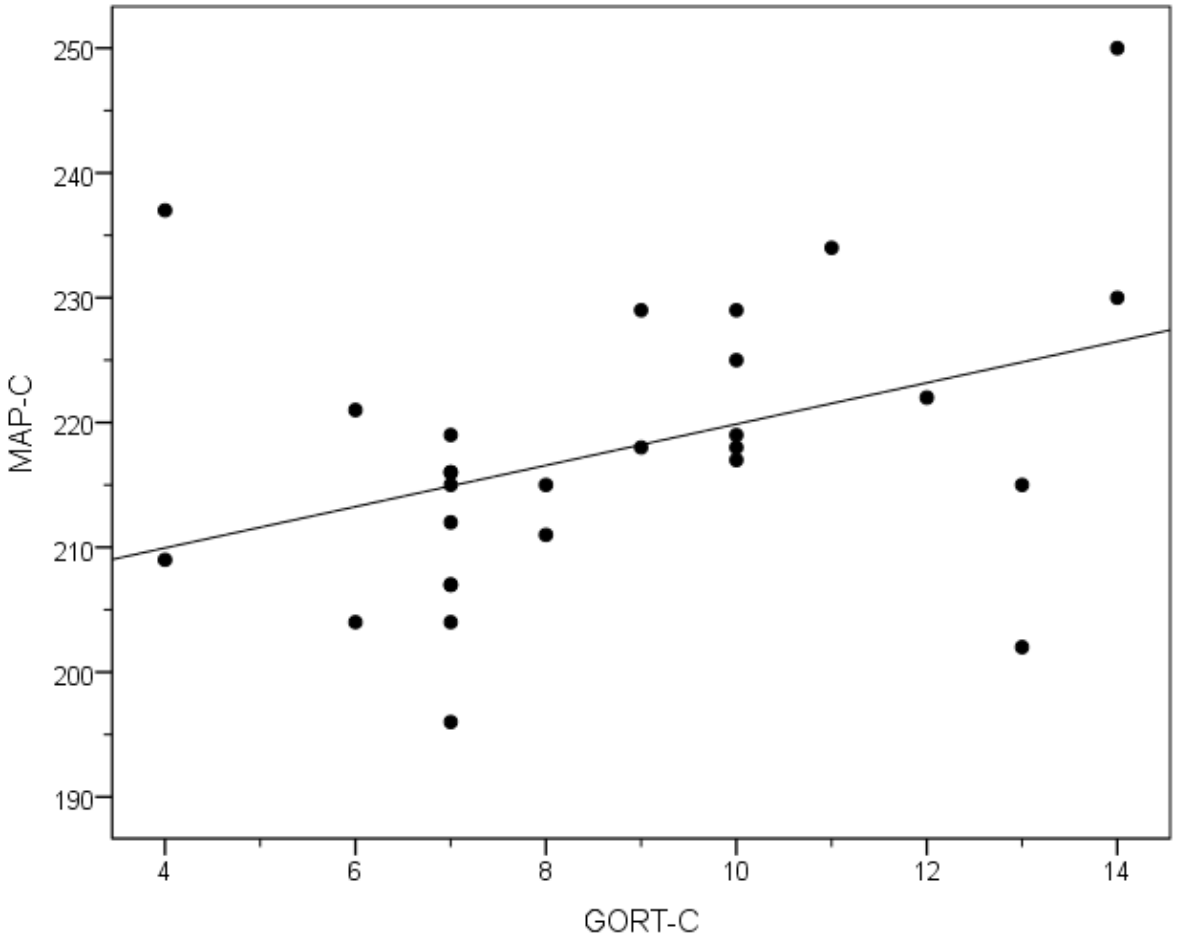

Figure 1. A scatterplot showing the relationship between GORT-C and MAP-C
$r = .39^*$, $r^2 = .15$, $p = .041$

