THE READING-WRITING CONNECTION: AN INVESTIGATION OF THE RELATIONSHIP BETWEEN READING ABILITY AND WRITING QUALITY ACROSS MULTIPLE GRADES AND THREE WRITING DISCOURSE MODES

Heather H. Koons

A dissertation submitted to the faculty of the University of North Carolina at Chapel Hill in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the School of Education

Chapel Hill
2008

Approved by
Advisor: Dr. Gregory J. Cizek
Reader: Dr. Jill Fitzgerald
Reader: Dr. Leigh A. Hall
Reader: Dr. A. Jackson Stenner
Reader: Dr. William B. Ware

ABSTRACT

HEATHER H. KOONS: The Reading-Writing Connection: An Investigation of the Relationship between Reading Ability and Writing Quality across Multiple Grades and Three Writing Discourse Modes
(Under the direction of Gregory J. Cizek)

This study examines the relationship between grade-level reading comprehension and writing quality at grades 4, 6, 8, 10 and 12. Data were collected from a total 521 students in one school district in Mississippi. Two essay scores each in narrative, informative, and persuasive writing were obtained for each student, which enabled a close examination of the relationship between overall writing quality and discourse mode. A many-faceted Rasch model was used to adjust for rater severity in the scoring of student essays and place essay scores on an equal-interval scale. Results of correlational and structural equation modeling analyses indicate a developmental trend in the relationship between reading comprehension and writing quality with the largest correlations at grades 8, 10, and 12. The relationship between reading comprehension and writing quality was not affected in a systematic way by the discourse mode of the writing prompt. Confirmatory factor analysis was used to model the relationship between reading comprehension and writing quality and evaluate the factor loadings between writing quality and the discourse mode indicator variables. The model in which the factor loadings for adjacent grades were held constant was found to fit for all adjacent-grade comparisons except between grade 4 and grade 6, suggesting an invariant factor structure for grades 6, 8, 10, and 12. Comparisons between male/female groups and

Black/White groups produced similar results. Implications of these findings for future research, measurement practice, and classroom instruction and assessment are provided.

For Marilyn A. Hughes

ACKNOWLEDGEMENTS

This journey would not have been possible without the encouragement of my husband, Darren Koons. Without his support, patience, and gift of time, I would not have made it to this point. I am tremendously grateful. Darren, Kyra, and Dylan provided daily doses of love and perspective, which kept me motivated and moving forward.

I am deeply indebted to Dr. Gregory Cizek, who spent so much time working with me during all stages of this process. I am also thankful for the guidance provided by my readers, Drs. Jill Fitzgerald, Leigh Hall, Jack Stenner, and William Ware. As I encountered bumps along the way, numerous people were instrumental in keeping me moving, particularly Dr. Jack Stenner, Dr. Ellie Sanford, Dr. Cathy Zimmer, Hal Burdick, Allen Johnson, and Dr. William Ware. In addition to their support and encouragement, they all provided invaluable technical expertise regarding various aspects of my research. I am enormously grateful for their insights and assistance.

The importance of the friends and family who cheered me on cannot be understated; I could not have done this work without them. I am especially thankful for the support of my father Dr. J. David Hughes, my sisters Dr. Kathleen Hughes and Julie Hughes, Kitty Hughes, Sally Stone, John and Karin Koons, and many friends and colleagues including Dr. Samantha Burg, Sarah McManus, Sofi Frankowski, Amber Sandvik, and Dee Brewer.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

In the current age of high-stakes accountability systems and focused political attention on the educational system in the United States, K-12 educators are under tremendous pressure to provide evidence of the effectiveness of the instructional strategies that they employ. Two areas that enjoy prominence in the area of educational accountability and are the focus of this research are reading and writing.

One example of the focus on reading by national political leaders is the requirement included in the No Child Left Behind Act of 2001 (NCLB, 2002) that all states participate in biennial assessments of reading administered by the National Assessment of Educational Progress (NAEP). Although the NAEP assessments measure students across the nation in grades 4, 8, and 12 in a wide variety of subject areas including science, economics, geography, mathematics, reading, and writing, the NCLB requirements single out reading and mathematics for required NAEP participation. Separate provisions for writing were made in NCLB by funding a National Writing Project. However, most states (45 in grade 4 and 44 in grade 8) chose to participate voluntarily in the most recent NAEP writing assessments for which results have been reported. Results from these standardized assessments provide information to the public about how well students have met specific performance standards as defined by the National Assessment Governing Board (NAGB), which establishes policy for the NAEP. NAEP results are used to monitor student progress in reading and writing over

subsequent testing administrations as well as to facilitate state-to-state comparisons of overall student performance.

Student results on NAEP assessments are reported in three categories: basic, proficient, and advanced. A fourth category, below basic is implicit in that all scores that fall below the basic level are referred to as below basic. According to NAEP policy definitions for the achievement levels, students who score at a basic level have a partial mastery of prerequisite knowledge and skills that are fundamental for proficient work at each grade. The proficient level represents solid academic performance for each grade assessed, and the advanced level signifies superior performance. The 2007 reports indicate that nationally, 33% and 31% of students in grades 4 and 8, respectively, scored at the proficient level in reading comprehension. Only 8 and 3% of 4th and 8th graders, respectively, scored at the advanced level. The most recent reports for grade 12 reading comprehension, (2005), showed that 35% scored at or above the proficient level and 5% scored at the advanced level (National Assessment of Educational Progress, NAEP, 2007).

The most recent results in writing (2002) are similar. Results for 4th, 8th, and 12th grades show that between 22 and 30% of students nationally are at a proficient level in writing and just 2% are at an advanced level (National Center for Education Statistics, NCES, 2003). Not surprisingly, such disappointing results lead to increased scrutiny of educators and their instructional programs. As teachers of reading and writing seek to improve the achievement of their students and look for ways to improve the efficacy of their instructional programs and the efficiency of their instructional practices, they may look to capitalize on the knowledge and skills shared by reading and writing.

Although research into the relationship between reading and writing dates back to the1930s, historically, most pedagogy has separated reading and writing instruction (Nelson & Calfee, 1998). In the 1970s and 1980s, an increasing number of researchers and practitioners pointed out the benefits of integrating reading and writing instruction (Durkin, 1988; Moffett & Wagner, 1983; Stauffer, 1980), and have continued to encourage their integration (Duke & Pressley, 2002; Hiebert, Pearson & Taylor, 1998; Shanahan, 1990). As educators have begun to incorporate findings about the benefits of reading and writing into curriculum (Shanahan, 1990) research continues. This research has influenced thinking about the nature of reading and writing and how to advance student learning in both areas. The research has also provided a foundation for additional investigations into the reading-writing relationship. Today there is a large body of educational research focused on the topic of reading and writing connections. Many volumes, chapters, and literature reviews are devoted to the topic (e.g., Heller, 1995; Irwin & Doyle, 1992; Langer, 1986; Nelson & Calfee, 1998; Shanahan, 1990, 2006; Stotsky, 1983; Tierney & Shanahan, 1996). However, although much has been learned about reading-writing connections, many areas remain to be more fully explored.

Intuitively, it makes sense that instruction in reading will improve student performance on reading assessments and instruction in writing will improve student performance on writing assessments. Research supports both of these conjectures (Berninger, 2000; Stanovich, 2000). The results of research investigating the impact of reading instruction on writing performance and the impact of writing instruction on reading performance are not as straightforward. Some studies have shown that instruction in either reading or writing improved outcomes in the other area (Bond & Dykstra, 1967; Crowhurst, 1991; DeVries

1970, Stahl, Pagnucco, & Sutters, 1996), whereas others showed little or no effect (Shanahan & Lomax, 1986, 1988). A fairly strong relationship has been found between reading more, between more exposure to text and higher reading achievement (e.g., Stanovich, 2000), and between reading and writing ability; better readers are generally better writers and better writers are generally better readers (Juel, 1988; Loban, 1963; Woodfin, 1968). Student survey results from NAEP 2002 writing administrations (National Center for Education Statistics, 2007) provide additional evidence that more reading is associated with higher writing scores. Grade 8 results showed that students who reported reading for fun or on their own almost every day had an average writing scale score of 168 and students who reported reading for fun or on their own 1-2 times a week had an average writing scale score of 155; those who never or hardly read for fun on their own had an average writing scale score of 143. Although these scores all fall within the basic range (114 -173), the differences between them are statistically significant. Grade 4 results showed a similar pattern (National Center for Education Statistics, 2007).

A growing body of research on the reading-writing relationship has used varied approaches to provide insight into this relationship. Each investigation into the reading-writing connection has as its foundation an approach that reflects specific beliefs about what comprises reading and writing and their relationship, although these are not always explicitly stated. The current research is guided by the perspective that reading and writing share elements of a common knowledge base as well as overlapping cognitive processing (Fitzgerald & Shanahan, 2000) and that valuable information about these relationships can be gained by examining relationships between student performance in reading and writing. This approach is situated among other approaches; the most prominent are introduced below and

elaborated upon in the literature review. The following section describes the approach used in the current research.

Theoretical Approach: Shared Knowledge and Cognitive Processes

Though quite varied, research on reading-writing connections can be grouped into three main theoretical approaches: (1) rhetorical relations, (2) procedural connections, and (3) shared knowledge and cognitive processes (Fitzgerald & Shanahan, 2000). The *rhetorical relations* approach to studying the reading-writing connection focuses on the similarities between communication experiences shared by readers and writers. Both reading and writing are viewed as meaning-making activities. In contrast to the view that reading is a somewhat passive act and a reader simply receives meaning from text, the reader, much like a writer, is actively engaged in creating new meaning. The *procedural connections* approach focuses on the functional impact of reading on writing and writing on reading as well as their impact on external goals such as performance in academic content areas (e.g., history). The third main approach to research on the reading-writing connection adopts the *shared knowledge and cognitive processes* perspective, which examines the commonalities in the processes and knowledge shared by reading and writing. The literature review discusses all three approaches. However, because the current research adopts the shared knowledge and cognitive processes approach, the following section addresses that approach only.

The majority of research into the reading-writing connection has utilized the shared knowledge and cognitive processes (SKCP) approach. The premise behind this approach is that reading and writing are connected because they "depend on identical or similar knowledge representations, cognitive processes, and contexts and contextual constraints"

(Fitzgerald & Shanahan, 2000, p. 40). The SKCP approach is broad enough to include elements of the rhetorical relations and procedural connections approaches. For example, the idea that both writing and reading are meaning-making activities (a fundamental aspect of the rhetorical relations approach) is compatible with the SKCP approach; the investigation of rhetorical relations aspect of the reading-writing relationship is simply not the primary focus of research that adopts the SKCP approach.

According to Fitzgerald and Shanahan (2000), readers and writers rely on four common knowledge bases: domain or content knowledge, procedural knowledge, knowledge of specific features or components of written language, and metaknowledge. Domain or content knowledge refers to specific knowledge of the topic about which one is reading or writing. Procedural knowledge refers to the knowledge and skills needed to negotiate reading and writing (Langer, 1986). This can include relatively automatic processes, such as recalling information from memory, or more intentional strategies, such as predicting, summarizing, or questioning (Kellogg, 1994). Procedural knowledge is also described as the cognitive processes employed by readers and writers to access, use, and generate information during reading and writing (Shanahan, 2006). Knowledge of specific features or components of written language include knowledge of features from the word level to text level including phonemic, orthographic, morphological, lexical, syntactic, and discourse features. Metaknowledge includes knowing about the functions and purposes of reading and writing, knowing that readers and writers interact, monitoring one's own meaning-making, and monitoring word identification strategies (Fitzgerald & Shanahan, 2000). This list of shared features highlights the complexity of the reading and writing endeavors. Not surprisingly, as

researchers have focused on one or more of these shared features, the researchers have adopted varied research methods.

Two dominant avenues for investigations into the shared knowledge and cognitive processes between reading and writing are processed-based and performance-based correlational studies. Process-based studies focus on the processes that readers and writers bring to bear when reading and writing. These processes can include prediction, questioning, summarizing, reflecting, and hypothesizing; the studies typically use methods such as think-alouds, interviews, and observations. In particular, these methods have been utilized in studies of the metacognitive and procedural knowledge shared by reading and writing (Brandt, 1994; Langer, 1986; Ryan, 1985). Performance-based correlational studies typically consider the relationship between student scores on performance measures of reading and writing. In the case of either reading or writing, the measures being correlated could be measures of a cognitive process that underlies either reading or writing (e.g., brain functioning), precursor skills for reading or writing (e.g., phonemic awareness, vocabulary, handwriting, spelling), an overall performance measure (e.g., a holistic writing score), or a combination of these.

An implicit assumption in the performance-based studies is that the performance measures (tests) chosen for the study adequately and appropriately produce estimates of student ability that represent reading or writing ability or the aspects of reading or writing that are of interest. Reading and writing ability are examples of *latent* abilities, that is, abilities that we cannot measure directly and thus must measure using indirect means such as reading or writing assessments. In other words, the student scores are treated as external manifestations of the literacy knowledge or process: the latent ability.

This research study focuses on outcome measures of the reading and writing activities with the assumption that the degree to which someone can read well can be inferred from a score on a reading comprehension test and the degree to which someone can write well is reflected in a score on a writing test. This research focuses on reading comprehension of grade-level text and writing quality elicited by grade-appropriate writing prompts and how the relationship between them varies across grades and types of writing.

Reading Comprehension Scores

It is important to recognize that an examinee's score on a multiple-choice reading comprehension test represents the degree to which the student's comprehension of the pieces of text used in the assessment is in agreement with what the test developers believe is the correct interpretation of the text. However, the test construction process typically involves many stages of item review, so the items reflect ideas explicitly represented in the text and similarly understood by multiple individuals. This review process helps to ensure that responses to the text by skilled readers, which often depend greatly on the prior knowledge of the reader, will converge on the one correct answer that is most reasonable to a wide variety of readers. Similarly, a writer's score on any given essay represents what the essay raters believe is good writing, though these beliefs are typically guided by a specific set of scoring criteria, generally presented in the form of a rubric.

Writing Scores

The assessment of writing using extended responses such as essays is complicated by the presence of human judgment in the generation of the score. Unlike an objectively scored,

multiple-choice assessment, the essay score involves human raters who judge the quality of the writing. Engelhard (2000) addressed this issue regarding rater-mediated (RM) assessments:

> RM assessments do not provide direct information regarding examinee achievement because the examinee's responses must be mediated and interpreted through raters to obtain judgments about examinee achievement. One of the major concerns regarding RM assessments is that raters bring a variety of potential response biases that may unfairly affect their judgments regarding the quality of examinee responses. (p. 261)

Ideally, the ratings given to student essays would be the same regardless of which raters scored them. In reality, raters vary in their relative severity or leniency in assigning scores to essays (Engelhard, 1994, 2000; Hollenbeck, Tindal, & Almond, 1999). Thus a student whose essay was scored by a severe rater would be at a disadvantage compared to a student whose essay was scored by a lenient rater. Engelhard (1994) described a measurement model that can take into account the relative severity or leniency of the raters and adjust the estimates of writing ability accordingly. This model, the many-facet Rasch model, is used in the analysis of writing quality in this study. As a result, the students' writing scores are a better estimate of the quality of their writing and a more appropriate estimate to use when comparing reading and writing than scores that are unadjusted for rater differences.

The Writing Task

Another area in the assessment of writing that has received attention is the impact of the discourse mode of the prompt on the score received by the writer. Studies of student writing have shown that the mode of the prompt (e.g., narrative, expository, persuasive) has an impact on a student's writing quality score (Engelhard, Walker, Gordon, & Gabrielson, 1994; Kegley, 1986; Kuhlemeir, van den Bergh, & Wijnstra, 1989; Quellmalz, Capell & Chou,

1982; Sachse, 1984). However, the results have not all supported the relative ease of one type of discourse mode over another. Some have shown that students who write to narrative prompts earn higher scores on average than students who write in response to prompts in descriptive, expository, or persuasive modes (Engelhard et al., 1994; Kegley, 1986; Sachse, 1984). In contrast, Quellmalz et al. (1982), in a study of 11th and 12th-grade students, found that students who wrote to a narrative prompt received lower scores than students who wrote to an expository prompt. The researchers suggested that this finding may have been due to the increased emphasis on expository writing in high school writing instruction and the accompanying reduction in narrative writing.

Regardless of the findings about the relative effect of discourse mode on student performance, weaknesses in these studies are causes for concern when interpreting the findings. In many of these studies, the same students have not written essays in multiple modes, so any differences in scores across modes may be the result of student rather than mode variation. In addition, the impact of rater variations was not incorporated into the scoring protocols, so variation in scores may have been affected by rater variation. The current research addresses both of these limitations of prior studies.

Need for the Study

Although a great deal has been learned about the relationship between reading and writing, some areas have not yet been fully explored. In particular, little empirical information exists about the reading-writing connection across grade levels. Much of the work on reading-writing connections has examined relationships at specific grade levels with no systematic investigation across grade levels. The current study extends existing

knowledge by examining the relationship between reading ability and writing quality across grade levels. In addition, most of the existing research on the reading-writing relationship that uses student essays as an outcome measure for writing does not address rater impact on student scores. The current research utilizes the methods described by Engelhard (1992, 1994) to take into account rater severity when measuring writing. The resulting scores are a more accurate writing measure and thus may provide a clearer look at the reading-writing relationship. Several writing samples by the same student are analyzed, so the effect of mode can be more easily identified than if different students wrote in response to different modes. Finally, the present study uses the same outcome measure for reading and writing for all grade levels. The use of the same instruments across grades allows comparison across grade levels that is not confounded by an inconsistent operationalization of reading and writing across grade levels. In addition, the reading and writing instruments used were measures of culminating processes rather than precursory skills, so it is possible that a large proportion of the shared relations between reading and writing is captured.

Research Questions

The primary purpose of this study is to examine the relationship between reading ability and writing quality across grades 4, 6, 8, 10, and 12. A secondary purpose is to examine the relationship between mode of discourse and writing quality scores across grades levels as well as their relationship to reading. To achieve these purposes, three research questions and associated subquestions were formulated.

(1) What is the strength of the relationship between reading ability and writing quality across grade levels and is the strength of the relationship similar across grade levels?

(2) Is the strength of the relationship between reading ability and writing quality affected by the mode of the writing prompts (narrative, informational, persuasive)?

(a) What is the relationship between mode of the writing prompt and writing scores? (i.e., Are the modes similar in difficulty?)

(b) Is writing quality in the three modes (narrative, informational, persuasive) divergent enough that their relationship to reading should be modeled separately, as three factors, or are they similar enough to be modeled as one overall writing factor?

(3) Is the strength of the relationship between reading ability and writing quality similar across gender and race groups? Is the relationship between writing quality and the discourse mode of the prompt similar across gender and race groups?

Summary

As pressure mounts for educators to show improvement in student skills in reading and writing, much may be gained by better understanding the strength of the relationship between the two. An unintended consequence of the increased focus on test scores frequently voiced by concerned parents and educators is that curriculum decisions are being driven by what is likely to be on the upcoming test. Students in grades with no writing test may receive comparatively little or no instruction in writing, or writing instruction may be focused primarily or exclusively on the discourse mode that will be tested in the current or subsequent grade. This research provides new insights into the connection between reading and writing that will help educators find reasons to incorporate instruction in reading and writing in all modes into their curricula and classroom practices. As teachers focus more on

reading and writing in general and less on the specifics of each mode, students will likely gain more proficiency in skills that can transfer across modes and outside of the classroom. Faced with an ever-changing society and an economy that sometimes requires unanticipated career changes, students will be better served by instruction that prepares them to read and write broadly rather than in narrowly defined modes.

CHAPTER 2

LITERATURE REVIEW

This chapter provides a context for the current study by providing a review of literature related to the topics in reading-writing research that are most pertinent to the questions addressed and theoretical approach used by the research that is the focus of the present study. This chapter first provides a summary of the major approaches taken by research that addresses the reading-writing relationship. The perspective adopted by this research is described along with results of research studies that have adopted a similar approach. Next, this chapter explores issues related to assessing writing quality. Finally, the chapter presents a theoretical framework for approaching the question of developing readers and writers and its relationship to the current research.

Reading-Writing Research

Until the 1970s and 1980s, reading and writing were taught in the majority of schools in the United States as if they were separate endeavors (Nelson & Calfee, 1998). However, investigation into relationships between reading and writing has a long history in educational research, dating back to the 1930s (Stotsky, 1983). An underlying theme of this research has been the idea that if substantive connections exist between reading and writing, students may benefit from integrated instruction in reading and writing. Additionally, some research has investigated whether instruction in reading or writing can substitute for or supplement

instruction in the other. Generally, research has shown that "reading and writing rely on corresponding or correlated mental processes and isomorphic knowledge, though the nature of the relations between reading and writing is different at different age or grade levels" (Fitzgerald & Shanahan, 2000, p. 42). These correspondences can lead to more effective instruction when reading and writing are taught in combination, but have not been shown to be substantial enough for either reading or writing instruction to substitute for the other.

As educators have begun to incorporate findings about the benefits of reading and writing into curriculum (Shanahan, 1990), research continues. Today there is a large body of educational research focused on the topic of reading and writing connections. Many volumes, chapters, and literature reviews are devoted to the topic (e.g., Heller, 1995; Irwin & Doyle, 1992; Langer, 1986; Nelson & Calfee, 1998; Shanahan, 1990, 2006; Stotsky, 1983; Tierney & Shanahan, 1996). However, although much has been learned about reading-writing connections, many areas remain to be more fully explored. In particular, there is scant empirical information about the reading-writing connection across multiple grade levels.

The growing body of research on the reading-writing relationship has provided insight into this relationship by using varied approaches. Each investigation into the reading-writing connection has as its foundation an approach that reflects specific beliefs about what comprises reading and writing and the relationship between them, although the beliefs are not always explicitly stated. The approach taken by a researcher guides the types of questions asked about the reading-writing relationship and determines appropriate methods to employ when answering those questions. The approach employed in the current research is guided by the perspective that reading and writing share elements of a common knowledge base as well as overlapping cognitive processing (Fitzgerald & Shanahan, 2000) and that these elements

can be represented by scores on assessments of reading comprehension and writing. Valuable information about the relationship between reading and writing within and across grade levels can be gained by examining relationships between student scores on reading comprehension tests and holistic scores on essay writing. The shared knowledge and cognitive processes approach is situated among other approaches; these are described in the following section, with a particular focus on the approach and methods adopted in this research. The following reviews previous research into the reading—writing connection and then provides a discussion of reading comprehension and writing quality as the particular focus of this research.