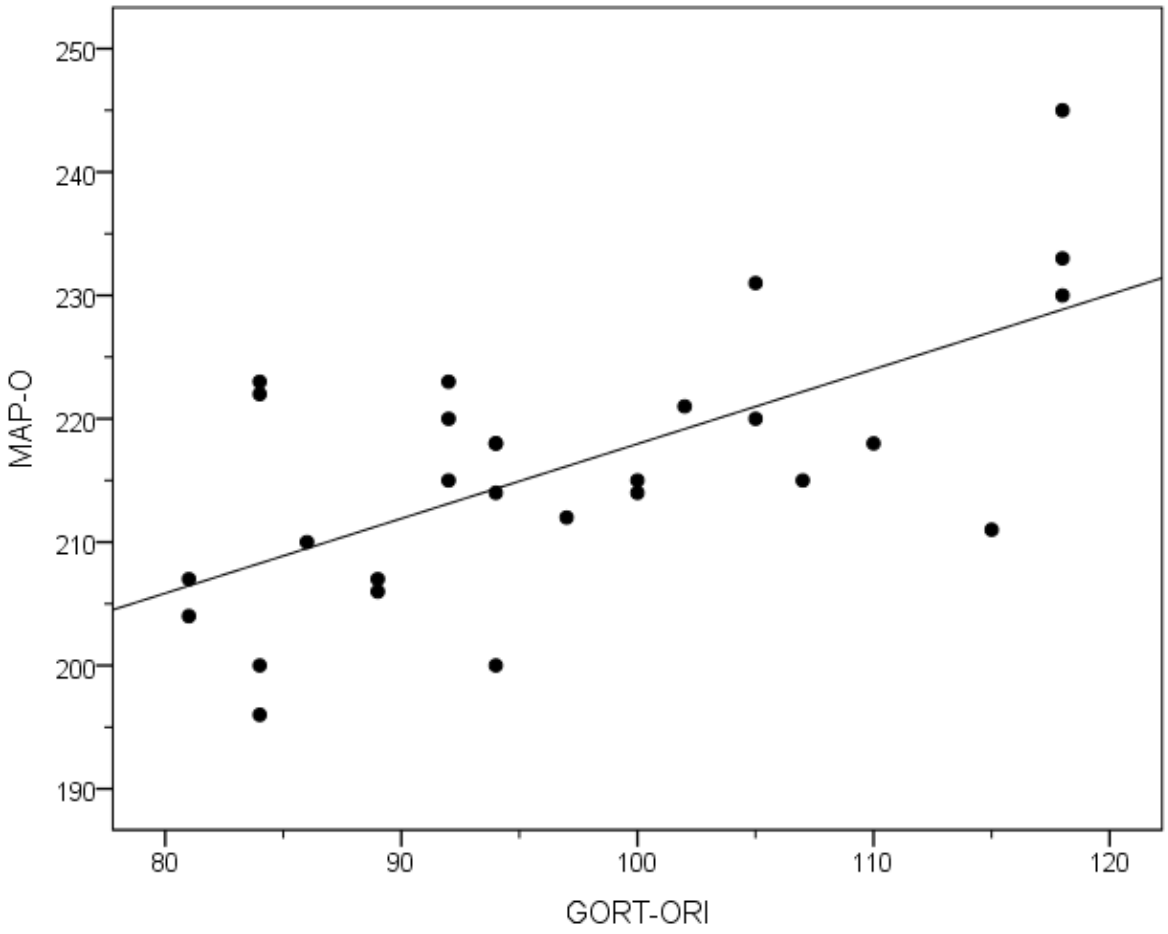

Figure 2. A scatterplot showing the relationship between GORT-ORI and MAP-O
r =.65**, r^2 = .42, p = .00