Three Approaches to the Study of the Relationship between Reading and Writing

Although the body of research on reading-writing connection is quite broad and diverse, it can be grouped into three main theoretical approaches: (1) rhetorical relations, (2) procedural connections, and (3) shared knowledge and cognitive processes (Fitzgerald & Shanahan, 2000). The current research adopts the shared knowledge and cognitive processes approach, so the following overview provides the most information on that approach. However, although the current research is grounded in the shared knowledge and cognitive processes approach, the beliefs about what constitutes reading comprehension and writing quality that undergird this research are compatible with other approaches to research on the reading-writing connection.

The Rhetorical Relations Approach

The rhetorical relations approach considers reading and writing to be communication activities; readers and writers gain insight into communication by being both the sender and receiver (Fitzgerald & Shanahan, 2000). In contrast to a view that considers writing a process of putting meaning on the written page (meaning production) and reading a process of getting meaning from the written page (meaning reception), the rhetorical relations approach considers both writing and reading to be meaning-making activities. Reading as a meaning-making activity is highlighted by intertextually informed research that views a text as an "intermediate, provisional, unfinished work, open to new amplification and interpretation, engendered by its existence in a complex set of shifting relations. . . . From out of many texts, the text becomes many more" (Hartman, 2004, p. 356). Wittrock (1984) stated that both reading and writing are generative processes in which readers and writers "create meanings by building relations between the text and what they know, believe, and experience" (p. 77). Similarly, Tierney and Pearson (1984) described reading and writing as

> essentially similar processes of meaning construction. Both are acts of composing. From a reader's perspective, meaning is created as a reader uses his background of experience together with the author's cues to come to grips both with what the writer is getting him to do or think and what the reader decides and creates for himself. As a writer writes, she uses her own background of experience to generate ideas and, in order to produce a text which is considerate to her idealized reader, filters these drafts through her judgements [*sic*] about what her reader's background of experience will be, what she want to say, and what she wants the reader to think or do. (p. 33)

Fitzgerald (1990) extended these ideas to describe the affective overlaps between reading and writing in addition to their shared meaning-making elements. The affective overlaps are the connections between reading and writing that go beyond the cognitive processes that are shared to include shared, noncognitive aspects. She described a universe in which reading and writing take place as an interaction among readers, writers, and texts. Within this

universe, a driving force is desire for "mind meeting" (p. 82) and both the reader and writer consider the other when engaging in their respective activities. When composing, writers are influenced by the perceived expectations of the reader as well as their own goals for their written work. Similarly, a reader considers the writer's purpose for writing when reading the text; the reader's beliefs about the writer's intentions are integrated into the reader's own understanding of what is read. For example, a reader who knows that the author of a medical report on the effects of tobacco use is funded by the tobacco industry may suspect that the findings are biased toward minimizing harmful effects of tobacco use. The writer of an article for a teen magazine commonly chooses a language style that would be most likely to engage a teen reader. Thus, the reading and writing experiences both involve an element of reaching out and seeking an understanding of the other person involved in the two-way interaction.

The Procedural Connections Approach

The procedural connections approach treats reading and writing as functional activities that can be combined to accomplish external goals (Fitzgerald & Shanahan, 2000). Typically, studies that employ a procedural connections approach investigate how various writing activities, in combination with reading, can enhance learning of academic material.

Some of the earliest studies on the reading-writing connection illustrate this approach. Dynes (1932) found that taking notes, outlining, and summarizing were superior to reading and rereading alone for immediate learning and for retention of information in high school history classes. In the 1970s, work in the area of writing-to-learn gained additional support by a group of educators who promoted writing in all subject areas (Bangert-Drowns, Hurley,

& Wilkinson, 2004). In particular, Emig (1977) claimed that writing and learning are similar processes and that writing is a unique mode of learning.

The natural extension of this idea—that any writing at all will improve learning—has not been supported by research. Rather, the degree to which the writing activity improves instructional outcomes in subject area content is related to the nature of the writing task. For example, Newell (1984) studied the effects of using note-taking, study guide questions, and essay writing on learning from prose passages in science and social studies. He found that students involved in essay writing gained the most knowledge related to key concepts and that the effect was greater for students who had little initial knowledge of the topic. Newell also had students think aloud as they wrote. He found that essay writing, as compared to note-taking or answering study guide questions, involved more cognitive operations and reasoning that went beyond the simple translation of ideas. Students who wrote essays made an effort to integrate the information from the passage into a coherent text, an operation not observed when students answered study guide questions or took notes. Newell hypothesized that the cognitive operations and reasoning involved in completing the essay-writing task contributed to the higher scores for learning from the passages.

Bangert-Drowns et al. (2004) meta-analysis of 48 writing-to-learn research studies supported the finding that all writing is not equally effective in increasing learning. They found that "implementations of writing-to-learn instruction, as represented in comparative studies, result fairly consistently in positive, albeit small, effects on content learning by conventional academic measures" (p. 39). In their review, 75% of the studies showed positive outcomes for the writing-to-learn over conventional instruction on the same content, with an average unweighted effect size of .26 standard deviations and an average weighted

effect size of .17. The weighted analysis weighted each effect by the inverse of its sampling error. The weighting strategy assumes that larger studies give better estimates (less sampling error) and thus gives more weight in the analysis to studies with larger samples. Like Newell (1984), Bangert-Drowns et al. (2004) found that the type of writing used in the intervention influenced the outcomes. When writing activities involved personal writing only, the outcomes were no different than for instruction without personal writing. However, instruction that involved writing assignments that required some level of metacognition about learning showed more positive results. Examples of metacognitive writing assignments that resulted in increased learning included those that required students to evaluate their current understandings, confusion, and feelings in relation to the subject matter.

A related area of investigation has examined whether instruction in some aspect of either reading or writing can increase achievement in the complementary skill. For example, DeVries (1970) found that grade 5 students who did extra expository reading instead of writing wrote better expository compositions at post-test time than students who wrote two themes a week. His results suggested that in some cases, reading practice may be more effective than writing practice in improving writing ability.

In another study of the impact of reading and writing activities on writing scores, Crowhurst (1991) implemented a 10-lesson intervention that involved instruction in a model structure for persuasive writing plus either practice writing persuasive essays (writing group) or reading persuasive text (reading group). Students in either the writing group or reading group received higher global quality scores for persuasive compositions than a control group that engaged in activities unrelated to persuasive reading or writing. Both the writing group and the reading group improved significantly in writing quality from pretest to post-test and

both scored significantly higher than the control group on the post-test, which supported the idea that both reading and writing can improve writing quality.

In general, however, the results of studies that examined the impact of reading activities or writing activities on the complementary skill are mixed. Some studies have shown that additional reading can improve components of writing, such as grammar and writing performance generally, more than more writing practice alone (Elley, Barham, Lamb & Wyllie, 1976; Mills, 1974). Other studies have not shown evidence of improvement in composition skills when students had a reading program alone (e.g., Shanahan, 1984). A possible explanation for these results can be found in the work of Shanahan and Lomax (1986), who examined the influence of reading on writing and the influence of writing on reading using structural equation analysis. They found that an interactive model in which reading and writing support each other was superior to a model in which reading skills caused writing skills or a model in which writing skills caused reading skills. According to Shanahan (2006),

> A common finding has been that some reading-to-writing or writing-to- reading learning is possible, but that instruction targeting skills in one or the other tends to be most effective at improving that dimension. . . . Reading and writing instruction can be usefully combined, but instruction in one or the other is unlikely to be an adequate replacement for the other if the goal is to develop students who can read and write well. (p. 177)

Although instruction in reading or writing may not adequately replace the other if the goal is to develop both areas, studies have shown that instruction in each area may enhance learning in the other as well. Thoughtfully integrating reading and writing in an instructional program, with careful attention paid to the kind of reading and writing tasks assigned, may be a more efficient means to improve performance in both.

The Shared Knowledge and Cognitive Processes Approach

The greatest amount of research into the reading-writing connection has been in the areas of shared knowledge and cognitive processes. The premise behind this perspective is that reading and writing are

> constellations of cognitive processes that depend on knowledge representations at various linguistic levels (phonemic, orthographic, semantic, syntactic, pragmatic). Reading and writing are connected, according to such views, because they depend on identical or similar knowledge representations, cognitive processes, and contexts and contextual constraints. Therefore, we should expect reading and writing to be quite similar, their developments should parallel each other closely, and some type of pedagogical combination may be useful in making learning more efficient. (Fitzgerald & Shanahan, 2000, p. 40)

The drive to make learning more efficient is powerful in educational research, so it is natural that a great deal of work has been conducted in an attempt to tease out the shared elements between reading and writing and explore whether instruction and improvement in one area (e.g., writing) can enhance instruction and improvement in the other (e.g., reading).

A foundation for any shared knowledge or cognitive processes is a basic set of neurophysiological abilities. A premise underlying an investigation into shared cognitive processes is that reading and writing also share neurophysiological abilities that enable the cognitive processes to function effectively. These include the processes that enable visual, phonological, and semantic systems to function appropriately. Berninger, Abbot, Abbot, Graham, and Richards (2002) did extensive work as part of a decade-long research project into the relationship between reading, writing, listening, and speaking. In one study, Berninger et al. (2002) used brain imaging techniques to compare the brain activity of boys with dyslexia who had reading and writing difficulties to the brain activity of boys who were what Berninger et al. referred to as "good readers." The researchers presented the boys with tasks that required listening to language and making judgments about phonological

characteristics of what they heard (e.g., whether or not the words rhymed). Results showed higher levels of chemical activation of lactate (a byproduct of brain metabolism) in the boys with learning disabilities, suggesting less efficient phonological processing. The boys with dyslexia were given a phonologically driven reading intervention over a summer session and follow-up sessions throughout the school year. One year after the initial brain scans, both the boys with dyslexia and the control group were re-imaged. There was no longer a difference in the levels of chemical activation of lactate between the two groups. Because the levels were stable for the control group only, the change for the dyslexic group was attributed to the instructional intervention, leading Berninger et al. to conclude that "the brain is both an independent variable that constrains learning and a dependent variable that may change in constrained ways in response to intervention" (p. 52).

Although reading and writing are affected by similar neuropsychological abilities they also rely on common knowledge bases and cognitive processes. According to Fitzgerald and Shanahan (2000), readers and writers rely on four common knowledge bases: domain or content knowledge, procedural knowledge, knowledge of specific features or components of written language, and metaknowledge. Domain or content knowledge refers to specific knowledge of the topic about which one is reading or writing. Procedural knowledge refers to the knowledge and skills needed to negotiate reading and writing (Langer, 1986). This can include relatively automatic processes such as recalling information from memory or more intentional strategies such as predicting, summarizing, or questioning (Kellogg, 1994). Knowledge of specific features or components of written language include knowledge of features including phonemic, orthographic, morphological, lexical, syntactic, and discourse features. Metaknowledge includes knowing about the functions and purposes of reading and

writing, knowing that readers and writers interact, monitoring one's own meaning-making, and monitoring word identification strategies (Fitzgerald & Shanahan, 2000).

Two major avenues for investigations into these aspects of the shared knowledge and cognitive processes between reading and writing are process-based studies and performance-based studies. The current research can be described as a performance-based study as it focuses on an analysis of performance measures of reading comprehension and writing quality. However, the focus on performance (i.e. test scores) does not negate a recognition that reading comprehension and writing draw upon similar processes and are both meaning-making endeavors.

Process-based studies. Process-based studies collect information about the cognitive processes that are shared between reading and writing. These can include processes such as questioning, drawing conclusions, integrating information, developing hypotheses, reflecting on ideas, summarizing, and making connections. Shanahan (2006) described the ability to employ the cognitive processes actively as procedural knowledge, which refers to knowing how to access, use, and generate information during reading and writing. This includes awareness of intentional strategies such as prediction, questioning, and summarizing (Langer, 1986). Methods such as think-alouds, interviews, and observations are often employed in these studies.

In a comprehensive process-based study, Langer (1986) examined the knowledge sources, reasoning operations, strategies, and monitoring behaviors of 67 3rd, 6th, and 9th grade children when they read and wrote stories and reports. She found that the behaviors were varied and complex and that they changed with age and difficulty of the task. However, she also found that the behaviors varied consistently between reading and writing. For

example, in both reading and writing, students in grade 9 were more likely than students in grade 3 to reflect on their ideas. In addition, the students' attempts to "develop hypotheses about the evolving meaning" (p. 77) of what they were reading or writing increased in both areas about 11% from grade 3 to grade 9. Thus Langer's work supports the notion that some of the processes involved in both the reading and writing tasks (e.g., degree of reflection on ideas) share a pattern of development in which older students engage in more of the behavior in both reading and writing activities than younger students.

Performance-based studies. Performance-based studies typically look at the relationship between performance measures of reading ability and writing ability. In the case of either reading or writing, the measures correlated could be an overall performance measure (e.g., an holistic writing score) or a prerequisite skill (e.g., phoenemic awareness or vocabulary knowledge to describe reading ability). Studies have found positive relationships between various aspects of reading and writing at all developmental levels (Abbott & Berninger, 1993; Hiebert, Englert, & Brennan, 1983; Juel, 1983; Loban, 1963; Perin, 1998; Shanahan, 1984; Woodfin, 1968).

In the longest longitudinal study to date of the relationship between reading and writing, Loban (1963) collected data on 220 students over 12 years (from 1st through 12th grades). He found a positive relationship between reading and writing and concluded that "those who read well also write well; those who read poorly also write poorly" (p. 75). This study measured reading ability using the *Stanford Achievement Test* and writing using single writing prompts. The writing prompt at each grade consisted of a single picture, fairly complicated in terms of its content, to which students were directed to write. In grades 10, 11, and 12, students were assessed 2 to 3 times each year. One essay was written in response

to a picture prompt and the other(s) to assigned topics (Loban, 1967). The students were given as much time as needed to complete the task. Loban grouped the students by writing ability into five groups: superior, good, inferior, illiterate, and primitive. He then scored the students' reading ability according to chronological age and found that "*every* subject ranked superior in writing is *reading above* his chronological age; *every* subject ranked illiterate or primitive in writing is *reading below* his chronological age" (Loban, 1963, p. 75, emphasis in original).

In a longitudinal study of reading development of students in 1st through 4th grades, Juel (1988) found that students who had less phonemic awareness at the beginning of first grade were significantly more likely to become poor readers and poor writers in their years from 1st through 4th grades. Perin (1998) found that poor adult readers were more likely to show less syntactic sophistication in their writing and made more word-level errors. Similarly, in a study of college students, Hiebert et al. (1983) found a significant correlation between overall reading and writing scores ($r = .35$, $p < .01$) and significant differences between students with high and low reading ability in their skill in writing paragraphs according to a specified text structure. Woodfin (1968) found that student scores on a reading test were more highly correlated with student writing in terms of effectiveness of expression of ideas ($r = .50$, $p < .01$) than specific writing features such as spelling, capitalization, punctuation, or usage. In a study of more than 600 students in grades 1 – 6, Abbott and Berninger (1993) found correlations between passage comprehension and a measure of narrative or expository writing quality (measured on a 5-point holistic scale) ranging from .22 to .54. Correlations between passage comprehension and a measure of narrative or expository writing fluency (measured as number of words and clauses written) were lower, ranging from .08 to .34.

Shanahan (1984) examined the relationship among four reading measures (vocabulary, phonics, comprehension, and sentence completion) and eight writing measures (vocabulary diversity, average t-unit length, number of episodes, number of story categories, number of information units, spelling accuracy, phonemic accuracy, and orthographic accuracy) for grade 2 and grade 5 students. He found correlations between the reading and writing measures that ranged from .14 (vocabulary and writing episodes) to .68 (phonics and orthographic accuracy) and concluded that "neither reading nor writing was found to be sufficient to explain more than 43% of the variance of the opposite set" (p. 475).

These are just a few examples in a large set of performance-based studies that compare student scores on various instruments that measure aspects of reading and writing. In cases in which the studies have reported correlations between the reading and writing scores, most have reported correlations in the .20 to .50 range (typically no more than 25% shared variance); correlations in this range have been fairly consistent regardless of the age studied or the particular components of reading or writing studied (Fitzgerald & Shanahan, 2000). The notable exception is work by Berninger et al. (2002), which found that with the use of multiple indicators for each factor, shared variance was in the 77–85% range for word recognition and spelling and in the 65–66% range for text-level comprehension and composition.

Given statements such as Squire's (1984) assertion that both "comprehending and composing seem basic reflections of the same process" (p. 24), we would expect correlations higher than .50 between measures of reading and writing. Various reasons for the unexpectedly low correlations have been suggested. For example, Fitzgerald and Shanahan (2000) noted,

> Most of these studies have been small scale (fewer than 50 participants), have been conducted at a single point in time (few longitudinal or cross-panel studies) and have usually focused on bivariate relations as opposed to multivariate relations. The low correlations have also been blamed on the low reliability of the writing measures used or because only a specific component of reading or writing had been studied (spelling or decoding, for example) which missed other areas of relation between reading and writing. (p. 41)

The current research addresses many of the shortcomings listed above. In particular, the measures used in this study, reading comprehension and writing quality are not component skills of either the reading or writing process. Reading comprehension and essay writing are culminating activities that involve the use of multiple skills (e.g., word recognition, grammatical knowledge) and thus are more likely to capture many of the elements shared by the two processes than instruments that measure precursor or component skills in isolation.

Performance Measures (Outcome Measures)

An implicit assumption in the performance-based studies is that the performance measures (instruments) chosen for the study adequately and appropriately produce estimates of student ability that represent reading or writing ability or the component(s) of reading or writing that are of interest. Reading and writing ability are examples of constructs or *latent* abilities; that is, abilities that cannot be directly measured and thus must be measured using indirect means such as reading or writing assessments. In other words, student scores are treated as external manifestations of the literacy knowledge or process: the latent ability. For example, Shanahan (1984) examined reading ability by measuring student performance on four measures: vocabulary, phonics, reading comprehension, and sentence completion. In the same study, he examined eight writing measures: vocabulary diversity, number of episodes, number of story categories, and number of information units, spelling accuracy, phonemic

accuracy, visual accuracy, and average t-unit length. In his study, student performance on the four reading measures and eight writing measures were modeled as appropriate representations of a student's overall reading and writing ability. Thus, because reading and writing are latent abilities and student scores are interpreted as external manifestations of reading and writing ability, the choice of assessment used to produce estimates of the ability will affect the nature of the conclusions that can be drawn from the results.

The current research focuses on reading comprehension measures to represent reading ability and measures of writing quality to represent writing ability. As Engelhard (2001) stated, "One way to think about a reading test is to view it as an operational definition of a latent construct—reading" (p. 2). Writing tests (e.g., writing quality scores) can also be viewed as operational definitions of the latent variable—writing ability. For the current research, reading comprehension and writing quality measures were chosen to examine the reading-writing relationship because both abilities have precursory skills but are not typically viewed as precursors of something more complex or advanced. Just as successfully navigating down a ski slope involves multiple precursor skills and strategies ranging from simply balancing on skis to using poles and muscle strength to navigate effectively around various moving and stationary obstacles, both reading comprehension and writing are complex activities that rely on multiple skills and strategies. Precursory skiing skills are necessary for a successful run; reading skills are prerequisites for reading comprehension. Writing skills are prerequisites for an outcome of quality writing.

The view of reading comprehension taken in this research is one that incorporates a constellation of skills that a reader uses in the task of understanding text. Reading ability is defined as the ability to comprehend text. Because grade-appropriate texts are used in the

reading comprehension measures, for the purposes of this study the two terms are used interchangeably.

Pressley (2000) described the complexity of reading comprehension (i.e., text comprehension) in the following description of text comprehension:

> Text comprehension begins with decoding of words, processing of those words in relation to one another to understand the many small ideas in the text, and then, both unconsciously and consciously, operating on the ideas in the text to construct the overall meaning encoded in the text. Of course, the meaning constructed by the reader is a function of the ideas explicitly represented in the text and the reader's response to those ideas, responses that often depend greatly on the prior knowledge of the reader (Anderson & Pearson, 1984; Rosenblatt, 1978). The many active processes of reading—prediction, construction of images during reading, monitoring of comprehension and rereading, summarization, and interpretation—depend greatly on prior knowledge, with skilled reading being an articulation of prior knowledge and these active reading processes. (p. 551)

With just a few alterations, a very similar description could apply to the production of an essay: the meaning constructed by the writer is a function of the ideas explicitly represented in the text (using the writer's skill in utilizing words and presenting those words in relation to one another to convey ideas and construct the overall meaning of the text) and the reader's response to those ideas, which often depend greatly on the prior knowledge of the reader.

The relationship between the ability to make meaning by reading text (reading comprehension of grade-level text) and writing quality and how the relationship between them varies across grades and modes of writing is the focus of this research.

Reading Comprehension and Writing Quality Research

Many performance-based studies of the reading-writing connection use multiple measures of reading and writing and examine the correlations between these measures. Shanahan's (1984) study discussed earlier is a good example. These studies typically

consider correlations between various aspects of reading and writing to evaluate which relationships are strongest. When the study involves students who have advanced beyond kindergarten or first grade, formal instruments of reading comprehension and writing quality are often included in the array of instruments used. Shanahan's study used a test of passage comprehension as one of his reading instruments. Although two writing samples were obtained for each student, no overall or holistic writing quality score was assigned to the essays. Instead, they were analyzed in terms of vocabulary diversity, organizational structure (number of episodes), spelling, and t-unit length. T-unit length is defined as the average number of words per independent clause with all dependent clauses attached, and is often used as an index of syntactic complexity (Hunt, 1965). The correlation between passage comprehension and these various writing measures ranged from .21 (number of episodes per writing sample) to .58 (spelling) in grade 2 and from .19 (average t-unit length) to .56 (spelling) in grade 5. Shanahan's sample included 256 grade 2 students and 251 grade 5 students, so he was able to compare the relationship between reading and writing across grades. He found that the only significant ($p < .05$) difference in reading between the grades was in the relative importance of vocabulary knowledge. The knowledge of word meaning was more important to the reading process in grade 5 than in grade 2.

Shanahan's (1984) work is similar to the research described in the current study in that Shanahan studied the relationship between reading and writing across grades, though the current research analyzes five different grade levels rather than the two included in Shanahan's work. In an interesting extension of Shanahan's (1984) analysis, Shanahan and Lomax (1986) used Shanahan's 1984 data set and structural equation modeling (SEM) to compare and evaluate three alternative theoretical models of the reading-writing relationship.