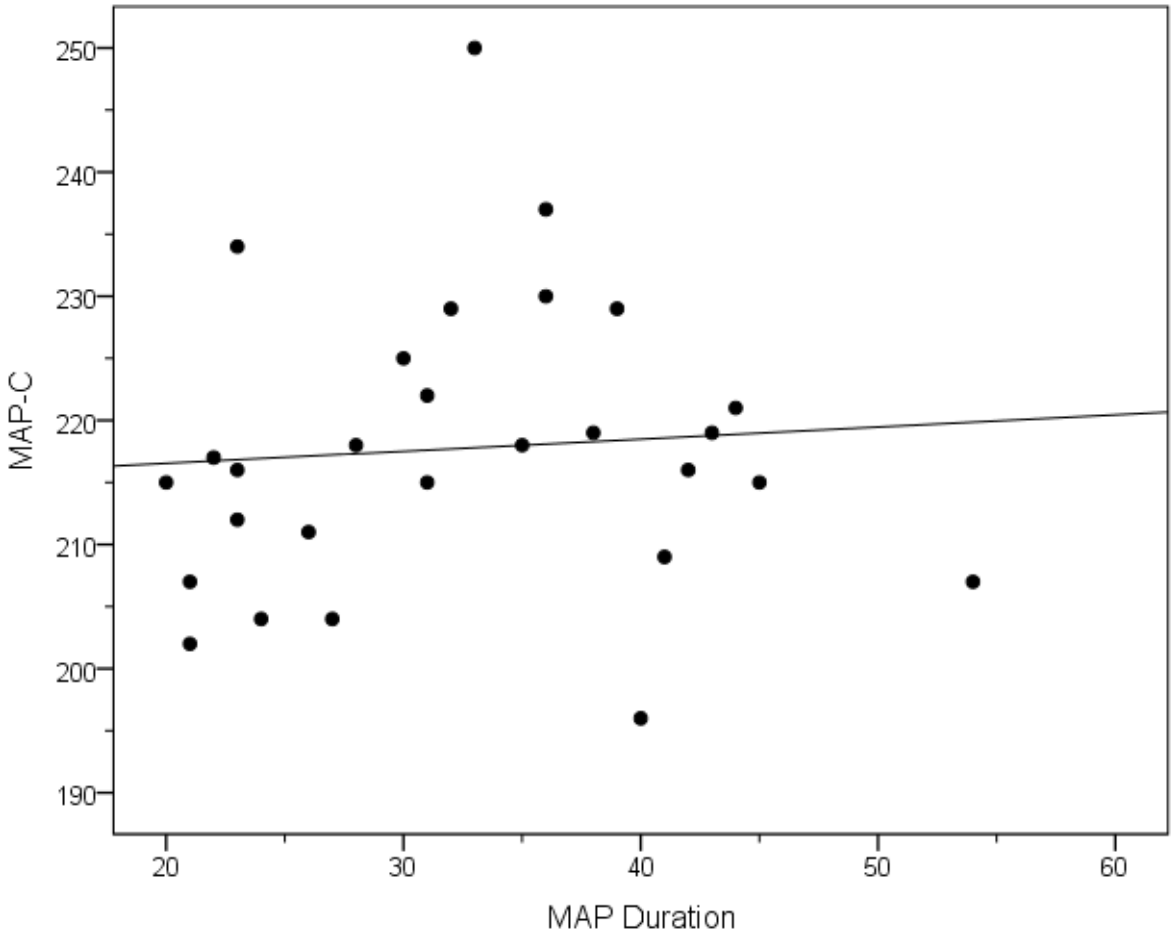

Figure 3. A scatterplot showing the relationship between MAP Duration and MAP-C
$r = .075$, $r^2 = .006$

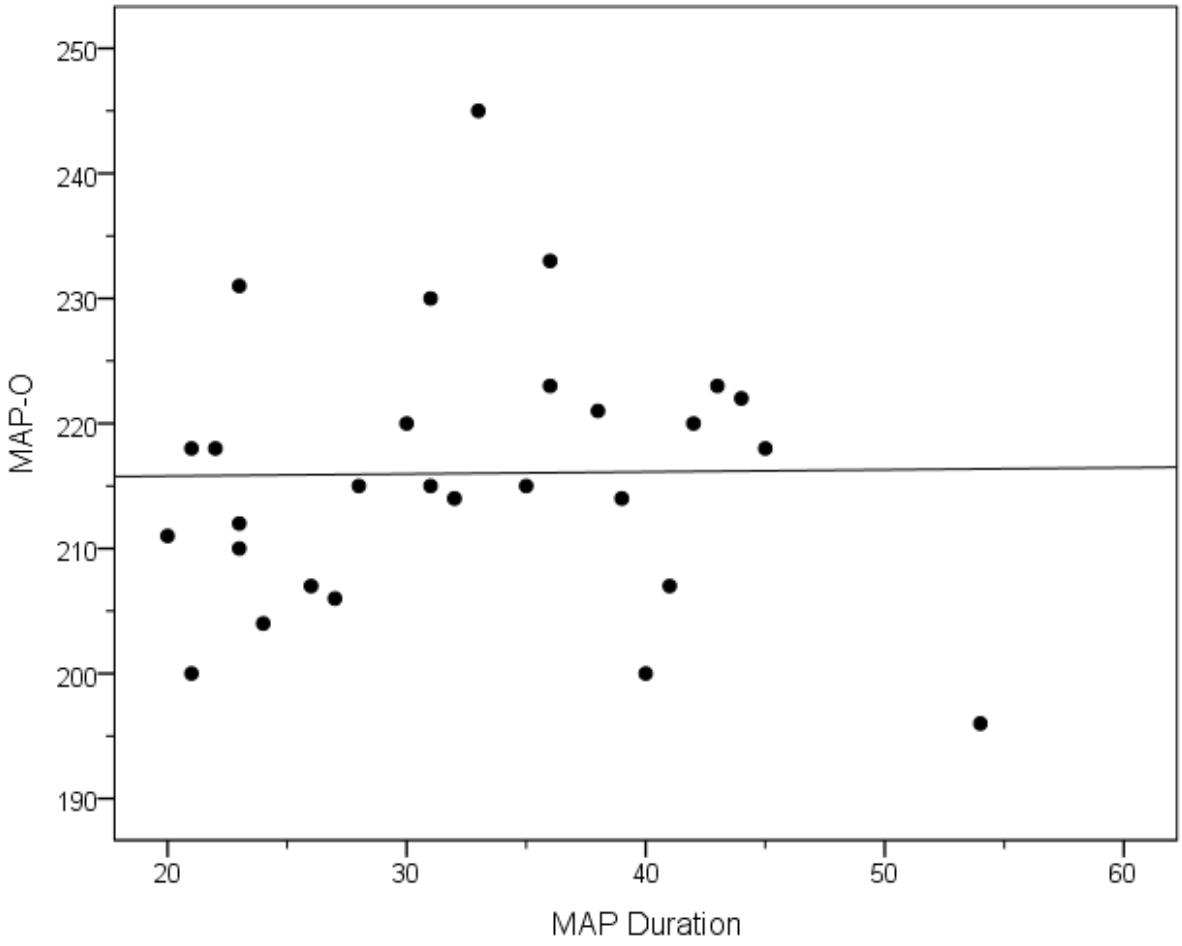

Figure 4. A scatterplot showing the relationship between MAP Duration and MAP-O
$r = .014$, $r^2 = .0002$

CHAPTER V

DISCUSSION

This study examined the relationship between scores obtained on the GORT-5, a clinical measure of reading, and the reading portion of the MAP, an academic assessment. Both tests yield two standardized scores for purposes of comparison, one that exclusively measures reading comprehension (i.e. GORT-C and MAP-C), and another that offers an overall score of reading ability (i.e. GORT-ORI and MAP-O). The GORT-5 and MAP measure reading in different ways, however. The GORT-5 is an oral reading test that measures an individual's ability comprehend text, and to read accurately and quickly. The MAP is a silent reading test that measures comprehension in a variety of reading genres. The GORT-ORI score includes comprehension, rate, and accuracy of oral reading whereas the MAP-O score includes comprehension, word structure, phonological awareness, vocabulary, and concepts of print. Therefore, the overall reading scores do not directly consider the same subset of skills. It can be argued, however, that both tests do in fact, assess the same skills because vocabulary, phonological awareness, concepts of print, etc., as measured by the MAP, are necessary to having good reading rate and accuracy, as measured by the GORT-5. Because of this, one would expect the scores obtained on the GORT-5 and MAP to be highly correlated.

As expected, the scores of overall reading ability (i.e. GORT-ORI and MAP-O) were highly correlated. It was surprising however, that the correlation between comprehension

scores (i.e. GORT-C and MAP-C) only yielded a fair to moderate relationship (Gray, 2004). To explain these correlations, we must understand the basic structural differences between the two tests and consider how such differences impact comprehension. Notable structural differences between the GORT-5 and MAP include the text medium display, types of comprehension questions, and inclusion/exclusion of oral reading fluency.

The GORT-5 is a paper-based test, whereas the MAP is computer-based. Due to increasing technological advances and technology use in education, there is a growing body of research to examine the effect of text medium (e.g. paper vs. computer) on reading comprehension. The findings of this study are consistent with Srivastava & Gray (2012) who found no difference in student performance based on text medium display. A few participants in this study showed a slight advantage for performance on the MAP, demonstrated by relatively high scores on the MAP and relatively low scores on the GORT-5, however, due to the small sample size and the few number of participants demonstrating this trend, the findings of this study indicate no significant advantage between text mediums. Therefore, differences in performance cannot be attributed to text medium display.

Another difference between the GORT-5 and the MAP is the type of comprehension questions used to assess readers' knowledge of the text. After each story of the GORT-5, participants were asked five open-ended questions which required them to recall main ideas, supporting details, and make inferences about the text. The majority of comprehension questions asked the reader to recall specific details, which can be taxing on readers' memory. For example, "What was the one word used to describe the boat?"

In contrast, the MAP also required participants to recall main ideas, supporting details, and make inferences about the text but significantly less emphasis was placed on recalling details. There was a greater balance amongst types of questions which seemed to be less taxing on memory and more reflective of actual reading comprehension. Additionally, questions were asked on a multiple-choice basis and there are obvious contextual differences between the two tests. The GORT-5 assesses comprehension of short stories only, whereas the MAP targets a variety of genres including poetry, recipes, maps, etc. Therefore, it seems that the MAP measures reading comprehension in a more comprehensive way than the GORT-5, which may have contributed to the relatively low correlation between the two tests.

Perhaps the most fundamental difference between the MAP and the GORT-5 is the inclusion of oral reading fluency in the scoring of the GORT-5, compared to silent reading comprehension in the MAP. The GORT-5 relies on oral reading fluency, which is a combined score of reading rate (i.e. speed of reading) and accuracy (i.e. number of reading errors) to determine the participants' basal and ceiling levels. Ceiling levels were obtained by achieving a Fluency score of 0, 1, or 2 (out of 10), on two consecutive passages. Almost always, students reached a point where they scored 0 out of 5 comprehension questions correctly, but still maintained a reading Fluency score high enough to warrant further testing.

For example, many participants had to continue reading higher-level passages simply to obtain a ceiling level, despite the passages being too difficult for them to understand. It was not uncommon for participants to have poor comprehension scores on passages that they were able to read with decent fluency. This demonstrates that reading fluency was not

necessarily related to reading comprehension and that GORT-5 scores may not be reflective of the participants' reading comprehension abilities.

Research supports the idea that students must demonstrate reading fluency skills that at least equate to the difficulty level of their text, otherwise, their cognitive resources are expended on decoding, and therefore cannot be utilized as readily for comprehension (Adams, 2004). However, as seen in this sample, it is possible, and perhaps even common, for students to have higher decoding abilities than comprehension abilities. In this case, a test that assumes a relationship between oral reading fluency and reading comprehension (e.g. GORT-5) may not provide the most accurate picture of student's reading comprehension.