Their analysis included three latent variables for reading (word analysis, vocabulary, and comprehension) and four for writing (spelling, vocabulary diversity, syntax, and story structure). Each latent variable was represented by one, two, or three indictor variables (i.e., test scores). In their writing-to-read model, all of the relations between the reading and writing latent variables emanated from the writing variables, suggesting that the writing factors influenced the reading factors and not the reverse. In their reading-to-write model, all of the relationships were reversed, with the reading factors modeled to influence the writing factors and not the reverse. Their third model was an interactive model in which some writing factors influenced the reading factors and vice versa. Shanahan and Lomax compared the theoretical models both within each grade and compared them across grade 2 and grade 5. The researchers found that the model that represented an interaction between reading and writing factors provided the best fit to the data in both grades. Shanahan and Lomax also found that the reading-to-write model fit the data better than the writing-to-read model, which suggests that reading has a relatively larger influence on writing factors than the influence of writing on reading factors. Shanahan and Lomax's use of SEM techniques to explore the relationship between reading within and across grades is similar to the use of SEM in the current research and represents one of the earliest examples of SEM techniques used to examine the reading-writing relationship.

A second study that used SEM techniques to explore the reading-writing relationship across grades is Abbott and Berninger's (1993) study of 600 students in grades 1-6. The researchers examined the relationship between various indicators of oral language/verbal reasoning, written language, and reading in an effort to describe the characteristics of developing writers. Abbott and Berninger used SEM to assess whether the relationship

between their latent structures was consistent across grades as well as the degree to which they varied. The study used three indicators for the reading factor, all from the *Woodcock Reading Mastery Test-Revised* (WRMT-R; Woodcock, 1987): two measures of word recognition in isolation and one measure of reading comprehension. The reading comprehension measure was the passage comprehension subtest of the WRMT-R, which is a cloze task. The writing instrument measures included measures of both compositional fluency (number of words produced and number of clauses produced) and compositional quality (scored on a 5-point scale) of timed writing in the narrative and expository modes. Interestingly, the correlations between the quality and fluency measures and the predictor (composition) were not similar enough to combine them in the same structural analysis. Subsequent analyses were conducted with writing quality and writing fluency separately, which provided a way to compare their relative relationship with reading comprehension across grades.

The writing quality measures were more strongly related than the fluency measures to passage comprehension. The correlations between passage comprehension and fluency measures were in the .08–.34 range in grades 1-3 and lower (.01–.20) for students in grades 4-6. The correlations between passage comprehension and quality measures of writing were higher, ranging from .22–.54 in grades 1-3 and from .24–.53 in grades 4-6.

Abbott and Berninger's (1993) SEM analysis revealed that, as in Shanahan's (1984) study, the contributions of some factors to measured writing ability varied by grade. For example, the researchers found that in first grade, both oral language/verbal reasoning and reading were significantly related to composition quality, but in grades 2 and 3, only reading had a significant incremental contribution to composition quality. As part of the analysis,

Abbott and Berninger examined whether the relationships among the variables (paths) were similar across grades. To do this, they fixed the paths between the oral language/verbal reasoning factor, the composition quality factor, and the reading factors to be equal across grades. For grades 1, 2, and 3, this approach significantly decreased the fit of the model. In particular, the researchers found that maintaining consistency in the pathway between reading and composition quality across grades was the largest contributor to the decreased fit of the model. This result is consistent with their finding of the changing (increasing) role of reading ability in composition quality. Similar results were found when the paths were constrained to be equal for grades 4, 5, and 6. However, the reduction in fit was not as large as seen in grades 1, 2, and 3, which suggests that the relationships among the factors is more stable in the intermediate grades.

In an extension of these analyses, Berninger et. al. (2002) studied the bidirectional relationships between the components of the SEM model used in Abbott and Berninger's (1993) study. Berninger et al. found that the relative influence of reading on writing was greater than the influence of writing on reading, a finding similar to that of Shanahan and Lomax's 1986 work. Berninger et al. also found that reading comprehension was a better predictor of compositional quality than it was of compositional fluency. Reading comprehension exerted a direct, significant influence on compositional quality at all grade levels (1-6) but on compositional fluency in grades 1, 2, 3, and 6 only. In addition, at all grade levels, the sizes of the path coefficients were larger for the path from reading comprehension to compositional quality than for the path from reading comprehension to compositional fluency. In contrast, the influence of compositional quality was significant in predicting reading comprehension only in grades 4-6. Thus, although reading comprehension

predicted compositional quality at all grade levels, compositional quality predicted reading comprehension at the upper three grades. The authors noted that this developmentally mediated asymmetry in the influences of reading and writing on each other supports the idea that writing and reading are not inverse processes. However, it is clear from Berninger et al.'s research that reading and writing are strongly related in all of the grades that they studied.

Heck and Crislip (2001) conducted another study that examined the relationship between reading comprehension and writing quality by examining the relationship between reading and direct and indirect measures of writing for a sample of 3,300 third graders. The direct writing measure used in their study was a score on the *Stanford Writing Assessment Edition 1* (Stanford Achievement Test, 1983) consisting of a single draft written in a 25-minute time period. The indirect writing sample used in this study came from the multiple-choice *Stanford Achievement Test* (ed. 8;Gardner, Rudman, Carlsen, & Merwin, 1985) language test, which consisted of language mechanics, language expression spelling, and listening. The reading measure used in this study also came from student scores on the *Stanford Achievement Test*. The reading score comprised word study skills, vocabulary, and reading comprehension. Students' reading scores correlated at .57 with their direct writing scores and .91 with their indirect writing scores.

Carrell and Connor (1991) also examined the relationship between reading and writing as part of their study of 33 English as a second language (ESL) college students' reading and writing ability, although it was not the primary focus of the research. The researchers examined student reading comprehension of descriptive and persuasive texts using multiple-choice comprehension questions and a recall task. The writing of descriptive and persuasive tasks was scored using a holistic scale as well as a qualitative scale that addressed the quality

of the description for descriptive essays and strength of the argument for persuasive essays. Carrell and Connor calculated Pearson product moment correlations between the reading and writing scores within each genre and found that only 3 of the 8 correlations were significant. The three significant correlations included persuasive reading multiple-choice scores and persuasive holistic writing scores ($r = .58$, $p < .001$), descriptive recall reading and descriptive holistic writing ($r = .45$, $p < .01$), and persuasive recall reading and persuasive holistic writing ($r = .48$, $p < .01$).

Although in all of these studies, the findings support a strong relationship between reading comprehension and writing, none of the studies focused exclusively on reading comprehension and writing quality. For example, in Heck and Crislip's (2001) study, the main focus was on writing. The reading comprehension piece of the student's reading score was just one element of the total reading score that was used to analyze the relationship between reading and writing. The current research focuses exclusively on reading comprehension and writing quality.

The Writing Task

A potentially confounding factor in all studies that use writing quality as an outcome measure of writing is the nature of the writing task used in the assessment. In particular, when the task involves writing an essay, as in the current research, many will wonder what impact, if any, the mode of the task had on the quality of the writing. Studies of student writing have shown that the mode of the task (e.g., narrative, expository, persuasive) has an impact on a student's writing performance (Engelhard et al., 1994; Kegley, 1986; Kuhlemeir et al., 1985; Quellmalz et al., 1982; Sachse, 1984). However, the results have not all

supported the relative ease of one type of discourse mode over another. For example, Sachse (1984) found that grade 4 students writing to narrative prompts on the *Texas Assessment of Basic Skills* (TABS) earned higher scores on average than students who wrote to either a descriptive or classificatory prompt. In another study, Kegley (1986) examined the scores obtained by 457 seventh-grade students who wrote to either a narrative, expository, descriptive, or persuasive writing prompt. She then matched students by their language score on the *California Test of Basic Skills* (CTBS) to create virtual students who had written to all four prompts. When she compared their results across writing modes, she found that, on average, the narrative essays received the highest scores and the persuasive essays received the lowest scores.

Engelhard et al. (1994) studied the scores on essays of 170,899 eighth-grade students who participated in statewide assessments of writing in Georgia in three consecutive years. Each student wrote an essay in response to either a narrative, descriptive, or expository prompt. Engelhard et al. found that essays written to the narrative prompts received the highest scores followed by descriptive and last, expository prompts. The differences between the mean ratings were significant in each case for ratings on the topic development scale (content/organization and style). The ratings for conventions (sentence formation, usage, and mechanics) were not significantly different between the narrative and descriptive essays. However, each student wrote to only one prompt, so this study does not provide information about how the same student responds to prompts for essays in different discourse modes.

In contrast to the findings that show higher scores for essays written to narrative prompts, in a study of 11th and 12th grade students, Quellmalz et al. (1982) found that students who wrote to a narrative prompt scored lower than students who wrote to an expository prompt.

The researchers suggested that this finding may have been due to the increased emphasis on expository writing in high school writing instruction and the accompanying reduction in narrative writing. In Carrell and Connor's (1991) study of the relationship between reading and writing descriptive and persuasive texts, each essay was given a holistic score and a qualitative score. The holistic score used a 6-point scale that addressed organization and development, appropriateness of details, unity and coherence, facility with language, and syntactic variety and word choice. The qualitative scores looked at the quality of the description for the descriptive essays and the strength of the argument for the persuasive essays. Carrell and Connor found no differences in the difficulty of the modes based on holistic scores of students' written essays, but did find significant differences ($p < .05$) between the students' qualitative scores. Carrell and Connor also investigated the genre differences as a function of students' language ability, measured by the *Michigan Test of English Language Proficiency* (MTELP). When language ability was taken into account, genre differences did not account for a significant portion of the remaining variance of either holistic or qualitative scores. Although the sample size used in this research ws small (N = 33), the results suggest that the mode of the writing task is not a dominant factor in determining a student's writing score.

Overall, these studies suggest that the question about the relative difficulty of the discourse mode has not been answered definitively. The current study will contribute to an understanding of the relative difficulty of the discourse mode by examining the quality of student writing in the informational, narrative, and persuasive modes. Notably, the research reported here uses data from two essays in each of the three discourse modes, for a total of

six essay scores for each student. The quantity of the writing data for each individual student represents an improvement over previous research on discourse modes.

The Writing Score

Assigning a score on a writing assessment that uses extended responses such as essays rather than multiple-choice options is complicated by the presence of human judgment in the generation of the score. Unlike an objectively scored, multiple-choice assessment, the essay score involves human raters who judge the quality of the writing. Although the actual essay scored by multiple raters is identical, a student may receive different scores from different raters. When the goal is a reliable estimate of a student's writing ability, this potential variation in scores across raters is problematic.

Ideally, the ratings given to student essays would be the same regardless of which raters scored them. In reality, raters have been found to vary in their relative severity or leniency in assigning scores to essays (Engelhard, 1994, 2000; Gyagenda & Engelhard, 1998; Hollenbeck et al., 1999). Thus a student whose essay was scored by a severe rater would be at a disadvantage compared to a student whose essay was scored by a lenient rater; the severe rater would be more likely to assign a lower score to the essay than the lenient rater. When scores on student essays are used to make educational decisions, the variation introduced by raters has implications for those decisions. For example, a student whose essay is scored by one or more severe raters may be inappropriately assigned to a low performance level and receive unneeded writing remediation. In contrast, a student who has the same level of writing skill but whose essay is scored by more lenient raters, may be assigned to a higher performance level and be ineligible for needed writing assistance. Because test scores

frequently have instructional implications for students, every effort should be made to account for score variation resulting from intervening variables that do not reflect a student's true writing ability.

Engelhard (1992) described a measurement model that can take into account the relative severity or leniency of the raters who score the essays. According to Engelhard (1992),

> Another way to think about the FACETS model is to view it as an equating model with the raters viewed as analogous to test forms that may vary in difficulty; because different students are rated by different raters, it may be necessary to "equate" or statistically adjust for differences in rater severity. (p. 96)

The student's overall estimate of writing ability score can be adjusted based on the combination of raters who scored that student's essays. This model, the many-faceted Rasch (FACETS) model, developed by Linacre (1988), is an extension of the Rasch measurement model (Andrich, 1988; Rasch, 1980; Wright & Masters, 1982).

The Rasch measurement model, also known as the 1-parameter model in item response theory (IRT), has several desirable measurement characteristics, one of which is that ability estimates are reported on an equal-interval scale (Bond, 2007; Hambleton, Swaminathan, & Rogers, 1991). On an equal-interval scale, each score point increase represents the same amount of increase in the amount of the ability being measured by the instrument regardless of the point on the scale. That is, the amount of ability increase needed to move from point 1.1 to point 1.2 on the scale is the same amount as is required to move from point 2.3 to point 2.4. In contrast, the amount of writing ability needed to move from 1 to 2 on a raw score point scale may be very different from the amount needed to move from a 3 to a 4 on the same scale. Using IRT to estimate a person's writing ability rather than simply the raters' raw scores enables more precise ordering of the persons along a continuum. This is an additional benefit of using FACETS to produce examinee writing scores.

The many-faceted Rasch model can be used for measurements that include multiple *facets*. *Facets* are elements of the measurement (e.g., writing) situation that influence the observed rating of a student's essay. The *facets* analyzed in the research reported here are student writing ability, rater severity, and difficulty of the discourse mode. Other elements of the writing situation may affect the observed rating of a student's essay but are not addressed in the current research. Individual student characteristics also may affect a student's writing ability and thus the observed score. These student characteristics can include race, gender, age, social class, and others. In keeping with the focus on the three primary facets, the measurement model underlying the current research is shown in Figure 2.1.

In this model, both the relative severity of the raters and the relative difficulty of the discourse modes are viewed as intervening variables that may affect the observed ratings of student writing quality. However, only one—rater severity—is included in the FACETS calculations of the observed writing ability score. The degree to which the relative difficulty of the discourse mode affects the observed writing quality scores is a primary focus of this research, and is examined in detail using structural equation analysis as part of this research study. The FACETS analysis used to obtain the writing quality score for students does not include discourse mode as a facet because subsequent analysis requires any variation due to discourse mode to be retained in the student scores.

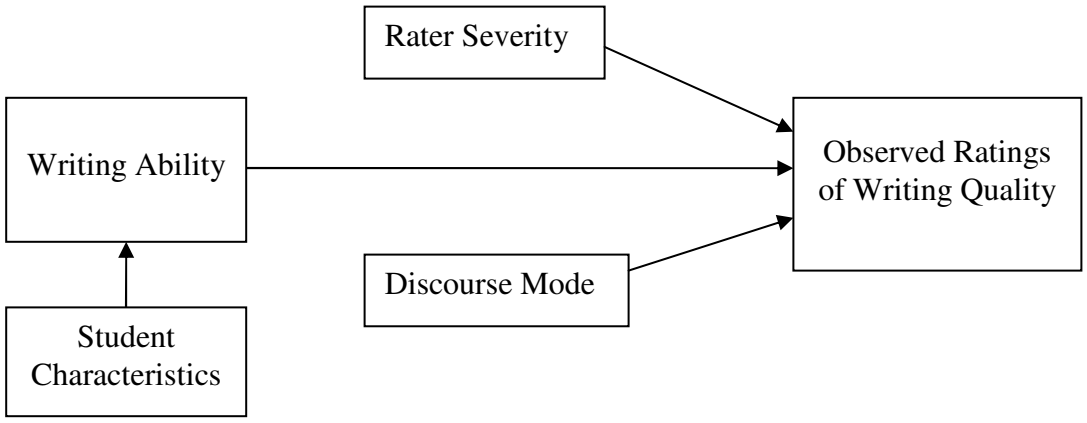

Figure 2.1. Measurement model for the assessment of writing quality.

An overall writing score for the student as well as scores for each separate essay can be obtained from FACETS. The scores obtained from FACETS for each essay were used in the subsequent analyses.

Developmental Framework/Progression

In her extensive process-based study referred to previously, Langer (1986) discovered that although reading and writing behaviors are complex and vary with age and difficulty of the task, the behaviors varied consistently between reading and writing. The current research provides the opportunity to add to Langer's work by examining the degree to which the relationship between reading comprehension and writing quality vary by grade level and by writing task.

Fitzgerald and Shanahan (2000) provided a framework for the development of the shared components of reading and writing. The framework is organized around the four basic types of knowledge that readers and writers share described earlier and listed below:

(1) Metaknowledge,

(2) Domain knowledge about substance and content (developing knowledge of the world),

(3) Knowledge about universal text attributes, and

(4) Procedural knowledge.

Fitzgerald and Shanahan (2000) suggested that although the nature of what is learned by readers and writers changes as the person develops, the changes occur on the same developmental track. As individuals develop as readers and writers, they move from one developmental stage to the next, and the elements of each of the four types of knowledge that they must learn change. Fitzgerald and Shanahan posited six developmental stages in reading and writing development, each with a different set of competencies that the individual is mastering within each of the four basic types of knowledge. For example, in the first stage, literacy roots, the procedural knowledge a child learns is the concept of a book. In the fifth stage, the student learns how to see from another's viewpoint and how to analyze and critique. Table 2.1 shows the six stages and corresponding elements of the four knowledge categories.

Different writing modes may tap into different aspects of this hypothesized developmental progression. For example, in Stage 4, students in grades 4-8 begin to use reading and writing to learn. At this stage, as they begin to utilize information text differently, and perhaps when they begin to be exposed to more informational text, the quality of their writing in the informative modes may improve (Kamberelis & Bovino, 1999). Similarly, in Stage 5, students begin to know how to see from another's viewpoint. Reaching this stage of development may have an impact on student's writing performance in the

persuasive mode, since effective persuasion requires skill in addressing issues from others' views. The developmental framework proposed by Fitzgerald and Shanahan (2000) provides a lens through which to examine variations in the relative quality of student writing in the three discourse modes across grades.

Table 2.1

Critical Knowledge at Various Stages of Reading and Writing

	Metaknowledge	Domain Knowledge about Substance and Content	Knowledge about Universal Text Attributes	Procedural Knowledge
Stage 1 Literacy roots Birth-age 6	Knowing about functions and purposes of reading and writing		Graphophonics: phonological awareness, grapheme awareness	Concepts of book
Stage 2 Initial literacy Grades 1-2 Ages 6-7	Knowing that readers and writers interact Monitoring for meaning Monitoring word making		Graphophonics: letter and word making, grapheme awareness, morphology Syntax of sentences	How to use strategies to read and write words

Table 2.1, Cont'd.

	Metaknowledge	Domain Knowledge about Substance and Content	Knowledge about Universal Text Attributes	Procedural Knowledge
Stage 3 Confirmation, fluency, ungluing from print Grades 2-3 Ages 7-8			Higher level graphophonics: morphology—big words	Instantiating smooth integration of all strategies and processes How to "make" big words
Stage 4 Reading and writing for learning the new: a first step Grades 4-8 Ages 9-13	Meta-comprehension	**(prior knowledge: using reading and writing to learn)** Semantics: Vocabulary meaning Meaning created through context of connected text	Syntax Of sentences Of larger chunks of text (i.e. text structures)	Knowing how to create and use word meanings Knowing how to create and use meaningful connected text
Stage 5 Multiple viewpoints High School Ages 14-18	Meta-comprehension	Semantics: Vocabulary meaning Meaning created through context of connected text	Syntax of chunks of text (e.g. text structure)	**Knowing how to see from another's viewpoint** **Knowing how to analyze and critique**
Stage 6 Construction and reconstruction: a worldview College Age 18 and above	Metaknowledge about how readers and writers interact			Knowing how to see from another's viewpoint Knowing how to analyze and critique

Summary

The research conducted on the nature of the reading-writing relationship has shown that that although reading and writing have processes, skills, and knowledge bases in common, performance measures typically do not capture large portions of the variance in the scores—typically around 25% (Fitzgerald & Shanahan, 2000). However, these studies have not focused on reading comprehension and writing quality, with particular attention paid to the

instruments used to estimate student ability in both areas. The current research addresses many of the shortcomings of previous performance–based research. The reading and writing measures used in the current research are holistic measures rather than precursory components of either process, so it is likely that a greater proportion of the shared relations between reading and writing will be captured.

Research question 2 directly addresses the relationship between discourse mode and the writing quality scores. Because this research uses two samples of each discourse mode and a total of six essays for each student, the data provide a deeper look at the issue of variation across discourse modes than does much previous research. The multiple measures used for reading comprehension and writing quality also provide a more robust measure of each ability than those used in much of the previous research. In addition, the sample size used in the research reported here addresses the lack of power evident in some of the previous research; the current study examines more than 500 students across six grades (4, 6, 8, 10, 12), and thus provides information about the reading-writing relation across grades. Most important, the same instruments for reading comprehension and writing quality are used across grades. This consistency in instruments makes the grade-grade comparisons more meaningful than if different instruments had been used across grades.

CHAPTER 3

METHODS

This study used data collected from a sample of 521 students in grades 4, 6, 8 10, and 12 from one school district in a small town in the northeast corner of Mississippi. For all grades, the writing data were collected in November and December, 2005 and the reading data were collected in January, 2006. At all grade levels, students responded to two writing prompts each week for a total of six essays over three weeks. Less than two months later, the same students were administered two reading tests within a 3-week period. The student essays were scored by human raters from a professional organization that specializes in large-scale essay scoring for high-stakes testing programs. Each essay was scored by four separate raters, and the average of their scores was used as the total score for each essay. The reading tests were dichotomously scored, multiple-choice tests. The following section provides additional description of the participants in the study, the instruments and scoring procedures, and the data analysis plan.

Participants

The student sample consisted of 521 students in grades 4, 6, 8, 10, and 12 from one school district in Mississippi. All of the 10th- and 12th-grade students attended the same high school and all of the 6th- and 8th-grade students attended the same middle school. The 4th-grade students attended one of two elementary schools that participated in the study. The 4th- and 6th-grade samples were obtained from seven different teachers in each grade, each of

whom had one full class of students participate in the study. The 8th-, 10th-, and 12th-grade students were enrolled in classes taught by one grade-level teacher.

In an effort to enhance the confidence in inferences that could be made about the relationship between reading and writing from this data set, a decision was made to include in the sample of students only those who had scores on both reading tests and all six essays. Table 3.1 summarizes the student demographic data for the sample (N = 521) and provides a comparison to state and national populations.