Although the relationship between reading fluency and comprehension was not a primary research question, it was observed that these factors seemed to demonstrate an inconsistent pattern for these participants, despite popular belief that reading fluency correlates with comprehension. After these observations were made with performance on the GORT-5, the relationship between reading fluency and comprehension was further examined in participant performance on the MAP. If reading fluency is related to reading comprehension, one would expect slow/inaccurate readers to generally have low comprehension and fast/accurate readers to generally have high comprehension. In contrast to the GORT-5, the MAP does not factor reading rate or accuracy into their assessment. Individuals are given an unlimited amount of time to complete the test and each question is presented based on the correctness of the previous response. If an individual answers correctly, a harder question will be presented. If he/she answers incorrectly, an easier question is presented.

Correlational analyses were conducted to determine the relationship between MAP Duration (i.e. length of time needed to complete the test) and performance outcomes (i.e. MAP-O and MAP-C). No correlations were found, however, meaning the length of time it took students to complete the test was not related to how well they scored on the comprehension portion of the test. So, on the MAP, silent reading fluency was not related to comprehension. The mean score for participants with the three fastest times was 208. The mean score for participants with the three slowest times was 216. For participants in this study, those that read slower actually performed better.

This study did not support the general idea that reading fluency is related to comprehension, such that students can have higher reading fluency skills than comprehension skills, which poses the question: What are the assessment implications for students whose reading fluency exceeds comprehension? Or for students who demonstrate poor oral reading fluency, despite maintaining adequate silent reading fluency?

These questions raise concern for the use of GORT-5 as a diagnostic tool for reading comprehension. The GORT-5 provides valuable information to a clinician, but it may rely too heavily on reading fluency in its calculation of reading ability. On the other hand, the MAP relies strictly on students' comprehension as a means for generating further test questions. In this way, it does not matter how long it takes a student to read the text or answer questions. Subsequent questions are presented based on the correctness of the previous answer and offer easier or harder questions, thereafter. Given that reading fluency was not related to comprehension in the current study, it is understandable that the

correlation between the comprehension scores on the GORT-5 and MAP were not more strongly correlated, since the GORT-5 places so much emphasis on reading fluency.

The goal of this study was to establish validity for the MAP by attempting to a find a strong correlation between the MAP and the GORT-5, a reliable and valid clinical test. Results showed a strong correlation between the overall reading scores, and a fair-to-moderate correlation between the comprehension sub scores. Several possible explanations were discussed as to why the comprehension scores were not more highly correlated.

First, consideration was given to how differences in text medium may have impacted comprehension. It was determined that the participants in this study showed no significant advantage between text mediums and therefore, differences in performance cannot be attributed to text medium display.

Second, the type of comprehension questions asked was different per test. The GORT-5 prominently favors supporting details, whereas the MAP includes a more balanced variety of comprehension questions targeting main idea, supporting details, and inferencing.

Third, the construct of the GORT-5 assumes that oral reading fluency is related to reading comprehension, whereas the MAP utilizes silent reading and does not score the test based on the individual reading rate and accuracy. In this particular sample, reading fluency was not necessarily related to reading comprehension, which raises concern for the use of the GORT-5 as a diagnostic tool of reading comprehension impairment.

Finally, it is important to note the limitations of this study. Results should be interpreted with caution due to two primary factors. First, the sample size was small (n=28). This study should be replicated with a larger sample so that the benefits of limitations of the

MAP can be more clearly identified.

Second, the MAP and the GORT-5 have many structural differences, most notably the inclusion of oral reading fluency in the GORT-5 versus silent reading comprehension in the MAP. A clinical test of silent reading comprehension could have been used as more of a direct comparison to the MAP, however, the GORT-5 was chosen because it is a widely used clinical measure of reading comprehension, both clinically, and in research. Therefore, using it in this correlation study offered consistency with current research and clinical practice. The GORT-5 has a silent reading counterpart, the Gray Silent Reading Tests (GSRT), which should be considered in future research.

In conclusion, overall reading scores from the GORT-5 and MAP were strongly correlated and reading comprehension scores were fair to moderately correlated. The GORT-5 and MAP assess reading in different ways. In particular, the GORT-5 includes oral reading fluency as part of an individual's overall reading score. The current study questions the relationship between reading fluency and comprehension.

The data did not provide strong enough evidence to support the validity of the MAP, however, as a part of a comprehensive speech-language assessment, it is beneficial to consider the MAP scores, along with other academic assessments, to supplement standardized clinical tests in order to provide an in-depth view of a student's performance.

APPENDICES

APPENDIX A
RECRUITMENT LETTER

UNIVERSITY OF UND NORTH DAKOTA

DEPARTMENT OF COMMUNICATION SCIENCES AND DISORDERS
SPEECH, LANGUAGE AND HEARING CLINIC
MONTGOMERY HALL ROOM 101
290 CENTENNIAL DRIVE STOP 8040
GRAND FORKS, NORTH DAKOTA 58202-8040
(701) 777-3232
FAX (701) 777-4578

Dear Parents/Guardians,

I am a speech-language pathologist and researcher at the University of North Dakota. My research team is conducting a study comparing scores on the MAP test to other clinically administered tests. The MAP test is a computer-based test that is administered by your child's school. If your child chooses to participate, I will need your permission to access these scores. The clinical tests will include one test of reading comprehension and three tests of vocabulary knowledge, one of which entails using your child's current weekly spelling list. For your convenience, my research team can schedule testing sessions at your child's school during after school hours. Weekend and evening sessions are available at UND, as well. Each testing session should last approximately 1 hour. Your child will be compensated with $20 for their participation.