Table 3.1

Demographic Information by Grade with Comparisons to District, State and National Averages

	Gender Number (%)		Ethnicity[1,2] Number (%)				
Grade	Male	Female	W	B	A	H	O
4 (n = 115)	65 (57)	50(43)	49(43)	54(47)	2(2)	5(4)	0(0)
6 (n = 122)	54(44)	68(56)	61(50)	53(43)	1(1)	3(2)	2(2)
8 (n = 106)	53(50)	53(50)	47(44)	54(51)	0(0)	5(5)	0(0)
10 (n = 85)	33(39)	52(61)	53(62)	39(36)	0(0)	1(1)	0(0)
12 (n = 93)	46(49)	47(51)	63(68)	27(29)	0(0)	2(2)	1(1)
Total (N = 521)	251(48)	270(52)	227(44)	273(52)	3(1)	16(3)	3(1)
District[3]	(51)	(49)	(54)	(42)	(1)	(3)	(0)
Mississippi[4,5]	(48)	(52)	(47)	(51)	(1)	(1)	(.2)
Nation[4,5]	(51)	(49)	(57)	(17)	(5)	(20)	(1)

[1]Percentages do not add to 100 due to missing data
[2]W = White; B = Black; A =Asian; H =Hispanic; O =Other
[3]Obtained from the Mississippi Department of Education (2007)
[4]Estimated from the *National Assessment of Education Progress: Reading* (2005)
[5]Estimated from the *Digest of Education Statistics* (2005)

The socioeconomic status of the students in the school district from which the sample population was taken is similar to that of Mississippi's overall student population, with 71% of the district's student population and 70.6% of the state's population qualifying as economically disadvantaged. However, based on student test scores, the academic

achievement of students in the district in the 2005-2006 school year was higher than the statewide average. Student scores on the Mississippi Curriculum Test, the standardized reading assessment administered to students in grades 2-8, were approximately 10% higher for students in the district than for students statewide. In grade 10, the difference was not as large. Statewide 78% of the students passed the English II Subject Area Test and 83.5% of the district students passed (Mississippi Department of Education, 2007).

Instruments and Scoring Procedures

The instruments used to collect data on student reading and writing ability for this study included two administrations of alternate forms of a reading comprehension test and six administrations of a standardized writing test. The reading instrument and scoring procedures are described first, followed by a description of the writing instrument and the scoring procedures used for writing.

Reading Comprehension

Administration of alternate forms of the *Scholastic Reading Inventory* (SRI; Scholastic, Inc., 1999) was used to obtain estimates of student reading ability. Each grade was administered the Form A and Form B versions of the same level of test. Forms A and B are parallel forms that can be used interchangeably. The two alternate reading comprehension test forms were administered to provide a more robust measure of student reading ability. Grade 4 was administered level 14, grade 6 was administered level 16, grade 8 was administered level 17, and grades 10 and 12 were both administered level 18. Each level of the SRI administered to a grade contained reading material at a difficulty level appropriate for the tested grade. The reported reliability of the test forms ranges from $r_{xx'}$ .83 in grade 3 to $r_{xx'}$ .90 in grade 10 (Scholastic, Inc., 2007).

The SRI measures reading ability by employing an item type that is well suited to measure the constellation of skills required for reading. According to the SRT Technical Guide (2007),

> The SRI is designed to measure how well readers comprehend literary and expository texts. It measures reading comprehension by focusing on the skills readers use to understand written materials sampled from various comtent areas. These skills include referring to details in the passage, drawing conclusions, and making comparisons and generalizations. SRI does not require prior knowledge of ideas beyond the test passages, vocabulary taken out of context, or formal logic. (p. 8)

In addition, vocabulary comprehension skills and an array of precursor reading skills are accessed by the reader to correctly answer the test items. The SRI is a multiple-choice test. Each item consists of a short paragraph, typically between 40 and 125 words, followed by a sentence with one missing word. The student must fill in the missing word with 1 of 4 choices that are all semantically correct completions of the sentence. Only one of the choices is correct, given the context of the paragraph. As a result, the student must be able to comprehend the paragraph to choose the correct word that completes the sentence below the paragraph. An example of the type of item found on the SRI test (SRI Educator's Guide, 2003) is shown in Figure 3.1.

> Richard tried hard to please his father. He fixed their food and washed the dishes. He mowed the lawn and trimmed the bushes. He even sewed on buttons and mended their clothes. Carol had taught him to do many things.
>
> He was _____
>
> A. old
> B. slow
> C. busy*
> D. hot

Figure 3.1. Example of an SRI test item.

A benefit of using multiple-choice items to assess reading comprehension is that the format avoids the pitfall of confounding reading and writing (Jenkins, Johnson, & Hileman, 2004). In addition, because each item addresses a separate piece of text, issues of interitem dependencies are minimized (Thissen, Nelson, Rosa, & McLeod, 2001). These desirable features, in addition to the use of grade-level text and a design that minimizes the ability for a reader to guess the correct answer simply by reading the completion sentence, makes the SRI item format an excellent tool for measuring reading ability.

The Scholastic Reading Inventory is based upon the Lexile Framework for Reading® and comprises the same item type that was used to develop the Lexile Framework for Reading and validate the test score interpretations. The Lexile Framework has been linked to many widely-known reading tests, and the test scores are highly correlated. For example, the correlation between the Standord Achievement Tests (Tenth Edition) and the Lexile Framework is .93; the correlation between the Gates-MacGinintie Reading Test (Version 4.0) and the Lexile Framework is .92 (SRI Technical Guide, 2007, p. 19). These strong correlations provide evidence that the SRI is also a reading test, because it utilizes the same item type as the Lexile Framework for Reading and was developed in collaboration with the creators of the Lexile Framework for Reading.

The text for the reading comprehension passages on the SRI comes from published works in both fiction and nonfiction. The paragraph texts were not selected to represent a specific discourse mode, though they may contain elements of one or more modes. Thus, the reading comprehension tests as a whole can be viewed as somewhat mode-neutral. As such, they serve as an unbiased example against which to compare writing across the discourse modes.

All levels of the SRI forms have the same basic structure: one paragraph followed by a sentence with one word missing. The levels (i.e., 14, 16, 17, and 18) differ only in the difficulty of the text and answer choices. A deliberate decision was made to keep the format of the assessment instrument consistent across grade levels. Because the focus of this research is on the relationship between reading and writing, with a focus on writing mode, it is important to minimize any confounding elements by keeping the reading instrument consistent across grades.

An examinee's ability score on each test was estimated using the Rasch model (1-parameter IRT model) with the program Winsteps version 3.6 (Linacre, 2007). The ability estimates are reported on a logit scale, which comprises equal-interval units. The scale used in this study ran from approximately -6 to +6, with higher values representing higher levels of reading comprehension ability. Reporting in a scale with equal interval units means that the difference in reading ability between the scores 1.3 and 1.4 is the same as the amount of difference in reading ability between the scores -.2 and -.3, greatly simplifying relative score interpretations. The average ability estimate in logits for the two tests was used as the overall estimate of reading ability for each student. Each participant was administered the two tests within a 2-week window in January, 2006. The alternate forms reliability estimate for the two administrations ($r_{xx'}$ = .62 - .85) was used as the reliability estimate for the reading comprehension measure.

Writing Quality

Estimates of writing quality were obtained by administering six essay prompts to each student. The prompts administered were released prompts from the National Assessment of Educational Progress writing assessment from the 1998 and 2002 administrations (National

Center for Education Statistics, 2005). The six prompts represent three types of discourse mode: narrative (N), informative (I), and persuasive (P). Descriptions of each discourse mode as defined by NAEP are shown in Table 3.2.

Table 3.2

Discourse Mode Descriptions

Discourse Mode	Discourse Mode Description
Narrative	Narrative writing involves the production of stories or personal essays. Practice with these forms helps writers to develop an ear for language. Also, informative and persuasive writing can benefit from many of the strategies used in narrative writing. For example, there must be an effective ordering of events when relating an incident as part of a report. Sometimes narrative writing contributes to an awareness of the world as the writer creates, manipulates, and interprets reality. Such writing—whether fact or fiction, personal essay, or creative narrative—requires close observation of people, objects, and places. Further, this type of writing fosters creativity, imagination, and speculation by allowing the writer to express thoughts and then stand back, as a more detached observer might, and grasp more fully what is being felt and why. Thus, narrative writing offers a special opportunity to analyze and understand emotions and actions.
Informative	Informative writing focuses primarily on the subject-matter element in communication. This type of writing is used to share knowledge and to convey messages, instructions, and ideas. Like all writing, informative writing may be filtered through the writer's impressions, understanding, and feelings. Used as a means of exploration, informative writing helps both the writer and the reader to learn new ideas and to reexamine old conclusions. Informative writing may also involve reporting on events or experiences, or analyzing concepts and relationships, including developing hypotheses and generalizations. Any of these types of informative writing can be based on the writer's personal knowledge and experience or on information newly presented to the writer that must be understood in order to complete a task. Usually, informative writing involves a mix of the familiar and the new, and both are clarified in the process of writing. Depending on the task, writing based on either personal experience or factual information may span the range of thinking skills from recall to analysis to evaluation.

Table 3.2, Cont'd

Discourse Mode	Discourse Mode Description
Persuasive	Persuasive writing emphasizes the reader. Its primary aim is to influence others to take some action or to bring about change. Persuasive writing may contain much information—facts, details, examples, comparisons, statistics, or anecdotes. Its main purpose, however, is not simply to inform but to persuade. This type of writing involves a clear awareness of what arguments might most affect the audience being addressed. Writing persuasively also requires the use of critical thinking skills such as analysis, inference, synthesis, and evaluation. Persuasive writing is called for in a variety of situations. It may involve responding to a request for advice by giving an opinion and providing sound reasons to support it. It may also involve presenting an argument in such a way that a particular audience will find it convincing. When there is opposition, persuasive writing may entail refuting arguments that are contrary to the writer's point of view. In all persuasive writing, authors must choose the approach they will use. They may, for instance, use emotional or logical appeals or an accommodating or demanding tone. Regardless of the situation or approach, persuasive writers must be concerned with having a particular desired effect upon their readers, beyond merely adding to knowledge of the topic presented.

The prompts used for NAEP assessments are good examples of the types of writing typically asked of students in the three discourse modes. An example of a prompt for each mode for grade 8 is shown in Table 3.3. The complete set of prompts used in the study is provided in Appendix A.

Table 3.3

Example Grade 8 Prompt for Each Discourse Mode

Mode	Prompt
Narrative (8AN)	Imagine this situation! A noise outside awakens you one night. You look out the window and see a spaceship. The door of the spaceship opens, and out walks a space creature. What does the creature look like? What do you do? Write a story about what happens next.
Informational (8AI)	A public television network is seeking ideas for a new series of shows that would be educational for teenagers. The series will include ten one-hour episodes and will be shown once a week. Some of the titles under consideration are: "Great Cities of the World" "Women in History" "Nature Walks" "American Legends" Choose one of these titles. Write a letter to the network president describing your ideas for a new educational series. In your letter, describe what one episode might be like. Use specific examples of what information you would include in the episode so the network president will be able to imagine what the series would be like.
Persuasive (8AP)	Many people think that students are not learning enough in school. They want to shorten most school vacations and make students spend more of the year in school. Other people think that lengthening the school year and shortening vacations is a bad idea because students use their vacations to learn important things outside of school. What is your opinion? Write a letter to your school board either in favor or against lengthening the school year. Give specific reasons to support your opinion that will convince the school board to agree with you.

NAEP administers writing tests at grades 4, 8, and 12. Because the study design included students in grades 6 and 10, students in these grades were administered NAEP prompts from the grade above and the grade below. For example, students in grade 6 were administered three prompts (one each of narrative, informative, and persuasive) from grade 4

and three prompts (one each of narrative, informative, and persuasive) from grade 8. The administration design is shown in Table 3.4. Each discourse mode (narrative, informative, persuasive) is represented by two prompts: A and B.

Table 3.4

Prompt Administration Design by Grade

Grade	Prompt*					
4	4AN	4AI	4AP	4BN	4BI	4BP
6	4BN	4BI	4BP	8AN	8AI	8AP
8	8AN	8AI	8AP	8BN	8BI	8BP
10	8BN	8BI	8BP	12AN	12AI	12AP
12	12AN	12AI	12AP	12BN	12BI	12BP

* Leading digit for each cell entry indicates grade level of the NAEP prompt administered; the first alphabetic character (A or B) indicates which of the two possible prompts was administered; the final character indicates the type of prompt (N = narrative; I = informative; P = persuasive).

Each participant responded to two prompts per week, on Tuesday and Thursday, for three weeks in the fall of 2005. Two weeks fell before the week of Thanksgiving and the third fell the week after Thanksgiving. The prompts were randomized to negate ordering effects. On any given day, each class of participants responded to a different prompt. The prompts were randomized within grade to ensure that, as a whole, student writing scores for a specific mode were not biased due to learning effects, student fatigue, lack of motivation toward the end of the assessment period, or other unknown factors. The administration order design is shown in Table 3.5.

Table 3.5

Distribution of Prompts by Day within Grade

Class	Day 1	Day 2	Day 3	Day 4	Day 5	Day 6
1	AN	AI	AP	BN	BI	BP
2	BN	BP	BI	AN	AP	AI
3	BP	BI	BN	AP	AI	AN
4	AP	AN	AI	BP	BN	BI
5	AI	AP	AN	BI	BP	BN
6	BI	BN	BP	AI	AN	AP

Note. In grades 6 and 10 the lower grade prompt was treated as the A prompt in the randomization design.

Professional essay raters with extensive essay scoring experience from an established testing company scored all of the essays. Each essay was scored by four separate human raters and the average score was used as the student score for each essay. There were a total of 19 raters divided into groups that scored specific modes. To ensure that the writing prompt and scoring process were well aligned, the rubrics used to score the NAEP essays were also used to score the essays in this study. The essays were scored according to the grade of the prompt, so all grade 4 essays were scored together using the grade 4 rubrics. Likewise, the grade 8 essays and the grade 12 essays were scored as separate groups. The ratings on the NAEP rubrics range from 1-6 with the score points labeled as follows: Unsatisfactory (1), Insufficient (2), Uneven (3), Sufficient (4), Skillful (5), and Excellent (6). An example rubric for the grade 8 persuasive essay is given in Table 3.6. The full set of rubrics is provided in Appendix B.

Table 3.6

Rubric for the Grade 8 Persuasive Essay

Score Point	Description
Excellent (6)	• Takes a clear position and develops it consistently with well-chosen reasons and/or examples across the response. • Well organized with strong transitions. • Sustains variety in sentence structure and exhibits good word choice. • Errors in grammar, spelling, and punctuation are few and do not interfere with understanding.
Skillful (5)	• Takes a clear position and develops it with reasons and/or examples in parts of the response. • Clearly organized but may lack some transitions and/or have occasional lapses in continuity. • Exhibits some variety in sentence structure and some good word choices. • Errors in grammar, spelling, and punctuation do not interfere with understanding.
Sufficient (4)	• Takes a clear position and supports it with some reasons and/or examples. • Organized with ideas that are generally related, but there are few or no transitions. • Exhibits control over sentence boundaries and sentence structure, but sentences and word choice may be simple and unvaried. • Errors in grammar, spelling, and punctuation do not interfere with understanding.
Uneven (3)	May be characterized by one or more of the following: • Takes a position and offers support, but may be unclear, repetitive, list-like, or undeveloped. • Unevenly organized; the response may be disjointed. • Exhibits uneven control over sentence boundaries and sentence structure; may have some inaccurate word choices. • Errors in grammar, spelling, and punctuation sometimes interfere with understanding.
Insufficient (2)	May be characterized by one or more of the following: • Takes a position, but may be very unclear, very undeveloped, or very repetitive. • Very disorganized; thoughts are tenuously connected OR the response is too brief to detect organization. • Minimal control over sentence boundaries and sentence structure; word choice may often be inaccurate. • Errors in grammar or usage (such as missing words or incorrect word use or word order), spelling, and punctuation interfere with understanding in much of the response.
Unsatisfactory (1)	May be characterized by one or more of the following: • Attempts to take a position (addresses topic) but is incoherent OR takes a position but provides no support; may only paraphrase the prompt. • Has no apparent organization OR consists of a single statement. • Minimal or no control over sentence boundaries and sentence structure; word choice may be inaccurate in much or all of the response. • A multiplicity of errors in grammar or usage (such as missing words or incorrect word use or word order), spelling, and punctuation severely impedes understanding across the response.

Data Analysis

This section provides information on each research question and the data analytic procedures used to address the question. The study addressed the overall relationship between reading and writing within grades and across grades using correlational analysis as well as SEM. Unadjusted correlations were used to examine the relationship between student scores for reading and writing. The construct relationship between reading and writing was examined with the disattenuated correlations between reading and writing within grades. Particular attention was paid to the writing scores, and the many-facet Rasch (FACETS) model was employed to adjust for variation in student scores due to rater severity. An initial simple SEM model of the reading-writing relationship was used to evaluate the degree of relationship between reading and writing across grades as it related to the discourse mode of the writing prompt. Next, the measurement models proposed for writing were adjusted to incorporate writing in all three discourse modes and evaluated using confirmatory factor analysis (CFA) in SEM. Alternate models were evaluated and analyzed according to alignment with previous research findings and statistical fit. The final model chosen was used to compare the reading-writing relationship across grades, with a close examination of how factor loadings varied across grades, gender, and race.

A procedure using the many-faceted Rasch (FACETS) model (Linacre, 1988) was used to calibrate the raters and produce scores for the student essays that were adjusted for the relative severity of the raters. The FACETS procedure produces writing scores that are more accurate measures of a student's writing ability than are raw scores unadjusted for rater variation. An additional benefit of using FACETS is that the writing scores are reported on

an interval scale rather than an ordinal scale. All analyses, with the exception of the GENOVA analysis, were conducted with scores adjusted for rater severity.

Research question 1 concerned the overall relationship between reading comprehension and writing quality scores. This question was addressed using correlational analysis. To begin the analysis for Research question 1, the unadjusted correlations between the average of the two reading scores and the average of the six writing scores were calculated for each grade level (4, 6, 8, 10, 12) and compared across grades. However, because the interest of this research is in the reading comprehension and writing quality constructs generally and is not limited to only the test scores produced for this study, the disattenuated correlations between reading and writing were also calculated.

The disattenuated correlation calculation addresses the fact that when two sets of measures are correlated, measurement error lowers the observed correlation coefficient below what it would likely be if the measurements of reading and writing were made without error (in other words, if the instruments used to measure reading and writing yielded perfectly reliable scores). The reliabilities of the reading and writing instruments are the proportion of observed variance not due to measurement error. Measurement error can be removed from the correlation coefficient for the measures using the formula proposed by Spearman (1904) and shown below.

$$r_{TxTy} = \frac{r_{xy}}{\sqrt{r_{xx'} r_{yy'}}} \qquad (1)$$

Where
r_{TxTy} = the disattenuated correlation between reading and writing,
r_{xy} = the observed correlation between reading and writing,
$r_{xx'}$ = the reliability of the reading measure, and
$r_{yy'}$ = the reliability of the writing measure.

The reliabilities of the reading comprehension scores used in the calculations to disattenuate the reading-writing correlation coefficients were obtained through the alternate forms reliability coefficients for the two reading test administrations. The reliabilities of the writing quality scores were obtained using a generalizability analysis (G study and d study) conducted separately for each grade level. The program GENOVA (Crick & Brennan, 1983) was used to conduct the generalizability analysis.

The generalizability analysis followed a 2-facet, nested design. The students (persons) were the object of measurement (*p*), and the two facets were items/essays (*i*) and raters/judges (*j*). The items were nested within the judges because discrete groups of judges evaluated essays written to specific modes. The item and judge facets were both crossed with persons because each person wrote an essay in response to all of the same prompts within each grade and each person was evaluated by each group of judges. This design is represented by *p* x (*j:i*; Shavelson & Webb, 1991). The generalizability analysis produces a coefficient called the generalizability coefficient (G-coefficient), which is analogous to a reliability coefficient in classical test theory (Brennan, 2001). The G-coefficient was used as the reliability coefficient for the writing scores.

Research question 2a concerned the relationship between the discourse modes and their observed difficulties. This question was also addressed with a correlational analysis. Correlations between scores on each of the writing modes as well as means and standard deviations were calculated for each grade level and compared. This analysis provided an initial look at whether student writing quality scores show similar patterns across discourse modes. More in-depth analysis of the relationship between writing and the discourse modes is a focus of the subsequent analyses.

Reading-Writing Relationship and the Discourse Mode

Research question 2 addresses the relationship between reading and writing, with a particular focus on the discourse mode of the writing prompt. To evaluate the degree to which the relationship between reading and writing is affected by the discourse mode used to assess writing quality, a simple structural equation model was developed. This model is shown in Figure 3.2. It represents the hypothesis that reading comprehension (reading) and writing quality (writing) are latent constructs and that the reading comprehension tests are indicators of the reading comprehension construct and essay scores are indicators of the writing quality construct. In this simple model, essays in only one discourse mode are modeled as indicators of writing so that the relative impact of the discourse mode used as the indicator variable on the reading-writing relationship can be evaluated.

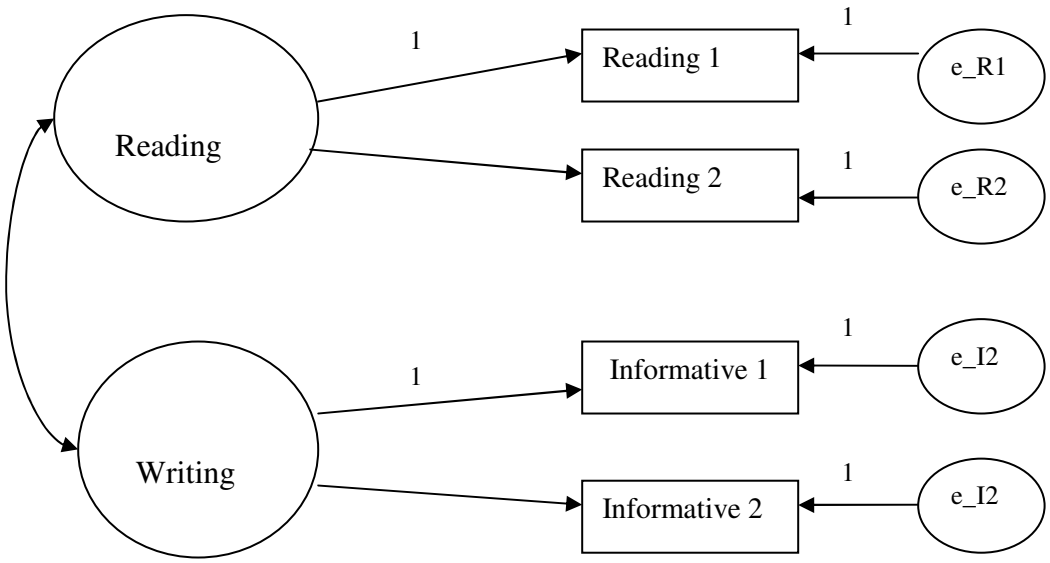

Figure 3.2. Simple model of the reading-writing relationship.