If your child is interested in participating in this study, please contact me via email: sarah.robinson@und.edu or by phone 701-777-1490.

Thank you,

Sarah Robinson, PhD, CCC-SLP

APPENDIX B
CONSENT FORM

PARENTAL CONSENT TO PARTICIPATE IN RESEARCH

TITLE:	A comparison of students' reading and vocabulary performance on MAP testing to performance on clinical measures
PROJECT DIRECTOR:	Sarah Robinson
PHONE #	777-3723
DEPARTMENT:	Communication Sciences and Disorders

STATEMENT OF RESEARCH
A person who is to participate in the research must give his or her informed consent to such participation. This consent must be based on an understanding of the nature and risks of the research. This document provides information that is important for this understanding. Research projects include only subjects who choose to take part. Please take your time in making your decision as to whether to allow your child to participate. If you have questions at any time, please ask.

WHAT IS THE PURPOSE OF THIS STUDY?
We invite your child to take part in a research study conducted by Dr. Sarah Robinson from the Department of Communication Sciences and Disorders at the University of North Dakota. The purpose of the study is to compare your child's score on sections of the MAP test (which s/he takes at school) to tests we are going to give him/her today.

HOW MANY PEOPLE WILL PARTICIPATE?
Approximately 80 fourth grade students will be selected to participate in this study. All of the students selected will need to complete the MAP testing in May (at school).

HOW LONG WILL MY CHILD BE IN THIS STUDY?
The testing session for this study will take approximately one hour. There will be only one testing session.

WHAT WILL HAPPEN DURING THIS STUDY?
There are two parts to the study.

1. The first part is the testing session. We will administer two standardized tests to evaluate your child's vocabulary abilities and one test to evaluate his/her reading abilities. This testing session will be video recorded.
2. The second part of the study is the MAP testing. The Grand Forks Public Schools administer the MAP testing to all students at school. With your permission, we will access your child's score for the May testing session.

WHAT ARE THE RISKS OF THE STUDY?
Participation in this study involves the following risks.

1. Your child may become uninterested, fatigured or frustrated during the testing session. We will offer appropriate breaks to use the restroom, get a drink of water, or walk around as needed. The tests that we are administrating are routinely used by speech-language pathologists during assessments.
2. It is possible that your child may become embarrassed if s/he does not know some of the items being tested. All participants will be assured that the items increase in difficulty and they will not know some or many of the words. They will be encouraged to guess if they are not sure or they will be told to respond "I don't know."
3. Your child may feel uncomfortable being video recorded during the testing session. Students will be assured that only the researcher and the research asssitants will have access to the video recordings. They will also be assured that we record sessions so that we can make sure that the evaluator has not made any mistakes.

WHAT ARE THE BENEFITS OF THIS STUDY?
Your child may benefit by knowing that s/he has helped in the research process. You will also have access to your child's vocabulary and reading comprehension scores. In the future, others may benefit by learning about what MAP scores tell educators.

ARE COSTS INVOLVED IN THIS STUDY?
You will not have any costs for allowing your child to participate in this research study. Upon completion of the testing session, your child will receive a $20 gift card.

WHO IS FUNDING THE STUDY?
The University of North Dakota and the research team are receiving no payments from other agencies, organizations, or companies to conduct this research study.

CONFIDENTIALITY
Confidentiality will be maintained to the extent allowed by law. We will make every effort to ensure that a loss in confidentiality does not occur. We will store all written records in a locked cabinet. We will store computer files related to your child's data under password protection. When the research program is complete, we will write up the results of the study as a research report. Your child will not be identified in any way except as a subject number. Our research records may be reviewed by Government agencies and the University of North Dakota Institutional Review Board.

IS THIS STUDY VOLUNTARY?
Your child's participation is voluntary. You or your child may choose not to participate or to discontinue participation at any time without penalty. Your decision whether or not to participate will not affect your current or future relations with the University of North Dakota.

INJURY DUE TO PARTICIPATION
If your child is injured as a direct result of being in this study, neither the University of North Dakota nor the principal investigator, Sarah Robinson, will pay for any care, lost wages, or provide other financial compensation. Please refer to the "Risks of the Study" section above for a list of possible risks of participating in the study.

CONTACTS AND QUESTIONS?
Sarah Robinson is the researcher conducting this study. You may ask any questions you have now. If you later have questions, concerns, or complaints about the research, please contact Sarah Robinson at 777-3723 during the day.

If you have questions regarding your rights as a research subject, or if you have any concerns or complaints about the research, you may contact the University of North Dakota Institutional Review Board at (701) 777-4279. Please call this number if you cannot reach research staff, or you wish to talk with someone else.