The latent construct is represented by an oval shape and the indicators are represented by rectangles. The double-sided arrow pointing from reading to writing represents the relationship between the two latent constructs. The one-sided arrows pointing from the latent

constructs to the indicators represent the presumed direct causal effect of the latent variable (e.g., writing) on the observed measure (e.g., essay score). The statistical estimates of these direct effects are called *factor loadings*. In CFA, the factor loadings are generally interpreted as regression coefficients that may be in unstandardized (covariance) or standardized (correlation) form. Indicators assumed to be caused by latent variables are called *effect indicators* (Kline, 1998). Finally, an arrow points to the indicator/essay term from the measurement error term (e); the error term represents all sources of variation found in the indicator (e.g., essay score) that are not explained by the indicator's latent variable. Values of 1 are assigned to some of the pathways to set a scale (i.e., point of reference) for the loadings. The scale for the error terms are typically set to 1 as is one loading for each group of loadings for one factor.

The model in Figure 3.2 was used to evaluate whether changing the writing indicator variables (mode of the essay) had an effect on the relationship between the reading comprehension and writing quality latent constructs. For each grade, the standardized estimates (correlation) between reading and writing for each grade were obtained for the three separate writing modes. For subsequent analyses, scores from all six essays, two in each mode, were used as the indicator variables for the writing quality construct.

Confirmatory Factor Analyses

Prior to conducting further analysis of the reading-writing connection, it was necessary to determine the best way to model writing with all six essay scores. Research question 2b investigated whether writing performance in the three modes (narrative, informational, persuasive) is divergent enough that their relationship to reading should be modeled

separately (that is, as three factors, or are they similar enough to be modeled as one overall writing factor). Confirmatory factor analysis techniques using structural equation modeling were used to address this question.

As described in Chapter 2, many researchers have found that students writing essays in the narrative, informational, and persuasive modes produce essays that do not have comparable scores. In some cases, the resulting scores are different enough that any conclusions about a student's overall writing ability would be questioned if based on writing in only one mode. To examine the effect of mode more closely, CFA in SEM was used to model two ways that writing mode might be related to overall writing scores (the latent writing ability). All of the SEM analyses were conducted with Amos 7.0 (Arbuckle, 2006).

The 1-factor model is shown in Figure 3.3 and the 3-factor model is shown in Figure 3.4. The 1-factor model represents the hypothesis that the essay scores (N1, N2, I1, I2, P1, P2) are all indicators of one latent construct: writing.

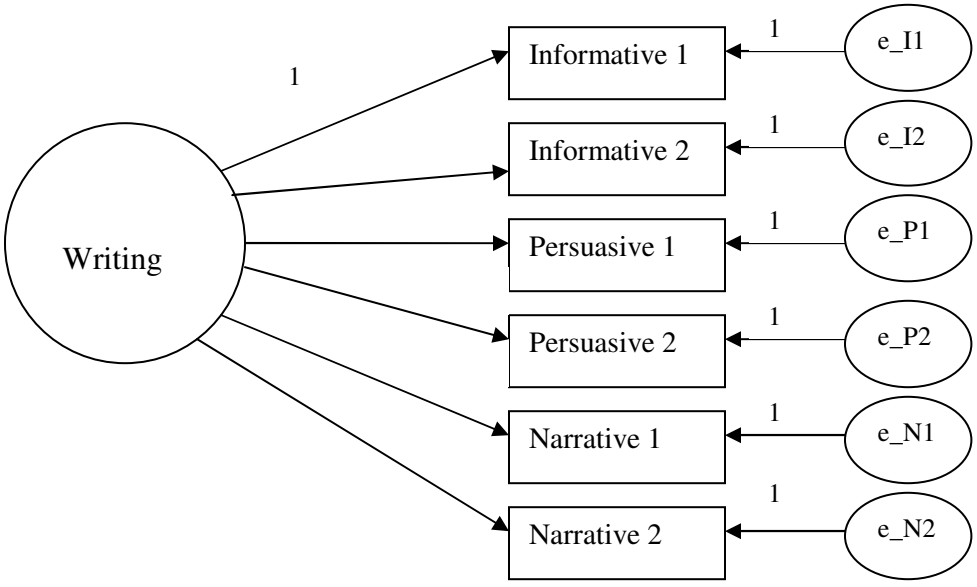

Figure 3.3. One-factor model for writing.

The 3-factor model for writing tests the hypothesis that writing in different discourse modes is sufficiently different that each mode of writing can be considered a separate latent construct. The 3-factor model represents the hypothesis that the essay scores for each separate mode (N1, N2, I1, I2, P1, P2) are indicators of a mode-specific latent construct: narrative writing, informative writing, or persuasive writing. The one-sided arrows pointing from each mode of writing to the indicators represent the presumed direct causal effect of the latent variables (mode-specific writing) on the observed measure (mode-specific essay score). The double-sided arrows from each latent construct represent the relationship between them.

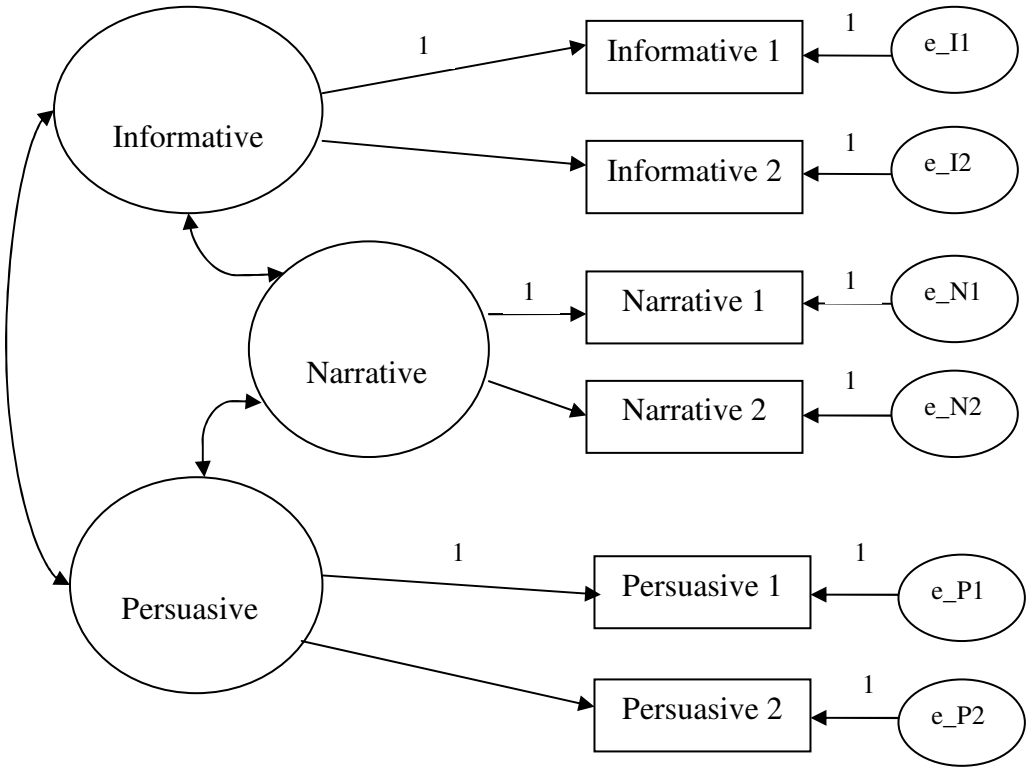

Figure 3.4. Three-factor model for writing.

The 1-factor model includes six observed indicator variables for a total of 21 sample moments [(6*7)/2] and 12 distinct parameters to be estimated. The 3-factor model includes the same six observed indicator variables for a total of 21 sample moments. The covariances between the factors mean that three additional parameters must be estimated, which increases the number of parameters to be estimated and reduces the degrees of freedom. Table 3.7 shows the summary of the calculation for degrees of freedom for each model.

Table 3.7

Calculation of Degrees of Freedom for the 1 Factor and 3 Factor Models

	1-Factor Model	3-Factor Model
Number of distinct sample moments	21	21
Number of distinct parameters to be estimated	12	15
Degrees of freedom	9	6

The fit of these two models to the data was analyzed based on recommendations provided by Hu and Bentler (1999) and described in detail later in this section. One model was chosen for the subsequent analyses on the writing and reading data. The relative parsimony of the two models was an important consideration in choosing the most appropriate model. If the initial CFA analysis found that the 1-factor model was not significantly different from the 3-factor model, then the 1-factor model would be selected for further analyses because it contains fewer parameters to estimate. Kline (1998) has suggested having at least a 10:1 ratio between sample size and parameters. With 12 parameters in the 1-factor model, 15 parameters in the 3-factor model, and a sample size of between 85 (grade 10) and 122 (grade 6), the present study has ratios that range from below the recommended ratio to adequate. At the lowest bound, the ratio for grade 10 using the 3-factor model is approximately 6:1. At the

upper bound, the ratio for grade 6 using the 1-factor model is approximately 10:1. Given the relative number of parameters for each model and the low sample size, a more parsimonious model is preferred.

Hu and Bentler (1999) provided guidance regarding appropriate values for evaluating the fit of structural equation models. Their recommendations are used as the basis for this analysis. The statistics used to evaluate the fit of the two models are shown in Table 3.8.

Table 3.8

Fit Statistics Used to Evaluate Models

Statistic	Adequate Fit Indicated by
Chi-squared (χ^2)	$p \geq .05$
(χ^2) / df	≤ 3
Standardized Root Mean Residual(SRMR)	$\leq .08$
Goodness of Fit Index (GFI)	$\geq .95$
Root Mean Squared Approximation (RMSEA)	$\leq .06$
Change in χ^2 and significance of the change [$\chi^2_{diff}(p)$]	$p \geq .05$

Once the most appropriate way to model writing was determined, the remaining research questions were addressed using the chosen model.

The aspect of research question 1a regarding the similarity in relationship between reading comprehension and writing performance across grade levels was also addressed with CFA. The essential question was whether group membership (grade) moderates the relationship between reading and writing. The 1-factor model for writing identified in the previous analysis was expanded to add reading as a second latent variable with the two reading tests as indicators of reading ability. Because the sample sizes in this study were small, the 1-factor model was adjusted to reduce the number of indicator variables and thus the number of parameters that needed to be estimated. The two indicators for each discourse

mode were combined to make one indicator variable that represented one discourse mode each. Thus the six writing indicators were reduced to three, one for each mode of writing, which eliminated three parameters. This model is shown in Figure 3.4.

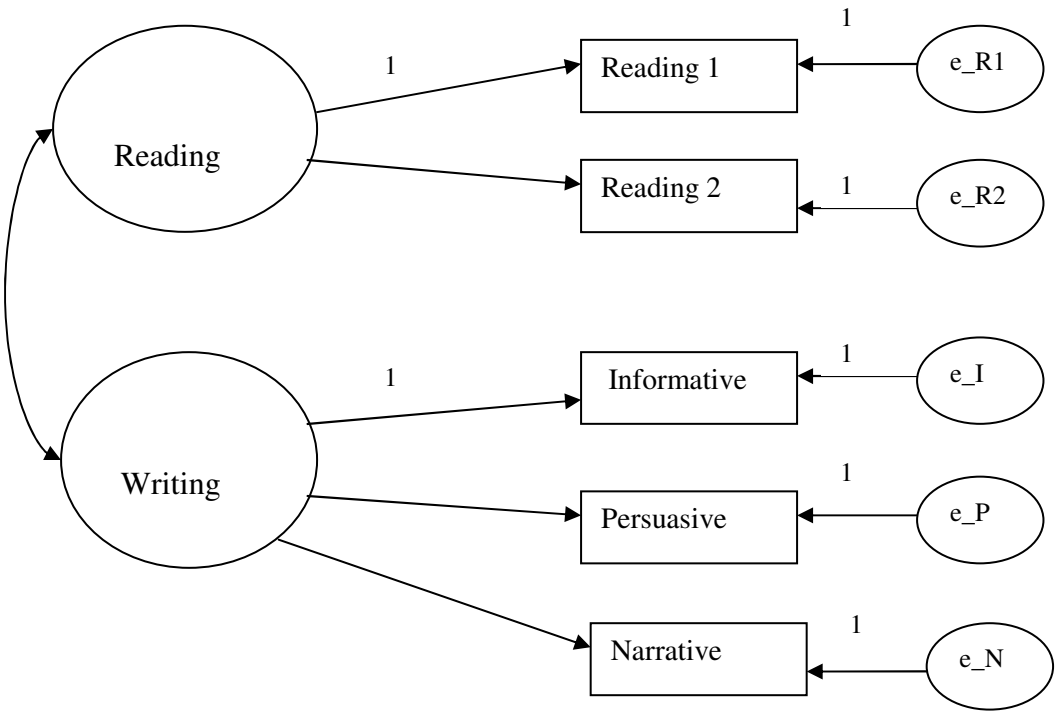

Figure 3.5. CFA model for the reading-writing relationship.

To analyze further the reading-writing relationship across grades and explore developmental trends, multiple-group analyses were performed using each grade as a group. This analysis addressed research question 2, which concerned the relationship between writing ability and the mode of the writing prompts (narrative, informational, persuasive). The same analysis procedure also shed light on research question 3, which concerns whether performances across writing modes are similar across gender and race groups.

Comparisons across grades were conducted in a sequential, pairwise fashion to determine whether the relationship between reading comprehension and writing quality was stable across adjacent grades. The fit of the models was evaluated under two conditions: (1) when the covariances were allowed to vary freely for each grade separately, and (2) when the covariance between reading and writing were constrained to be equal at adjacent grade levels.

In subsequent analyses, the factor weights for the indicator variables for writing were examined. To determine whether the factor loadings for the indicator variables were stable across adjacent grades, comparisons across grades were conducted in a sequential, pairwise fashion, as in the previous analysis. The fit of the models was evaluated under two conditions: (1) when the factor loadings were allowed to vary freely for each grade separately, and (2) when the factor loadings between writing and its indicators were constrained to be equal at adjacent grade levels. The examination of relative differences across modes and grades also served to provide insight into the developmental framework for reading and writing proposed by Fitzgerald and Shanahan (2000).

A final analysis was conducted to compare the factor loadings between writing and its indicator variables across gender and race groups.

Summary and Limitations

The series of analyses conducted in this research focus on the relationship between reading comprehension and writing quality both within grades and across grades. Although a great deal of attention has been paid to the development of skills precursory to reading comprehension and essay writing, less attention has been given to the specific abilities of

reading comprehension and writing quality, particularly how the relationship between them may change across grades. To develop a better understanding of the relationship between reading comprehension and writing quality, several key questions have been addressed. The first question addresses the relationship between reading comprehension and writing quality scores. Do the scores on a measure of reading comprehension account for a large amount of the variance in scores on a writing performance measure? The second and third questions address the relationship between discourse mode and writing quality scores and how this may affect the strength of the relationship between reading ability and writing quality. A general analysis using correlations between scores was followed by confirmatory factor analysis to examine relationships between writing scores on the three discourse modes. Next, multiple-group analyses were conducted to assess the stability of the reading-writing relationship across adjacent grades and identify any developmental trends. Differences between gender and race groups were also explored.

Some limitations to the analysis plan should be noted. First, the modest sample size used in this study is a limitation in terms of statistical power and in terms of the generalizabity of the findings presented in the next chapter. Ideally, much larger groups of students (200 + per grade level) would be used to conduct the SEM analyses. Second, because the tests used in this study carried no consequences for the students, the results may not reflect the students' best efforts. The somewhat low correlations between writing scores within grade, especially at higher grade levels, suggest that there may be error in the scores due to student motivational factors. Finally, the method of scoring the essays of grade 6 and 10 may have affected the scores for those grades. Half of the essays for each group were scored with the essays of the lower grade set and half with the essays of the higher grade set. For example,

the essays written to prompts 4BN, 4BI, and 4BP for grade 6 were scored with the essays from the full grade 4 sample and the essays written to prompts 8AN, 8AI, and 8AP for grade 6 were scored with the full grade 8 sample. The scores for the group A and group B sets of essays for grades 6 and 10 were more divergent than in other grades, suggesting that the scoring method had an effect on the results. These limitations notwithstanding, the study design and procedures offered a strong opportunity to investigate the relationship between reading and writing across grade levels and to probe more deeply into the relationship between writing ability and discourse mode of the writing prompt. The results of these analyses are presented in the next chapter.

CHAPTER 4

RESULTS

Using data from approximately 500 students across grades 4, 6, 8, 10, and 12, the strength of the relationship between reading ability and writing quality across grade levels was examined with a particular focus on the impact of writing prompt mode on writing quality scores. Correlational analyses as well as structural equation modeling were used to assess the strength of the relationship between reading comprehension and writing quality within and cross grades. Structural equation modeling was used to evaluate the factor structure of reading ability and writing quality using reading comprehension test scores and essay scores as the indicator variables. The factor structure of writing was evaluated as well as the factor structure of a reading-writing model. The reading-writing model was also evaluated for factor structure invariance across grades, gender, and race.

The Reading-Writing Connection

Research question 1 concerned the strength of the relationship between reading ability and writing quality at each grade level as well as the relationship between these variables across grades. To address this, the Pearson product moment correlation between reading scores and writing quality scores was calculated for each grade. Next, because the interest of this research is in the reading ability and writing quality constructs generally, and not only the test scores produced for this study, the disattenuated correlations between reading and writing were calculated using the formula provided in Chapter 3. The disattenuation formula

requires the reliability estimates for both the reading measure and the writing measure. These estimates and the results of the correlational analysis are provided in Table 4.1, which also shows the r^2 values for the disattenuated correlations. The r^2 value represents the degree to which variability in one measure is attributable to variability in the other measure. Thus, in grade 4, the data show that 47% of the variance in reading ability can be explained by the variability of the writing quality scores and at grade 10, 74% of the variability can be explained in this manner.

Table 4.1

Reliability Estimates for Reading Ability and Writing Quality and the Correlations Between Them

Grade	Reading Reliability (r_{xx})	Writing Reliability (*G Coefficient*)	Reading Writing Correlation (r_{xy})	Disattenuated Correlation (r_{xy})	Squared Correlation (r^2)
4	0.85	0.82	0.57	0.68	0.47
6	0.81	0.83	0.48	0.59	0.34
8	0.76	0.83	0.67	0.84	0.71
10	0.64	0.84	0.63	0.86	0.74
12	0.62	0.84	0.61	0.85	0.71

In general, these correlations indicate that there is a strong relationship between reading ability and writing quality in all grades, but that it varies by grade. The relationship is stronger in the upper grades than it is in the lower grades. In grades 4 and 6, the average disattenuated correlation between reading ability and writing quality is .635, and the average is .85 for grades 8, 10, and 12. The large size of the correlations at all grades is noteworthy. With the exception of grade 6, the unadjusted correlations are higher than the .50 reported by Shanahan (2006) to be the highest typically found in performance-based studies of the reading-writing connection.

The Discourse Mode

The focus of research question 2 was on the discourse mode of the writing prompt and the relationship of the discourse mode to writing quality scores in general and the reading-writing relationship across grades. A bivariate correlation analysis was conducted to examine the relative difficulty of the discourse mode and an SEM analysis using a simple model of the reading-writing relationship was conducted to address the impact of discourse modes on the reading-writing relationship.

A preliminary view of the relationship between the discourse mode and the writing quality scores was obtained by calculating bivariate correlations between the writing quality scores. These are shown in Table 4.2 along with the means and standard deviations for each set of scores. All correlations are significant at the $p < 0.01$ level with only one exception. The correlation between informative 1 and narrative 2 in grade 10 is significant at the $p < 0.05$ level.

In general, these correlations reveal a pattern of strong relationships between the writing modes but do not support a clear pattern that distinguishes within from between mode relationships. In many cases, the correlations within the same discourse mode (e.g., Informative 1 and Informative 2) are lower than the correlations between different modes (e.g., Informative 1 and Narrative 1). Only in grade 10 is the highest correlation between prompts of the same mode. In grade 10, the highest correlation occurs between Informative 1 and Informative 2. In grades 4, 6, and 8, the highest correlation occurs between informative and narrative, whereas in grade 12, it occurs between informative and persuasive.

The means for each prompt show that in general, the easiest prompt was narrative and the most difficult was persuasive. However, the pattern of difficulties within the two extremes is

not consistent. The second easiest prompt was either informative or persuasive, not narrative. The second most difficult prompt was persuasive in grade 8 only. This suggests that although a trend seems to be present in the data, the difficulty of the prompt cannot be predicted by the discourse mode alone.

Table 4.2

Intercorrelations, Means, and Standard Deviations for Writing Quality Scores in Different Modes

Subscale	INF 1	NAR 1	PER 1	INF 2	NAR 2	PER 2
Grade 4 (n = 115)						
1. Informative 1	—	.53	.48	.58	.68	.44
2. Narrative 1		—	.58	.66	.62	.51
3. Persuasive 1			—	.57	.42	.42
4. Informative 2				—	.59	.59
5. Narrative 2					—	.56
6. Persuasive 2						—
Mean (SD)	0.50 (1.73)	0.80 (2.16)	0.68 (2.58)	0.37 (2.13)	0.17 (1.92)	-0.55 (2.36)
Grade 6 (n = 122)						
1. Informative 1	—	.63	.57	.47	.57	.51
2. Narrative 1		—	.55	.55	.62	.54
3. Persuasive 1			—	.49	.54	.42
4. Informative 2				—	.49	.56
5. Narrative 2					—	.56
6. Persuasive 2						—
Mean (SD)	1.23 (1.97)	1.44 (1.75)	1.27 (1.96)	-0.13 (2.01)	0.30 (1.51)	-1.98 (1.66)
Grade 8 (n = 106)						
1. Informative 1	—	.64	.71	.66	.61	.66
2. Narrative 1		—	.63	.67	.74	.67
3. Persuasive 1			—	.75	.65	.70
4. Informative 2				—	.79	.76
5. Narrative 2					—	.69
6. Persuasive 2						—
Mean (SD)	0.63 (2.15)	0.99 (1.98)	-0.55 (2.29)	0.04 (2.41)	0.28 (1.99)	-0.04 (2.50)
Grade 10 (n = 85)						
1. Informative 1	—	.47	.48	.59	.26	.35
2. Narrative 1		—	.48	.51	.48	.50
3. Persuasive 1			—	.48	.44	.41
4. Informative 2				—	.33	.56
5. Narrative 2					—	.49
6. Persuasive 2						—
Mean (SD)	1.97 (1.98)	1.45 (1.71)	1.50 (1.61)	1.28 (2.10)	1.45 (1.71)	-0.04 (2.24)

Table 4.2, Cont'd.

Subscale	INF 1	NAR 1	PER 1	INF 2	NAR 2	PER 2
			Grade 12 (n = 93)			
1. Informative 1	—	.63	.70	.54	.33	.50
2. Narrative 1		—	.65	.57	.50	.51
3. Persuasive 1			—	.66	.42	.63
4. Informative 2				—	.34	.59
5. Narrative 2					—	.34
6. Persuasive 2						—
Mean (SD)	0.57 (2.36)	0.94 (2.47)	-0.06 (2.69)	-0.03 (2.19)	1.55 (2.22)	1.45 (2.61)

Simple Reading-Writing Model Using SEM

To examine the degree to which the strength of the relationship between reading and writing was affected by the mode of the indicator variables for writing, the simple SEM model shown in Figure 4.1 was proposed. Within each grade, writing was modeled using essays for each indicator mode separately. Figure 4.1 shows the informative essays as the indicator variables. Narrative and persuasive essays also served as the sole indicator variables for writing in the analysis for each grade.

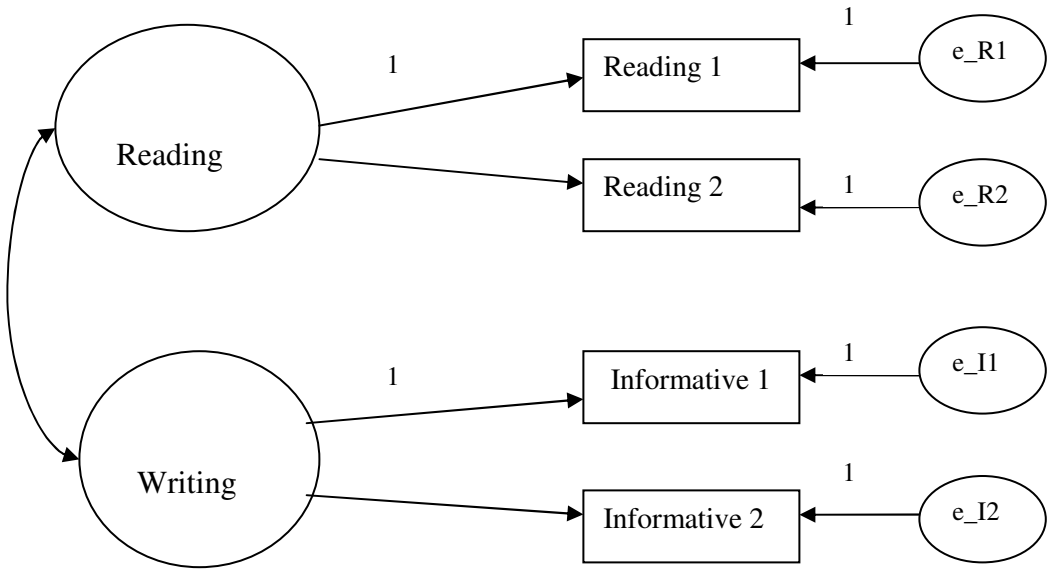

Figure 4.1. Simple reading-writing model.

The results shown in Table 4.3 indicate only a very slight trend of lower correlations when persuasive essays were used as the indicator variables for writing. The model that used narrative essays as the indicator variable for writing resulted in the highest correlation between the reading and writing latent variables in grades 4, 10, and 12 and the lowest correlation in grade 8. When informative essays were used as the indicator variables for writing, grades 6 and 8 showed the highest correlation between reading and writing and grade 12 the lowest. Persuasive essay scores as indicator variables did not result in the highest correlation of the modes in any of the grades but it produced the lowest correlation only in 3 of the 5 grades (4, 6, and 10).

Table 4.3

Reading-Writing Correlation by Grade and Discourse Mode

Discourse Mode	Grade 4	Grade 6	Grade 8	Grade 10	Grade 12
Informative	.60	.64	.81	.72	.61
Narrative	.75	.53	.61	.76	.74
Persuasive	.58	.51	.76	.66	.71

Overall, the results of this analysis do not strongly support the hypothesis that writing in a specific discourse mode produces consistently higher correlations between the constructs of reading comprehension and writing quality.

The Writing Model Using Indicators in all Discourse Modes

The next analysis focused on the writing quality scores and established the best SEM model to use for comparing the factor structures of reading and writing when all six writing scores were used in the same model. The first step was to determine the appropriate way to

model writing. The 1-factor and 3-factor models described in Chapter 3 were compared to determine which best fit the writing data. The fit of the one factor model was first evaluated to determine whether it appropriately modeled the data. The fit statistics for the one factor model are shown in Table 4.4.

Table 4.4

Fit Indices for the One Factor Writing Model by Grade

Grade	χ^2	df	χ^2/df	GFI	RMSEA	SRMR
4 (n = 115)	22.01**	9	2.45	.94	.11	.04
6 (n = 122)	10.28	9	1.14	.97	.03	.03
8 (n = 106)	21.10*	9	2.34	.94	.11	.03
10 (n = 85)	17.92*	9	2.00	.93	.11	.06
12 (n = 93)	11.23	9	1.25	.96	.05	.04

Note. GFI = Goodness of Fit Index; RMSEA = Root Mean Square Error Approximation; SRMR = Standardized Root Mean Residual.
*p < .05. ** p < .01.

The results for the one factor model are somewhat mixed. Suggesting a poor fit are the χ^2 values and the RMSEA for grades 4, 8 and 10. The significance levels of the χ^2 values for these three grades are less than the optimal of $p \geq .05$, and the RMSEA values are higher than .06. However, the GFI values for all grades are either above .95 or just slightly below, suggesting an adequate fit. In addition, the χ^2/df values are all below 3, and the SRMR values are all well below .08, which also suggests an adequate model fit.

To determine whether the 3-factor model would fit the data better than the 1-factor model, the 3-factor model was run for all of the grades. The results for the 3-factor model established that the 1-factor model is a better way to model the writing data. For all grades, running the 3-factor model produced the following message in AMOS: “This solution is not admissible.” Examination of the results showed that the error-adjusted correlations between

narrative, informative, and persuasive writing were greater than 1.00 in all cases in grades 4, 6, 8, and 12 and for the narrative-persuasive correlation in grade 10, which produced the nonpermissible solution. This type of error falls into a class of errors often called *Heywood cases*. Although Heywood cases have several identified causes (Chen, Bollen, Paxton, Curran, & Kirby, 2001), three causes are the most likely in this case: low sample size (Boomsma, 1983, 1985), combined with only two indicators per factor (Anderson & Gerbing, 1984) and model misspecification (Van Driel, 1978). The model misspecification for the 3-factor model is likely caused by there being too many factors in the model. The narrative, persuasive, and informative modes are so highly related that it may not be meaningful, or mathematically possible, to model them as distinct factors for this type of analysis. This result addresses research question 2b concerning the best way to model writing as a latent variable and determined that the 1-factor model would be used to model writing for subsequent analyses. Figure 4.2 shows the standardized (correlations) and unstandardized results for the grade 8, 1-factor model. The unstandardized estimates are shown in parentheses. Figures for all grades are provided in Appendix C.

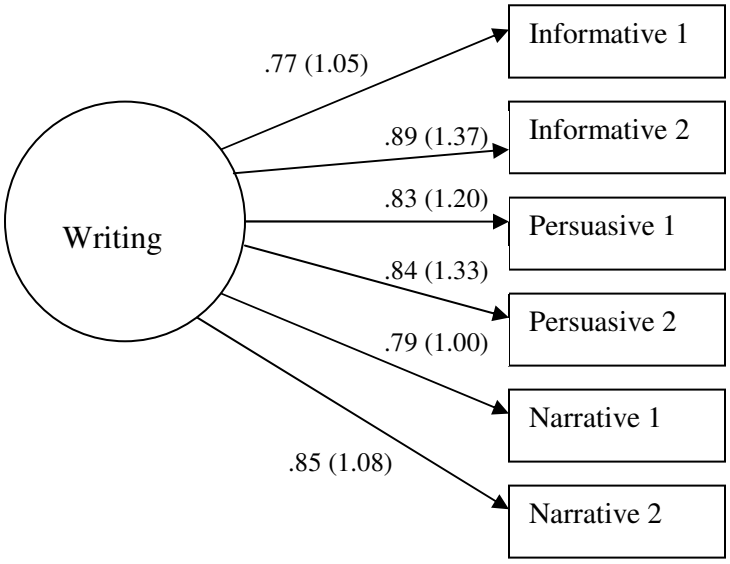

Figure 4.2. Standardized and unstandardized estimates for grade 8 one factor model.

The Reading-Writing Model

Once it was established that the 1-factor model was an appropriate way to model writing as a latent variable, the full model with writing as one latent variable and reading as the second latent variable could be examined. Figure 4.3 shows the confirmatory factor analysis model used to examine writing more closely and examine the reading-writing relationship for the remaining research questions. Because the research involves multiple grades, the results of the analyses for all grades are presented together in table form rather than separately in figure form. The correlations, means, and standard deviations for the models shown in Figure 4.3 are provided in Appendix D. Separate figures for each grade showing the standardized and unstandardized results are in Appendix E.

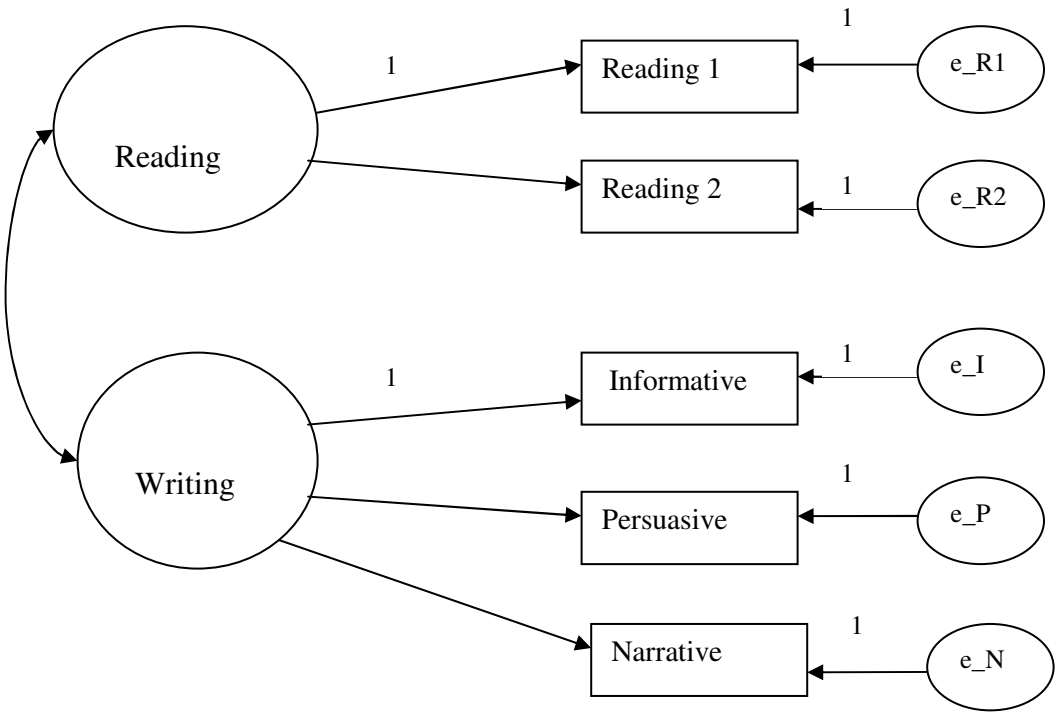

Figure 4.3. CFA model for the reading-writing relationship.

The first step in the analysis was to establish that the model shown in Figure 4.3 is an appropriate way to model the relationship between the reading comprehension and writing quality. Table 4.5 shows the fit statistics for this model for all grades. Overall, the indicators suggest that the data fit the model well: The χ^2 values are not significant at the .05 level, the χ^2/df values are below 3, the GFI indices are above .95, and the SRMR indices are below .08. The RMSEA indices are below .06 for grades 6 and 8. In grades 4, 10, and 12, the RMSEA indices were higher than .06. However, because all other indicators suggest that this model is an appropriate one to use, this model was retained for subsequent analyses. The hypothesis that the same factor analysis model holds for all grades is supported.

Table 4.5

Fit Indices for the Reading-Writing Model by Grade

Grade	χ^2	df	χ^2/df	GFI	RMSEA	SRMR
4 (n = 115)	7.21	4	1.80	.97	.08	.03
6 (n = 122)	3.99	4	1.00	.99	.00	.02
8 (n = 106)	4.46	4	1.12	.98	.03	.02
10 (n = 85)	6.21	4	1.55	.97	.08	.03
12 (n = 93)	7.17	4	1.79	.97	.09	.04

Note. GFI = Goodness of Fit Index; RMSEA = Root Mean Square Error Approximation; SRMR = Standardized Root Mean Residual.

The factor loadings, standardized estimates, and squared multiple correlations for the indicator variables are shown in Table 4.6 along with the standard error and critical ratio for the factor loadings. The critical ratio is obtained by dividing the factor loading by its standard error. Critical ratios above 1.96 indicate significance at the $p < .05$ level (Arbuckle, 1999). The estimates have been rounded to two decimal places for the purposes of presentation in the tables.

The factor loadings represent the direct effects of the factors on the indicators. Thus, if the amount of writing (or reading) ability were to increase one unit, the level of the corresponding indicators would be expected to increase by a unit amount proportional to the regression weight. For example, if grade 4 writing ability increased by 1.00 units, then the scores on the narrative prompts would be predicted to increase by 1.13 units. The factor loadings that were fixed to 1.0 to scale the latent variable remain 1 for the estimate. The standard errors of the estimate are not calculated and the estimates for the fixed loadings are not tested for significance (Kline, 1998).

As in the case of the research described here, when each indicator is specified to load on only a single factor, then the standardized estimate of the factor loading can be interpreted as

its correlation and the square of its factor loading equals the proportion of variance explained by the factor (Kline, 1998). For example, in grade 4, the standardized estimate for Writing → Narrative is .91, so the writing factor accounts for $(.91)^2$, or 82% of the variance in narrative-mode writing quality scores.

To investigate whether the relationship between reading comprehension and writing was stable across adjacent grades, the fit of the model was evaluated under two conditions for adjacent grades (i.e., grades 4 and 6, grades 6 and 8, grades 8 and 10, and grades 10 and 12). The first condition, model A, allowed reading and writing to covary freely for each grade; no constraints were imposed. The second condition, model B, restricted the covariance between reading and writing to be equal for both grades. Table 4.7 shows the results of this analysis. The same fit statistics used to evaluate the models described previously are presented, as well as the χ^2_{diff} values and the AIC values.

Table 4.6

Factor Loading Information for Reading-Writing Model

Model	Factor Loadings	Standard Error	Critical Ratio	Standardized Estimates (Correlations)	Squared Correlations
			Grade 4		
Reading →R1	1.00			.96	.92
Reading →R2	.80	.08	10.28	.88	.77
Writing → I	1.00			.85	.73
Writing → P	1.09	.12	9.52	.77	.59
Writing → N	1.13	.10	11.43	.91	.82
			Grade 6		
Reading → R1	1.00			.82	.68
Reading → R2	1.11	.14	8.1	1.00	1.00
Writing → I	1.00			.87	.76
Writing → P	.86	.08	11.12	.84	.71
Writing → N	.83	.08	11.06	.84	.70
			Grade 8		
Reading → R1	1.00			.87	.75
Reading → R2	1.25	.11	11.03	.93	.86
Writing → I	1.00			.94	.89
Writing → P	1.03	.06	15.99	.91	.84
Writing → N	.79	.06	12.98	.84	.71
			Grade 10		
Reading → R1	1.00			.85	.72
Reading → R2	1.11	.16	6.90	.82	.67
Writing → I	1.00			.74	.55
Writing → P	.99	.15	6.75	.82	.67
Writing → N	.84	.13	6.48	.77	.60
			Grade 12		
Reading → R1	1.00			.93	.86
Reading → R2	.75	.13	5.76	.70	.50
Writing → I	1.00			.86	.74
Writing → P	1.24	.13	9.6	.88	.78
Writing → N	.88	.11	7.85	.73	.54

Table 4.7

Fit Indices for Model A and Model B

Model	χ^2	Df	$\chi^2_{diff}(p)$	AIC	GFI	RMSEA	SRMR
Grades 4 and 6							
A	11.21	8		55.20	.98	.04	.03
B	12.24	9		54.24	.98	.04	.04
			1.03				
Grades 6 and 8							
A	8.45	8		52.45	.99	.02	.02
B	13.571	9		55.57	.98	.05	.04
			5.12*				
Grades 8 and 10							
A	10.68	8		54.68	.98	.04	.02
B	12.50	9		54.50	.98	.05	.03
			1.82				
Grades 10 and 12							
A	13.38	8		57.38	.97	.07	.04
B	13.67	9		55.67	.97	.05	.04
			0.29				

Note. AIC = Akaike Information Criterion; GFI = Goodness of Fit Index; RMSEA = Root Mean Square Error Approximation; SRMR = Standardized Root Mean Residual.
*$p < .05$

For all grade comparisons except the one between grades 6 and 8, the model fit did not diminish significantly when the covariance between the reading and writing latent factors were constrained to be equal. The χ^2_{diff} value is used to determine whether the change in χ^2 between two models is significant, so the lack of significance for these comparisons (grades 4 and 6, grades 8 and 10, and grades 10 and 12) suggest that models A and B fit the data equally well for those set of grade pairs. This suggests that only between grades 6 and 8 is the relationship between the reading and writing factors significantly different. In addition, for these grade comparisons (4 and 6, 8 and 10, and 10 and 12), the AIC values are smaller for model B than for model A. A smaller AIC value indicates the preferred model. However, the GFI, RMSEA, and SRMR indices for the grades 6 and 8 model B are within the range for

a well-fitting model, which when evaluated in conjunction with the significant χ^2 difference, suggests that although there are differences between the two grades, they are not substantial.

The next series of analyses addressed the components of research question 2 that concern the relationship between writing as a latent construct and the mode of the writing prompt. The hypothesis is that the factor loadings between the latent and indicator variables are not significantly different across grades. To test this hypothesis, the data for adjacent grades were compared in a set of analyses that examined the difference in the factor loadings between the latent variables and the indicator variables.

For each set of analyses, two models were fit to the data: models A and C. Model A did not include any constraints on the factor loadings between the latent variables and the indicator variables. Thus, the factor loadings for model A were the same as those obtained when each grade was run separately (see Table 4.6). These are shown in Tables 4.9 through 4.12 as the first number in each cell. For each set of grades, model C assumes a group-invariant factor pattern for writing. This means that the relative influence of the latent construct on each of the indicator variables should be the same for both grades. In other words, the influence of writing ability (e.g., the latent construct) on narrative writing quality scores would be the same in the grades being analyzed. To test this hypothesis, the regression weights between the latent variable and indicator variables are constrained to be equal. If the fit statistics show that model C fits the data as well as model A, or at least not significantly worse than model A, then the hypothesis would be supported.

Table 4.8 shows the fit statistics for models A and C. The same fit statistics used to evaluate models A and B are presented for this analysis as well.

Table 4.8

Fit Indices for Model A and Model C

Model	χ^2	Df	χ^2_{diff}	AIC	GFI	RMSEA	SRMR
			Grades 4 and 6				
A	11.21	8		55.20	.98	.04	.03
C	23.91*	11		61.91	.96	.07	.04
			11.28*				
			Grades 6 and 8				
A	8.45	8		52.45	.99	.02	.02
C	9.99	11		47.97	.98	.00	.03
			4.32				
			Grades 8 and 10				
A	10.68	8		54.68	.98	.04	.02
C	9.78	11		47.78	.98	.00	.02
			.84				
			Grades 10 and 12				
A	13.38	8		57.38	.97	.06	.04
C	18.57	11		56.57	.96	.06	.05
			5.19				

Note. AIC = Akaike Information Criterion; GFI = Goodness of Fit Index; RMSEA = Root Mean Square Error Approximation; SRMR = Standardized Root Mean Residual.
*$p \leq .05$

The hypothesis that the factor structure is invariant across grades is strongly supported by the comparisons between grades 6 and 8, grades 8 and 10, and grades 10 and 12. In these comparisons, model C fits the data quite well, and the χ^2_{diff} values are not significant. The lack of significance for these comparisons suggest that models A and C fit the data equally well. In addition, for these grade comparisons, the AIC values are smaller for model C than for model A. A smaller AIC value indicates the preferred model. For the comparison between grades 4 and 6, the fit statistics for the comparisons between models A and C do not suggest that model C is preferred over model A. However, they do show a fairly good fit for model C, providing additional, modest support for the factor invariance hypothesis across those grade comparisons.

The factor-loading information for models A and C for each grade comparison are shown in Tables 4.9 through 4.12. The factor loadings can provide insight into the relationship between the writing factor and the writing quality scores for the three discourse modes. In grade 4, the largest loading for writing is between writing and narrative (1.13). In grade 6, the largest loading for writing is between writing and informational (1.00). In grades 8 and 10, the factor loading for informational and persuasive are nearly identical (G6 I = 1.00, P = 1.03; G8 I = 1.00, P = .99) and in grade 12, the largest loading for writing is between writing and persuasive (1.24). Although in most cases, the differences are slight, the patterns do hint at some developmental trend in the degree to which writing ability is manifested in students' ability to produce essays in the different modes of writing.