AGREEMENT

The University of North Dakota Institutional Review Board has approved this consent form as signified by the committee's stamp. This consent form must be reviewed at least once each year and expires on the date indicated on the stamp. Your signature below indicates that you have read the information in this document and have had a chance to ask any questions you have about the study. Your signature also indicates that you have decided to let your child participate, and have been told that you can change your mind and withdraw your consent for your child's participation at any time. You have been given a copy of this consent form to keep. You have been told that by signing this consent form you are not giving up any of your child's legal rights.

_____	_____	_____
NAME OF CHILD PARTICIPANT (please print)	AGE	DATE

_____	_____
SIGNATURE OF PARENT OR GUARDIAN	DATE

_____	_____
SIGNATURE OF INVESTIGATOR	DATE

APPENDIX C
ASSENT FORM

TITLE:	A comparison of students' reading and vocabulary performance on MAP testing to performance on clinical measures
PROJECT DIRECTOR:	Sarah Robinson
PHONE #	777-3723
DEPARTMENT:	Communication Sciences and Disorders

I am doing a research study. A research study is a special way to find out about something. I want to find out if kids score the same or different on two tests. If you want to be in this study, you will have to take a vocabulary test where you will first point at pictures of the words that I say and then you will tell me what some words mean. It is OK if you don't know the answers. Some questions are very difficult and it is OK to guess if you aren't sure or just say "I don't know". Next we will talk about your spelling words. I will ask you to tell me what some of the words mean. You will also take a reading test. Some of the things that I ask you to read will be easy for you and other things will be hard. It is OK to guess or say that you don't know. Just try your best. We will video record the testing session so that I can make sure that we have scored your answers correctly.

I want to tell you about some things that may happen to you if you are in this study. You may get tired of answering my questions. Or you may get tired of sitting for a long time. We will take a break in between the tests so that you can stretch, walk around or get a drink of water. If you want to take a break at any other time, you can tell me.

Not everyone who is in this study will benefit. A benefit means that something good happens to you. If you decide to be in the study and take the tests, you will get to pick a $20 gift card. You will also be helping with research. I hope that other people will be able to learn something from what we find out in this study.

When we are done with the study, I will write a report about what we find out. I will not use your name in the report. You do not have to be in this study. It is up to you. If you want to be in the study, but change your mind later, you can stop being in the study.

If you want to be in this study, please sign your name.

Your name (printing is OK) Date

I certify that this study and the procedures involved have been explained in terms the child could understand and that he/she freely assented to participate in the study.

Signature of person obtaining assent Date

References

Ackerman, R., & Lauterman, T. (2012). Taking reading comprehension exams on screen or on paper? A metacognitive analysis of learning texts under time pressure. *Computers in human behavior, 28*(5), 1816-1828. doi: 10.1016/j.chb.2012.04.023

Adams, M. (1994). *Beginning to read: Thinking and learning about print.* Cambridge, MA: MIT Press.

Barab, S. A., Young, M. F., & Wang, J. (1999). The effects of navigational and generative activities in hypertext learning on problem solving and comprehension. *International Jouirnal of Instructional Media*, 26(3), 283-309.

Beishuizen, J., Stoutjesdijk, E., & Zanting, A. (1996). Using hypertext for studying and information search. *Journal of Educational Computing Research*, 15(4), 289-316.

Blohm, P. (1982). Computer-aided glossing and facilitated learning in prose recall. In J. A. Niles & L. A. Harris (Eds.) *New inquiries in reading research and instruction. Thirty-first yearbook of theNational Reading Conference* (pp. 24-28). Rochester, NY: National Reading Conference.

Catts, H.W., & Kamhi, A.G. (2005). *Language and reading disabilities*. Boston: Allyn and Bacon.

Chall, J. & Jacobs, V. (2003). Poor children's fourth grade slump. American Educator. Available at http://www.aft.org/pubs-reports/american_educator/spring2003/

College Board. (2010). *Total group profile report*. Retrieved from website:http://media collegeboard.com/digitalServices/pdf/research/2010-total-group-profile-report-cbs.pdf

College Board. (2011). *SAT benchmarks: Development of a college readiness benchmark and its relationship to secondary and postsecondary school performance* (Research Report 2011-5). Retrieved from website:http://research.collegeboard.org/sites/default/files/publications/2012/7/researchreport-2011-5-sat-college-readiness-benchmark-secondary-performance.pdf

College Board. (2013). *SAT report on college and career readiness*. Retrieved from website: http://media.collegeboard.com/homeOrg/content/pdf/sat-report-college-career-readiness-2013.pdf

College Board. (2013). *Total group profile report*. Retrieved from website: http://media.college board.com/digitalServices/pdf/research/2013/TotalGroup-2013.pdf

Coulter, G. (2004). Using one-to-one tutoring and proven reading strategies to improve reading performance with adjudicated youth. *Journal of Correctional Education, 55*(4), 321-333.

Erwin, T. D. (1991). *Assessing student learning and development: A guide to the principles, goals, and methods of determining college outcomes*. San Francisco: Jossey-Bass.