Table 4.9

Factor Loading Information for Grade 4 and Grade 6 Model Comparisons

	Factor Loadings	Standard Error	Critical Ratio	Standardized Estimates (Correlations)	Squared Correlations
Model A (Model C) Grade 4					
Reading → R1	1.00			.96 (.91)	.92 (.83)
Reading → R2	.80 (.92)	.08 (.07)	10.28 (13.34)	.88 (.92)	.77 (.85)
Writing → I	1.00			.85 (.89)	.73 (.79)
Writing → P	1.09 (.98)	.12 (.07)	9.52 (14.39)	.77 (.76)	.59 (.58)
Writing → N	1.13 (.97)	.10 (.06)	11.43 (15.63)	.91 (.87)	.82 (.76)
Model A (Model C) Grade 6					
Reading → R1	1.00			.82 (.89)	.68 (.79)
Reading → R2	1.11 (.92)	.14 (.07)	8.1 (13.34)	1.00 (.93)	1.00 (.87)
Writing → I	1.00			.87 (.83)	.76 (.69)
Writing → P	.86 (.98)	.08 (.07)	11.12 (14.39)	.84 (.85)	.71 (.73)
Writing → N	.83 (.97)	.08 (.06)	11.06 (15.63)	.84 (.86)	.71 (.74)

Table 4.10

Factor Loading Information for Grade 6 and Grade 8 Model Comparisons

	Factor Loadings	Standard Error	Critical Ratio	Standardized Estimates (Correlations)	Squared Correlations
Model A (Model C) Grade 6					
Reading → R1	1.00			.82 (.79)	.69 (.70)
Reading → R2	1.11 (1.21)	.14 (.09)	8.1 (13.27)	1.00 (1.00)	1.00 (1.00)
Writing → I	1.00			.87 (.86)	.76 (.74)
Writing → P	.86 (.96)	.08 (.05)	11.12 (19.15)	.84 (.87)	.71 (.76)
Writing → N	.83 (.81)	.08 (.05)	11.06 (17.21)	.84 (.82)	.70 (.67)
Model A (Model C) Grade 8					
Reading → R1	1.00			.87 (.88)	.75 (.77)
Reading → R2	1.25 (1.21)	.11 (.09)	11.03 (13.27)	.93 (.92)	.86 (.85)
Writing → I	1.00			.94 (.95)	.89 (.90)
Writing → P	1.03 (.96)	.06 (.05)	15.99 (19.15)	.91 (.90)	.84 (.81)
Writing → N	.79 (.81)	.06 (.05)	12.98 (17.21)	.84 (.85)	.71 (.72)

Table 4.11

Factor Loading Information for Grade 8 and Grade 10 Model Comparisons

	Factor Loadings	Standard Error	Critical Ratio	Standardized Estimates (Correlations)	Squared Correlations
Model A (Model C) Grade 8					
Reading → R1	1.00			.87 (.87)	.75 (.76)
Reading → R2	1.25 (1.21)	.11 (.09)	11.03 (13.05)	.93 (.92)	.86 (.85)
Writing → I	1.00			.94 (.94)	.89 (.89)
Writing → P	1.03 (1.02)	.06 (.06)	15.99 (17.35)	.91 (.91)	.84 (.83)
Writing → N	.79 (.80)	.06 (.06)	12.98 (14.70)	.84 (.84)	.71 (.71)
Model A (Model C) Grade 10					
Reading → R1	1.00			.85 (.82)	.72 (.68)
Reading → R2	1.11 (1.21)	.16 (.09)	6.90 (13.05)	.82 (.84)	.67 (.71)
Writing → I	1.00			.74 (.74)	.55 (.55)
Writing → P	.99 (1.02)	.15 (.06)	6.75 (17.35)	.82 (.83)	.67 (.69)
Writing → N	.84 (.80)	.13 (.06)	6.48 (14.70)	.77 (.75)	.60 (.57)

Table 4.12

Factor Loading Information for Grade 10 and Grade 12 Model Comparisons

	Factor Loadings	Standard Error	Critical Ratio	Standardized Estimates (Correlations)	Squared Correlations
Model A (Model C) Grade 10					
Reading → R1	1.00			.85 (.89)	.72 (.80)
Reading → R2	1.11 (.92)	.16 (.10)	6.90 (8.87)	.82 (.76)	.67 (.57)
Writing → I	1.00			.74 (.70)	.55 (.49)
Writing → P	.99 (1.15)	.15 (.10)	6.75 (11.59)	.82 (.85)	.67 (.72)
Writing → N	.84 (.89)	.13 (.09)	6.48 (10.20)	.77 (.76)	.60 (.58)
Model A (Model C) Grade 12					
Reading → R1	1.00			.93 (.87)	.86 (.76)
Reading → R2	.75 (.92)	.13 (.10)	5.76 (8.87)	.70 (.76)	.50 (.58)
Writing → I	1.00			.86 (.87)	.74 (.75)
Writing → P	1.24 (1.15)	.13 (.10)	9.6 (11.59)	.88 (.87)	.78 (.75)
Writing → N	.88 (.89)	.11 (.09)	7.85 (10.20)	.73 (.75)	.54 (.56)

Gender and Race Comparisons

The third research question addressed the relationship between reading and writing for groups of different gender and race. The sample sizes within grade were too small to conduct the desired SEM analysis between gender and race groups within grade. So, based on previous analyses, which showed that the factor structure being tested was relatively invariant across grades 6, 8, 10, and 12, the samples for those grades were combined for the cross race and gender analyses. The same type of analyses that were conducted for each grade were also conducted for males, females, Blacks, and Whites. Table 4.13 shows the fit indices for the model for each gender or race singly with no constraints imposed on the factor loadings.

Table 4.13

Fit Indices for the Reading-Writing Model by Gender and Race

Model	χ^2	df	χ^2/df	GFI	RMSEA	SRMR
Males (n = 186)	4.10	4	1.03	.99	.01	.01
Females (n = 220)	6.65	4	1.66	.99	.06	.01
Blacks (n = 165)	15.19*	4	3.79	.97	.13	.03
Whites (n = 224)	1.76	4	.44	1.00	.00	.01

Note. GFI = Goodness of Fit Index; RMSEA = Root Mean Square Error Approximation; SRMR = Standardized Root Mean Residual.
*p < .05

Next, the gender and race group models were tested for factor loading invariance. As in the cross-grade comparisons, model A imposed no constraints and model C fixed the factor loadings to be equal across gender or across race. Table 4.14 shows the fit statistics for both models.

Table 4.14

Fit Indices for Model A and Model C

Model	χ^2	Df	χ^2_{diff}	AIC	GFI	RMSEA	SRMR
Males and Females							
A	10.75	8		54.75	.99	.03	.01
C	13.70	11		51.70	.99	.03	.02
			2.95				
Blacks and Whites							
A	16.96	8		60.96	.98	.05	.03
C	19.31	11		57.31	.98	.04	.04
			2.35				

Note. AIC = Akaike Information Criterion; GFI = Goodness of Fit Index; RMSEA = Root Mean Square Error Approximation; SRMR = Standardized Root Mean Residual.

In both comparisons, Model C fit the data as well as Model A, which suggests that the factor loading for the groups can be considered invariant. The relationship between the reading and writing factors and their corresponding indicators is not significantly different

between males and females or between Whites and Blacks. The specific factor loading information for the groups is shown in Tables 4.15 and 4.16.

Table 4.15

Factor Loadings for Model Comparisons between Male and Female Students

	Factor Loadings	Standard Error	Critical Ratio	Standardized Regression Weight	Squared Multiple Correlations
Model A (Model C) Male					
Reading → R1	1.00			.85 (.85)	.73 (.72)
Reading → R2	1.08 (1.10)	.10 (.07)	9.8 (16.08)	.86 (.86)	.73 (.75)
Writing → I	1.00			.81 (.82)	.66 (.68)
Writing → P	1.14 (1.08)	.08 (.05)	13.86 (20.22)	.92 (.91)	.85 (.83)
Writing → N	.84 (.86)	.07 (.05)	12.22 (18.51)	.80 (.82)	.64 (.67)
Model A (Model C) Female					
Reading → R1	1.00			.84 (.84)	.70 (.70)
Reading → R2	1.11 (1.10)	.09 (.07)	12.67 (16.08)	.91 (.90)	.82 (.82)
Writing → I	1.00			.87 (.86)	.76 (.75)
Writing → P	1.02 (1.08)	.07 (.05)	14.35 (20.22)	.83 (.85)	.64 (.71)
Writing → N	.89 (.86)	.06 (.05)	13.96 (18.51)	.81 (.79)	.66 (.63)

Table 4.16

Factor Loadings for Model Comparisons Between Black and White Students

	Factor Loadings	Standard Error	Critical Ratio	Standardized Regression Weight	Squared Multiple Correlations
Model A (Model C) Black					
Reading → R1	1.00			.87 (.83)	.78 (.69)
Reading → R2	.96 (1.11)	.14 (.09)	6.99 (11.78)	.81 (.86)	.65 (.73)
Writing → I	1.00			.83 (.82)	.69 (.67)
Writing → P	.99 (1.05)	.10 (.06)	10.05 (17.98)	.80 (.83)	.64 (.68)
Writing → N	.86 (.84)	.09 (.05)	9.55 (16.62)	.75 (.73)	.56 (.53)
Model A (Model C) White					
Reading → R1	1.00			.77 (.79)	.59 (.63)
Reading → R2	1.19 (1.11)	.13 (.09)	9.33 (11.78)	.86 (.84)	.74 (.71)
Writing → I	1.00			.83 (.83)	.70 (.70)
Writing → P	1.08 (1.05)	.07 (.06)	14.89 (17.98)	.89 (.88)	.79 (.78)
Writing → N	.83 (.84)	.06 (.05)	13.63 (16.62)	.81 (.82)	.66 (.67)

There are slight differences between groups in the factor loading of the discourse modes, but these are smaller differences than between grade level groups. For male, female, and White groups, the largest loading is from writing to persuasive and the smallest is from writing to narrative. For the Black group, the size of the loading for persuasive and informative are virtually identical (.99 and 1.00 respectively) and the loading for narrative is the smallest.

The largest difference between the groups seems to be in the overall strength of the relationship between the reading and writing factors. As seen in Table 17, the strongest relationship between reading and writing is for the female students and the weakest relationship is for the Black students.

Table 4.17

Standardized Loading between Reading and Writing Factors by Group

Group	Reading-Writing Standardized Loading (Correlation)
Male	.65
Female	.73
Black	.54
White	.68

In general, the comparisons for males and females, Blacks and Whites show that the factor structure for the groups is not statistically different, although the data for the Black students did not fit the model well when run independently of the White group. The Black group also exhibited the weakest relationship between reading comprehension and writing quality scores, which further distinguishes the group of Black students from the other groups.

Summary

Three primary research questions were addressed using data from reading comprehension tests and writing quality scores on student essays in grades 4, 6, 8, 10, and 12. The first research question concerned the relationship between reading comprehension and writing quality. This relationship was shown to be strong in all grades, with disattenuated correlations ranging from a low of .59 in grade 6 to a high of .86 in grade 10. Thus, the shared construct variation between reading and writing accounted for by the other ranged from 34% in grade 6 to 74% in grade 10, with an average across grades of 59%.

The second research question concerned the influence of the discourse mode on the relationship between reading and writing. It also addressed the degree to which the relationship between writing and the indicator variables varied across grades. These questions were addressed using SEM. First, a simple model was proposed to examine how the discourse mode affected the reading-writing relationship. Second, all essays were used as part of the writing model, and it was determined that writing should be modeled as one factor. Next, a confirmatory factor analysis was conducted to evaluate a factor model that included both reading and writing as separate factors. The fit of a model when the reading-writing covariance or writing-indicator variable factor loadings were constrained to be equal across adjacent grades was also evaluated to determine whether the structure was consistent across grades and whether a developmental trend would emerge. Data strongly supported the invariance of the factor structure across grades 6, 8, 10, and 12 and, to a lesser degree, supported the invariance across grades 4 and 6. A closer look at the factor loadings for writing and the various discourse modes revealed a slight developmental trend across grades, with narrative writing quality scores being most influenced by the writing factor in grade 4,

followed by informational in grade 6 and persuasive in grade 12. In grades 8 and 10, both informative and persuasive writing were nearly identically affected by the writing factor, with narrative writing scores being the least affected.

The third research question addressed comparisons between male, female, Black, and White students. The same models that were used to compare grade level groups were used to compare the gender and race groups. Results showed that the factor structure for the reading-writing model was invariant across gender and race and that the pattern of factor loadings was also similar. However, the reading-writing model did not fit the Black group as well as it fit the other groups. In addition, the correlation between the reading and writing factors was lowest for the Black group.

The next chapter discusses these research findings in light of relevant research on reading comprehension and writing quality. Implications of these findings and suggestions for future research studies are also addressed.

CHAPTER 5

CONCLUSIONS AND DISCUSSION

The relationship between reading and writing has long been the focus of study, although various researchers have approached the connection between reading and writing from different theoretical perspectives. In general, although some have posited that reading and writing are different manifestations of single ability (Mehta, Foorman, Branum-Martin, & Taylor, 2005), the majority of existing research shares supports the notion articulated by Fitzgerald and Shanahan (2000) that "reading and writing rely on corresponding or correlated mental processes and isomorphic knowledge, though the nature of the relations between reading and writing is different at different age or grade levels" (p. 42). Further research into the nature of this relationship is theoretically relevant and can also yield practical benefits such as additional insights that may be used to design more effective instructional programs for teachers that enhance student ability to comprehend what they read and produce essays of high quality. The research presented here was undertaken in this context. Much of the previous research on the relationship between reading comprehension and writing quality has focused on precursor skills (e.g., alphabet knowledge, phonemic awareness, vocabulary knowledge). In contrast, although the current research focused on the actual targets of much instruction and many high-stakes assessments: reading comprehension and student writing products via essay scores.

The following sections of this chapter first address the study's limitations and then summarize some of the key findings and interpretations of those findings. Next, a discussion

of some factors that may have influenced the findings is presented followed by implications of the research. Finally, the chapter concludes with suggestions for future research in this area.

Limitations

Before addressing the key findings of this research, it is important to acknowledge some limitations of the study sample, design, and analysis. First, as discussed in Chapter 3, although the sample size in each grade was adequately large to conduct the research analyses, larger sample sizes would support more confidence in the results. Additionally, although the demographics of the sample closely matched those of the student population in the state where the study was conducted (see Table 3.1), the proportion of Black and White students in the sample differs from the national student population. Given the demographic characteristics of the sample, generalizing the results of this study beyond the context and characteristics of the data used here should be done with caution.

A second limitation is that this study did not seek to yield focused diagnostic results. The use of tests of overall reading comprehension and essay writing alone as indicators of the reading comprehension and writing quality latent constructs limited the diagnostic information that was provided. In the future, the inclusion of tests for precursor skills such as vocabulary knowledge or spelling, or the use of trait rather than holistic scoring for the essays may shed additional light on the relationship between reading comprehension and writing quality. Nonetheless, the decision to use indicators that focus on the outcome measures that are typically the target of instruction enabled this study to serve as an omnibus measure of the overall relationship between reading and writing. Any interpretation of these

results should be made within the context of the instruments and procedures used to measure the reading comprehension and writing quality constructs.

Overall, these limitations pose weak threats to the ability of this study to yield valid findings. Those findings, along with plausible interpretations are discussed in the following section.

Research Summary and Interpretations

This study had six major findings related to the three primary research questions. Research question 1 investigated the strength of the relationship between reading and writing quality at grades 4, 6, 8, 10, and 12 and the degree to which the strength of the relationship was similar across adjacent grades. This study found a strong relationship between reading and writing quality at all grade levels studied, but the strongest relationships at grades 8, 10, and 12. The amount of shared variance between reading comprehension and writing quality ranged from a low of 34% in grade 6 to a high of 74% in grade 10. SEM analysis revealed that the between-grades reading-writing relationship was significantly different only between grades 6 and 8. The difference in the strength of the relationship between reading and writing quality between all other adjacent grades was not significant.

Research question 2 examined whether the strength of the relationship between reading and writing quality was affected by the mode of the writing prompts (narrative, informational, persuasive). SEM results showed that the mode of the writing prompt used to elicit the student essay had a slight impact on the strength of the relationship between reading and writing quality, but the pattern varied across grades. No single discourse mode created a stronger relationship between reading and writing quality for all of the grades studied.

Research question 2a examined whether the discourse modes were similar in difficulty. This study found that the discourse modes varied slightly in difficulty, with narrative tending to be the easiest and persuasive the hardest, but the pattern was not consistent across all six essay scores or across grades.

Research question 2b examined the structure of writing quality as a latent variable to determine whether the three discourse modes (narrative, informational, persuasive) should be modeled separately, as three factors, or modeled as one overall writing factor. SEM analysis determined that writing quality should be modeled as one factor. An extension of this analysis examined whether the factor loadings from writing quality to the discourse modes were invariant across adjacent grades. Results showed that the discourse modes varied slightly in the relative size of their factor loadings from writing quality. The pattern showed a minor developmental trend, but shifts in the pattern were only significant between grades 4 and 6.

Research question 3 focused gender and race groups rather than grade levels. This analysis addressed the strength of the relationship between reading and writing quality as well as the relationship between writing quality and the discourse mode of the prompt. Results showed that the strength of the relationship between reading and writing quality was stronger for females than males and stronger for Black students than for White students. No significant differences were found in the patterns of factor loadings between writing quality and the discourse modes for male and female students nor for Black and White students.

Each of these findings is described in greater detail below and presented according to the respective research questions addressed in this study.

The first research question explored the relationship between reading comprehension and writing quality across grades 4, 6, 8, 10, and 12. Correlational analyses showed that the amount of shared variance between reading comprehension and writing quality, based on disattenuated correlations, was fairly high in all grades: G4 = 47%, G6 = 34%, G8 = 71%, G10 = 74%, and G12 = 71%. At only one grade level (grade 6) did the degree of shared variance fall within the range of below 40%, which is the level of shared variance reported by the majority of previous studies of the reading-writing connection (Shanahan, 2006). However, the amount of shared variance for the other four grades was greater than most of the previous studies, a noteworthy result given the large amount of previous research on the reading-writing connection. As in Berninger et al.'s (2002) research, which also showed high levels of shared variance (65-66% range for text-level comprehension and composition), this may have been due to the use of multiple indicators of reading comprehension and writing quality to represent each construct. Additionally, the research reported here addressed the issue of differential rater severity by using FACETS to adjust the student essay scores. This adjustment minimized the impact of differential rater severity on the essay scores. Reducing the confounding factor of human raters on the writing scores may have enabled more shared variance between reading and writing to be identified.

The shift from lower levels of shared variance in grades 4 and 6 to higher levels in grades 8, 10, and 12 suggest a developmental trend in the relationship between reading comprehension and writing quality. Previous research on the relationship between reading and writing has shown that as readers and writers develop the relationship between the components of reading and writing also change. Shanahan (1984) found that as readers became more proficient, the importance of sophisticated vocabulary and story structure to

writing achievement increased as did the importance of comprehension of larger units of text to reading achievement. Juel (1988) found that as students moved from 1st through 4th grade, the poor readers appeared to become poor writers. She found that the impact of ideas (based on a measure of oral story-telling) on writing quality increased from 8% of variance in 1st grade to 30% of the variance in 4th grade. Juel also found large differences in the amount of text read by good readers and writers and poor readers and writers. By the end of 4th grade, good readers and writers had read more than twice the number of words in running text in their basal readers than had the poor readers and writers. Juel linked the more frequent reading experiences of the good readers to better story ideas as well as knowledge of story structures and vocabulary with which to express those ideas, thus the good readers were also better writers. Juel also found that the correlations between reading comprehension and writing increased from .27 in 1st grade, to .39 in 2nd grade, .43 in 3rd grade and .52 in 4th grade, indicating that as students become more proficient readers and writers, the connections between the construct increases.

As students move through school and advance to higher grade levels, typically they are exposed in school to writing that is more complex and varied in terms of structure and vocabulary use than the writing encountered in previous grades. Eckhoff (1984) found that the characteristics of second-grade students' reading material influenced the characteristics of their students' writing samples. However, an awareness of these characteristics is affected by years of schooling. In a study of elementary school students, Korat and Schiff (2005) found that children's grade level, even more than their book-reading experience, was the most important predictor of children's knowledge of so-called good writing and writing difficulties

and also of their writing self-efficacy. Shanahan and Fitzgerald (2000) pointed out that reading and writing are not identical processes:

> Moving from reading to writing or from writing to reading is not like reversing directions on the same road. The differences in functional starting points can be enough to require different roads altogether. Consequently, reading is a somewhat easier task that writing. (p. 43)

It is possible that the increased relationship between reading ability and writing quality in higher grades is a function of students' writing ability and the attendant precursor skills improving and closing the gap between the two skills. A student's ability to use more complex vocabulary and conventions when writing and to show greater facility with writing structures causes writing quality (e.g., scores on writing tests) to align more with the student's ability to comprehend these words and structures when reading text (e.g., scores on reading tests). This provides a theoretical explanation for the larger correlation between reading comprehension and writing quality as seen in later grades and found in the research reported here.

The second research question addressed the impact of discourse mode on the reading-writing relationship as well as the relative difficulty of the discourse mode. Three discourse modes were examined: narrative, informational, and persuasive. An examination of the bivariate correlations of essay scores between the modes revealed no clear pattern of correlation either within mode or between modes. Essay score mean comparisons showed a slight trend of narrative prompts producing higher scores than persuasive prompts. Structural equation analysis using scores from a single discourse mode to model writing showed that the correlations between reading and writing varied depending upon the mode of the essay. However, the variation was fairly inconsistent across grades. Again, only a slight trend appeared, with narrative essays producing slightly larger correlations between reading

comprehension and writing quality in more grades than either informative or persuasive essays.

It is important to note that just two essay scores in each mode were used in the analysis, so these results should be interpreted with caution. More indicator variables for each latent construct may produce clearer trends. Engelhard et al. (1994) found that narrative prompts were easier than either descriptive or expository, but their research included only one prompt per student. Studies that use more prompts per discourse mode may shed more light on the issue of relative difficulty of discourse mode, although because of time, fatigue, and other concerns, it is often impractical or inappropriate to ask students—particularly those at lower grade levels—to produce more writing samples, especially if the writing is not part of regular instructional activities.

The next set of analyses in SEM modeled writing with all three discourse modes as indicator variables for writing. These analyses more fully explored the relationship between reading and writing across grades as well as the degree to which writing quality was reflected in the writing of different discourse modes. An initial analysis of two possible ways to model writing revealed that writing was best modeled as one latent factor with the three discourse modes as indicator variables. Next, two alternate models were compared for adjacent grades. In one model, the covariances between reading and writing were allowed to vary freely. In the second model, the covariances between reading comprehension and writing quality were constrained to be equal. This analysis explored whether the reading-writing relationship was similar in adjacent grades and simply replicated within an SEM framework the bivariate correlational analysis conducted to address research question 1. The comparison between grades 6 and 8 (the pair of grades for which the model constrained the covariances to be

equal) did not fit. This result paralleled the bivariate correlational analysis, which showed the largest difference in the reading-writing correlation between grades 6 and 8. This result provides some support for previous research by showing a developmental trend in the relationships between reading and writing.