Fish, M. C., & Feldman, S. C. (1987). A comparison of reading comprehension using print and microcomputer presentation. *Journal of Computer-Based Instruction*. 14, 57-61.

Gambrell, L. B., Bradley, V., & McLaughlin, E. (1985). Young children's comprehension and recall of computer screen displayed text. *Journal of Research in Reading.* 10(2). 156-163. doi: 10.1111/j,1467-9817.1987.tb00292.x

Gough, P.B. & Tunmer, W. (1986). Decoding, reading, and reading disability. *Remedial and Special Education*, 7, 6-10.

Gray, D.E. (2004). *Doing research in the real world.* Thousand Oaks, CA: Sage Publications Inc.

Grunau, R. E., Whitfield, M. F., & Davis, C. (2002). Pattern of learning disabilities in children with extremely low birth weights and broadly average intelligence. *Archives of Pediatrics & Adolescent Medicine, 156*(6), 615(6).

Heppner, F. H., Anderson, J.G. T., Farstrup, A.E., & Weiderman, N. H. (1985). Reading performance on a standardized test is better from print than from computer display. *Journal of Reading*, 28, 321-325.

Hoover, W.A., & Gough, P.B. (1990). The simple view of reading. *Reading and Writing*, 2, 127-160.

L'Allier, J. J. (1980). *An evaluation study of a computer-based lesson that adjusts reading level by monitoring on-task reader characteristics.* Unpublished doctoral dissertation, University of Minnesota, Minneapolis.

Lee, M.J., & Tedder, M.C. (2003). The effect of three different computer texts in readers' recall: Based on working memory capacity. *Computers in Human Behavior*, 19, 767-783.

Liu, Z. (2005). Reading behavior in the digital environment: Changes in reading behavior over the last ten years. *Journal of Documentation*, 61, 700-712.

Macedo-Rouet, M., Rouet, J., Epstein, I., & Fayard, P. (2003). Effects of online reading on popular science comprehension. *Science Communication*. 25(2), 99-128, doi: 10.1177/1075547003259209

National Center for Education Statistics. (June, 2013). *2012 long-term trend: Summary of major findings*. Retrieved from http://nationsreportcard.gov/ltt_2012/summary.aspx

National Center for Education Statistics. U.S. Department of Education, Institute of Education Sciences. (2014). *A First Look: 2013 Mathematics and* Reading. *National Assessment of Educational Progress at Grades 4 & 8.* 2014 (451). Retrieved from website: http://nces.ed.gov/nationsreportcard/subject/publications/main2013/pdf/2014451.pdf

NWEA. (2012). *Proctor handbook for client-server map users*. Retrieved from http://www.nwea.org/sites/www.nwea.org/files/resources/ProctorHandbook pdf.pdf

NWEA. (2011). *2011 normative data.* Retrieved from http://www.nwea.org/sites/www.nwea.org/files/resources/2011_Normative_Data_Overview.pdf

NWEA. (n.d.). *Map overview.* Retrieved from http://www.nwea.org/productsservices/assessments/map®

Pazzaglia, F., Toso, C., & Cacciamani, S. (2008). The specific involvement of verbal and visuospatial working memory in hypermedia learning. *British Journal of Educational Technology.* 39(1), 110-124.

Pham, A. V., Fine, J. G., & Semrud-Clikeman, M. (2011). The influence of inattention and rapid automatized naming on reading performance. *Archives of Clinical Neuropsychology*, *6*(3), 214-224. doi: 10.1093/arclin/acr014.

Pressley, M. (2006). *Reading instruction that works: The case for balanced teaching*. (3rd ed.). New York, NY: The Guilford Press.

Reinking, D., & Schreiner, R. (1985). The effects of computer-mediated text on measures of reading comprehension and reading behavior. *Reading Research Quarterly*, 20(5), 536-552. Retrieved from http://www.jstor.org/stable/747941.

Srivastava, P., & Gray, S. (2012). Computer-based and paper-based reading comprehension in adolescents with typical language development and language-learning disabilities. *Language, speech, and hearing services in schools*, *43*, 424-437. doi: 10.1044/0161-1461(2012/10-0108)

Storch, S.A., & Whitehurst, G.J. (2002). Oral language and code-related precursors to reading: Evidence from a longitudinal structural model. *Developmental Psychology*, 38, 934-947.

The Kingsbury Center at the Northwest Evaluation Association. (2010). North Dakota linking study: A study of the alignment of the NWEA RIT scale with the 11th grade North Dakota state assessment. Available at http://www.nwea.org/sites/www.nwea.org/files/ resources/ND%20 Linking%20Study%20August%202010%20rev%202010_1129.pdf

Whitehurst, G.J., & Lonigan, C.J. (1998). Child development and emergent literacy. *Child Development*, 69, 848-872.

Wiederholt, J. L. & Bryant, B. R. (2012). Gray Oral Reading Tests, Fifth Edition. Austin, TX; Pro-Ed.