To explore further the possible differences between adjacent grades, and in particular the role of discourse mode, an analysis was conducted to determine whether writing as a latent construct manifested in writing scores similarly across grades. As in the previous analysis, two models were compared for adjacent grades. The first model imposed no constraints on the factor loadings between the writing latent construct and indicator variables (e.g., narrative essay scores, informative essay scores, persuasive essay scores). The second model set the factor loadings to be equal for both grades. This analysis revealed that the factor loadings for the writing indicator variables did not differ significantly between grades except between grades 4 and 6. A closer look at the results showed that in grade 4, the loading from writing to narrative was the highest of the three modes (1.24), whereas in grade 6, it was the lowest (.83). In fact, in grades 8, 10, and 12, the factor loading for narrative was the lowest of the three modes. A higher factor loading between a latent construct and an indicator variable means that whatever skills students have in that latent construct (e.g., writing quality) are translated more to the measure represented by that indicator variable than those with lower factor loadings. Higher factor loadings reveal better indicators of the latent variable. In other words, in grade 4, narrative writing seems to be a better measure of student writing quality than either informative or persuasive writing. In all other grades, it seems to be the worst. However, the differences between the factor loadings were slight, so strong conclusions

should not be made. (The next section of this chapter discusses these findings in light of the developmental framework proposed by Fitzgerald & Shanahan, 2000).

A similar analysis was conducted to compare the factor loadings between writing quality and the indicator variables for male and female students and Black and White students. The model in which the factor loading for the writing quality indicator variables were constrained to be equal was not significantly different from the model in which the factor loadings were not constrained for either the male/female comparison or the Black/White comparison. Thus, for students within both groups, the writing quality construct generally seems to follow a similar pattern. It is worth noting that the loading between writing and narrative was the lowest for all groups, but this may have been because the grade 4 data were excluded from the gender and race analyses. The grade 4 data were excluded from the gender and race analyses because the change in the fit of the model for grades 4 and 6 became significantly worse ($p < .05$) when the factor loadings for the writing quality indicator variables were constrained to be equal. The fit did not change significantly for the other grades, indicating that the factor structure was stable across those grades, so it was appropriate to combine them for the gender and race comparisons.

Developmental Framework

Fitzgerald and Shanahan (2000) proposed a framework that outlined the stages in development shared by reading and writing. This framework was described in Chapter 2. Their framework addressed the cognitive process of reading and writing as they develop, but it can still be useful to note where their framework sheds light on the findings of the research presented here. Fitzgerald and Shanahan described six stages of developing reading and

writing. Stage 1 begins at birth and stage 6 describes reading and writing in college and beyond. Their stage 4 describes students in grades 4-8 and addresses reading and writing for learning the new. Stage 4 corresponds to the beginning of the developmental range of this study, and the description of the processes that emerge in stage 4 correspond to the trend found in factor loadings between writing quality and the three discourse modes. According to Fitzgerald and Shanahan, prior to stage 4, reading and writing focus primarily on narratives, but beginning in stage 4, "informational text become increasingly important" (p. 47). The present research indicates that between grades 4 and 6, the narrative essay permanently loses its status as the best indicator of student writing quality, which supports the shift described by Shanahan and Fitzgerald.

In grade 6, the largest factor loading is from writing quality to informative writing. Perhaps, to extend the ideas in Fitzgerald and Shanahan's (2000) framework beyond process to product, students in grade 6 able to show their understanding of informative text by writing in that mode. In grades 8 and 10, the factor loadings for informative and persuasive writing are nearly identical (G8: I = 1.00, P = 1.03; G10: I = 1.00, P = .99), perhaps reflecting another developmental shift.

According to Fitzgerald and Shanahan (2000), high school students are typically in stage 5. In this stage, developing readers and writers begin to know how to see from another's viewpoint and know how to analyze and critique when reading and writing. As writers apply their understanding of others' viewpoints, they may be better able to use their increasing skill in analyzing, critiquing, and revising their own work to produce persuasive essays. The results of this study support the emergence of persuasive writing as the best

indicator of student writing quality in grade 12, in which the factor loading for persuasive essays is the highest of the three modes.

It is interesting to note that the shifts in the relative strength of the factor loadings between writing quality and each discourse mode parallel key components of the developmental framework described by Fitzgerald and Shanahan (2000) throughout the grade range of the study sample. However, the structural equation model comparisons showed that the differences in factor loadings were significant only between grades 4 and 6, so these trends should not be over interpreted.

Impact of Instruction

Research has shown that children benefit both intellectually (Hart & Risley, 1992) and academically from exposure to words, print, and verbal stimuli when they are very young and from more and wider reading as students (Cunningham & Stanovich, 1991; Juel, 1988; Stanovich, 1986, 2000). Research has also shown that student reading practices have an impact on knowledge of writing and ability to write more complex text (Eckhoff, 1984; Korat & Schiff, 2005). When students enter school, much of their reading and writing experience is guided by instruction. Eckhoff (1984) found that the writing of second-grade students clearly showed features of the type of text contained in the students' basal readers. The more elaborate linguistic structures from one basal series was reflected in the writing of the students who read that series and not in the students who read a series containing less elaborate linguistic structures. Thus, the exposure of a student to various types of reading material may determine the skill with which the student can produce writing in various

discourse modes (Crowhurst, 1991; Duke, Bennett-Armistead, & Roberts, 2003; Kamberelis, 1999).

Mehta et al. (2005) examined the relative roles of teachers and students in predicting literacy outcomes for low SES students in kindergarten through grade 4. The literacy outcomes they studied included word reading, passage comprehension, and composition (i.e., writing). Mehta et al. found that of all of the literacy outcomes studied, teaching quality had the largest effect on writing. The other literacy outcomes were predicted primarily by a student's prior status on the measure. The teachers who taught writing well were those who went beyond mechanics to work on ideas, organization, language choice, and structure.

The findings of Mehta et al. (2005) support other research that shows that instruction can play an important role in the quality of student writing. The study of writing models was found to be a key element of effective adolescent instruction in a meta-analysis of writing instruction studies (Graham & Perin, 2007). Fitzgerald and Teasley (1986) found that instruction in narrative structure improved the overall quality of the students' narrative writing. Crowhurst (1991) found that students' ability to write persuasive essays was positively affected by both direct writing instruction in persuasion and the reading of persuasive text. However, it is not clear that instruction in one mode improves writing in other modes of discourse. For example, Troia and Graham (2002) found no instructional transfer from story writing to persuasive essay writing. The findings of Quallmelz et al. (1982) suggest that perhaps some components of writing transfer across discourse modes, whereas others do not. Quallmelz et al. found that scores for general impression, organization, and total were less stable between narrative and expository essays. Scores for

mechanics, focus, and support were unaffected by whether the mode was narrative or expository.

Does the pattern of writing in different discourse modes found in the present research reflect the instructional practices in place in the schools? Information about the curriculum in effect during the school year when the study was conducted (2005-2006) suggests that it may. The curriculum in place in Mississippi for grade 4 stated explicitly that students should be reading and writing for various purposes: to entertain, to inform, persuade, and describe (Thompson, Jones, Haynes, & Rucker, 2000). The focus on all three discourse modes continued in the curriculum through grade 12, which may help explain why there were only slight differences in student performance on the discourse modes across the grades and in the relative weight of the factor loadings from writing to the indicator variables for each mode.

At the same time, a high stakes statewide testing program was in place in Mississippi to monitor student ability in writing for statewide accountability purposes. Each year, students in grades 4 and 7 were expected to write one essay to a prompt that was either narrative, informative, or persuasive. In grade 10, students were expected to write two essays: one to a narrative prompt and one to an informative prompt. In the upper grades (8, 10, and 12), the factor loading from writing to narrative was the smallest, which runs counter to what would be expected by the emphasis on narrative writing in the statewide testing program.

Although the curriculum and statewide test clearly encouraged reading and writing instruction in narrative, informative, and persuasive modes in fourth grade, the curriculum in kindergarten through third grade emphasized study of narrative reading and writing. This early emphasis on narrative may help explain why, though the differences between modes

were slight, the factor loading between writing and narrative was the largest in grade 4 but not the other grades.

Implications

Two primary findings of this research have instructional implications. First, the finding that correlations between reading comprehension and writing quality were fairly high at all grade levels studied but were the highest in the upper grades (8, 10, 12) suggests that there is a greater influence of the processes and knowledge shared between reading and writing as students become more proficient in reading and writing. If this increased relationship between reading comprehension and writing quality is partially caused by improving student writing skills, then an early and strong emphasis on writing instruction may bring students to a point at which more benefit can be gained by combining reading and writing instruction. To quote Shanahan (1984), "The finding that the reading-writing relation changes with reading development suggest the possibility that writing curricula could be directly integrated into those materials currently used for the teaching of reading" (p. 475).

Tierney, Soter, O'Flahavan, and McGinley (1989) found that college students who wrote and read in combination exhibited more evidence of critical thinking than when reading and writing were separated or when reading was combined with knowledge activation or answering questions. If a goal of instruction is to increase students' critical-thinking activity, combining reading and writing in instruction in all content areas may be one way to do it. Students who have developed critical thinking skills, such as the college students in the previous study, or those who are developing those skills, such as students in the stage 5 described by Fitzgerald and Shanahan (2000), may gain additional benefits from the combination of reading and writing activities. Unfortunately, as Tierney et al. (1989) stated,

"The potential of reading and writing activities to enable thinking and learning often goes unrealized in instructional settings" (p. 136). Perhaps the research presented here will further encourage the use of integrated reading and writing activities in classrooms.

Second, the finding that the relationship between reading comprehension and writing quality is not markedly affected by the discourse mode of the writing prompt suggests that the shared variance between reading comprehension and writing quality may be of a general, rather than mode-specific, nature. Instruction in all modes of reading and writing can capitalize on the features of good reading and writing that are shared across discourse modes. By focusing on similarities among modes rather than their differences, it may be more likely that teachers faced with testing pressures will avoid focusing instruction only on the modes that are most likely to appear on statewide assessments. Some statewide testing programs include assessment of only one discourse mode per grade, though the statewide curriculum may require instruction in several discourse modes. A common complaint of teachers is that they feel pressured to teach only what is tested. For example, a grade 4 teacher may focus much of the reading and writing instruction in grade 4 on narrative if that is the mode tested on the statewide writing assessment. The present study does not directly address the transfer of skills from instruction in one mode to another, but the study findings support a close link between writing quality in the three modes addressed. These findings may support a more balanced approach to writing instruction in situations in which there is a lack of balance in areas tested in high-stakes testing programs.

Finally, the present study has methodological implications for future research. The use of a scoring procedure that adjusts for rater severity and produces more accurate measures of student writing quality may have identified more of the shared variance between reading

comprehension and writing quality. In addition, multiple essay scores for each student provide a more robust measure of student writing quality than found in most previous research. Adopting both of these study characteristics could benefit future research in the relationship between reading comprehension and writing quality.

Suggestions for Future Research

Just as this research was informed by the tremendous amount of work that preceded it, subsequent research can expand upon the issues that are herein brought to the fore. Few studies focus on the reading-writing connection in the middle and high school grades and fewer yet on the developmental nature of the relationship in those grades. The findings of this research point to the years between grades 4 and 8 as being noteworthy in terms of a shift in the strength of the relationship between reading comprehension and writing quality. Follow-up studies could examine the reasons and causes of the occurrence of this shift. One suggestion presented earlier is that students' writing skills become more sophisticated and close the gap between operationalizing reading comprehension processes and writing processes.

Additional work on the transfer of writing skills across modes is also needed. The more time teachers can spend on skills that transfer across discourse modes, the more effective overall instructional time can be. If teachers can ground students in skills that can be applied to writing and reading in all genres—skills such as idea generation, organization, focus, critique and conventions—the students will be better equipped for whichever future path they choose. By identifying which instructional strategies produce the best outcomes across discourse modes, we may encourage teachers to teach more broadly even when under

pressure to show evidence of student achievement in narrowly defined areas. Similarly, identifying instructional strategies that improve student writing in specific discourse modes but do not transfer to other modes would help teachers give appropriately instructional time to those strategies. For example, Fitzgerald and Teasley (1986) found that the improvements in student narrative writing due to direct instruction of narrative structure were realized quickly (six intensive 30–45 minute sessions over a period of two weeks) and were maintained over time.

Conclusion

This research study shows, like others before it, that there is a strong relationship between reading ability and writing quality. The procedures used in this study helped to reveal a stronger relationship than shown by many previous studies as well as a developmental shift in the relationship between grades 4 and 8. Analysis of the discourse modes of the writing prompts suggests only a minimal effect of mode. As pressure continues to mount for educators to show that their instructional programs are effective, capitalizing on the shared knowledge and processes of reading and writing, especially at the higher grades, may help them meet their instructional goals.

APPENDIX A

Writing Prompts

Prompt Label	Text of the Prompt
4AN	One morning a child looks out the window and discovers that a huge castle has appeared overnight. The child rushes outside to the castle and hears strange sounds coming from it. Someone is living in the castle! The castle door creaks open. The child goes in. Write a story about who the child meets and what happens inside the castle.
4AI	We all have favorite objects that we care about and would not want to give up. Think of one object that is important to you. For example, it could be a book, a piece of clothing, a game, or any object you care about. Write about your favorite object. Be sure to describe the object and explain why it is valuable or important to you.
4AP	Pretend you have a friend who is invisible and you would like other people to meet him or her. Write a letter to this invisible person. Convince your friend to become visible so others may meet him or her. In your letter, use details and examples.
4BN*	One day you wake up and go down to breakfast. You eat what you normally eat. Your breakfast is the last normal thing that happens to you all day.

Prompt

Label	Text of the Prompt
	Write a story called “The Very Unusual Day” about what happens that day, from right after breakfast until you go to bed again.
4BI	Describe what lunchtime is like for you on a school day. Be sure to tell about your lunchtime so that someone who has never had lunch with you on a school day can understand where you have lunch and what lunchtime is like.
4BP	Imagine this situation: Your favorite book is missing from your school library. It might be a book that you like to read over and over again. Or it might be a book that your teacher or parent has read to you. Some of your friends also like to read this book. The school librarian is not sure she wants to buy the book again. Write a letter to convince your school librarian to buy the book again. In your letter, give lots of reasons why the book should be in your school library.
8AN	Imagine this situation! A noise outside awakens you one night. You look out the window and see a spaceship. The door of the spaceship opens, and out walks a space creature. What does the creature look like? What do you do? Write a story about what happens next.

Prompt

Label	Text of the Prompt
8AI	A public television network is seeking ideas for a new series of shows that would be educational for teenagers. The series will include ten one-hour episodes and will be shown once a week. Some of the titles under consideration are: "Great Cities of the World" "Women in History" "Nature Walks" "American Legends" Choose one of these titles. Write a letter to the network describing your ideas for a new educational series. In your letter, describe what one episode might be like. Use specific examples of what information you would include in the episode so the network president will be able to imagine what the series would be like.
8AP	Many people think that students are not learning enough in school. They want to shorten most school vacations and make students spend more of the year in school. Other people think that lengthening the school year and shortening vacations is a bad idea because students use their vacations to learn important things outside of school. What is your opinion? Write a letter to your school board either in favor or against lengthening the school year. Give specific reasons to support your opinion that will convince the school board to agree with you.

Prompt Label	Text of the Prompt
8BN	Imagine that you wake up one morning to discover that you have become the President of the United States. Write a story about your first day as President.
8BI	A novel written in the 1950s describes a world where people are not allowed to read books. A small group of people who want to save books memorize them so that the books won't be forgotten. For example, an old man who has memorized the novel *The Call of the Wild* helps a young boy memorize it by reciting the story to him. In this way, the book is saved for the future. If you were told that you could save just one book for future generations, which book would you choose? Write an essay in which you discuss which book you would choose to save for future generations and what it is about the book that makes it important to save. Be sure to discuss in detail why the book is important to you and why it would be important to future generations.
8BP*	Suppose a research study showed that teenagers have low energy levels in the morning and that adults have low energy levels at night. The study recommends that teenagers should stay up later at night and sleep later in the morning. The study also recommends that adults go to bed earlier and get up earlier.

Prompt

Label Text of the Prompt

Write a letter to your principal arguing for or against the proposition that classes at your school should begin much later in the day. Be sure to give detailed reasons to support your argument and make it convincing.

12AN The following excerpt is from a poem by Walt Whitman.

There was a child who went forth every day,
And the first object he look'd upon, that
object he became,
And that object became part of him for
the day or a certain part of the day,
Or for many years or stretching cycles
of years.

Whitman's poem suggests that certain objects become important to us and remain important to us even if we no longer have them.

Write a story in which you tell about an object that remains important to the main character over a period of years. The main character could be you or someone you know.

In your story, describe the main character's first encounter with the object, why the object is so important to the character, and how, over the years, it remains a part of the character's life.

Prompt Label	Text of the Prompt
12AI	Your school has a program in which a twelfth grader acts as a mentor for a tenth grader at the beginning of each school year. The mentor's job is to help the tenth grader have a successful experience at your school. The tenth grader you are working with is worried about being able to write well enough for high school classes.
12 AN	Write a letter to your tenth grader explaining what kind of writing is expected in high school classes and what the student can do to be a successful writer in high school. As you plan your response, think about your own writing experiences. How would you describe "good" writing? What advice about writing has been helpful to you? What writing techniques do you use?
12AP	Your school is sponsoring a voter registration drive for 18-year-old high school students. You and three of your friends are talking about the project. Your friends say the following, Friend 1: "I'm working on the young voters' registration drive. Are you going to come to it and register? You're all 18, so you can do it. We're trying to help increase the number of young people who vote and it shouldn't be too hard—I read that the percentage of 18- to 30-year-olds who vote increased in recent years. We

Prompt Label	Text of the Prompt
	want that percentage to keep going up." Friend 2: "I'll be there. People should vote as soon as they turn 18. It's one of the responsibilities of living in a democracy." Friend 3: "I don't know if people should even bother to register. One vote in an election isn't going to change anything." Do you agree with friend 2 or 3? Write a response to your friends in which you explain whether you will or will not register to vote. Be sure to explain why and support your position with examples from your reading or experience. Try to convince the friend with whom you disagree that your position is the right one.
12BN*	A tall tale is a type of story that uses exaggeration to solve a real-life problem. As the story progresses, the main character demonstrates superhuman abilities to overcome ordinary obstacles. Imagine that you will participate in a "tall-tale writing contest" at your school. Write you own tall tale. You can write about yourself, someone you know, or someone you imagine. Be sure to give your main character whatever superhuman abilities are necessary to save the day.
12BI	A novel written in the 1950s describes a world where people are not allowed to read books. A small group of people who want to save books memorize them, so that the books won't be forgotten. For example, an old man who has memorized the

Prompt Label	Text of the Prompt
	novel *The Call of the Wild* helps a young boy memorize it by reciting the story to him. In this way, the book is saved for the future. If you were told that you could save just one book for future generations, which book would you choose? Write an essay in which you discuss which book you would choose to save for future generations and what it is about the book that makes it important to save. Be sure to discuss in detail why the book is important to you and why it would be important to future generations.
12BP	Who are our heroes? The media attention given to celebrities suggests that these people are today's heroes. Yet ordinary people perform extraordinary acts of courage every day that go virtually unnoticed. Are these people the real heroes? Write an essay in which you define heroism and argue who you think our heroes really are—mass-media stars, ordinary people, or maybe both. Be sure to use examples of specific celebrities, other people you have heard or read about, or people from your own community to support your position.

* The wording of the prompt was altered slightly from the NAEP wording.
Note: All of the A prompts were administered in 1998. All of the B prompts were administered in 2002.

APPENDIX B

Rubrics

Grade 4 Informative

Score Point	Description
6 Excellent	• Develops ideas well and uses specific, relevant details across the response. • Well organized with clear transitions. • Sustains varied sentence structure and exhibits specific word choices. • Exhibits control over sentence boundaries; errors in grammar, spelling, and mechanics do not interfere with understanding.
5 Skillful	• Develops ideas with some specific, relevant details. • Clearly organized; information is presented in an orderly way but response may lack transitions. • Exhibits some variety in sentence structure and exhibits some specific word choices. • Generally exhibits control over sentence boundaries; errors in grammar, spelling, and mechanics do not interfere with understanding.
4 Sufficient	• Clear but sparsely developed; may have few details. • Provides a clear sequence of information; provides pieces of information that are generally related to each other. • Generally has simple sentences and simple word choice; may exhibit uneven control over sentence boundaries. • Has sentences that consist mostly of complete, clear, distinct thoughts; errors in grammar, spelling, and mechanics generally do not interfere with understanding.

3 Uneven	May be characterized by one or more of the following: • Provides limited or incomplete information; may be list-like or have the quality of an outline. • Disorganized or provides a disjointed sequence of information. • Exhibits uneven control over sentence boundaries and may have some inaccurate word choices. • Errors in grammar, spelling, and mechanics sometimes interfere with understanding.
2 Insufficient	May be characterized by one or more of the following: • Provides little information and makes little attempt at development. • Very disorganized or too brief to detect organization. • Exhibits little control over sentence boundaries and sentence formation; word choice is inaccurate in much of the response. • Characterized by misspellings, missing words, incorrect word order; errors in grammar, spelling, and mechanics are severe enough to make understanding very difficult in much of the response.
1 Unsatisfactory	May be characterized by one or more of the following: • Attempts a response, but may only paraphrase the prompt or be extremely brief. • Exhibits no control over organization. • Exhibits no control over sentence formation; word choice is inaccurate across the response. • Characterized by misspellings, missing words, incorrect word order; errors in grammar, spelling, and mechanics severely impede understanding across the response.

Grade 8 Informative

Score Point	Description
6 Excellent	• Develops and shapes information with well-chosen details across the response. • Well organized with strong transitions. • Sustains variety in sentence structure and exhibits good word choice. • Errors in grammar, spelling, and punctuation are few and do not interfere with understanding.
5 Skillful	• Develops and shapes information with details in parts of the response. • Clearly organized, but may lack some transitions and/or have occasional lapses in continuity. • Exhibits some variety in sentence structure and some good word choices. • Errors in grammar, spelling, and punctuation do not interfere with understanding.
4 Sufficient	• Develops information with some details. • Organized with ideas that are generally related, but has few or no transitions. • Exhibits control over sentence boundaries and sentence structure, but sentences and word choice may be simple and unvaried. • Errors in grammar, spelling, and punctuation do not interfere with understanding.
3 Uneven	May be characterized by one or more of the following: • Presents some clear information, but is list-like, undeveloped, or repetitive OR offers no more than a well-written beginning. • Unevenly organized; the response may be disjointed. • Exhibits uneven control over sentence boundaries and sentence structure; may have some inaccurate word choices.

	• Errors in grammar, spelling, and punctuation sometimes interfere with understanding.
2 Insufficient	May be characterized by one or more of the following: • Presents fragmented information OR may be very repetitive OR may be very undeveloped. • Very disorganized; thoughts are tenuously connected OR the response is too brief to detect organization. • Minimal control over sentence boundaries and sentence structure; word choice may often be inaccurate. • Errors in grammar or usage (such as missing words or incorrect word use or word order), spelling, and punctuation interfere with understanding in much of the response.
1 Unsatisfactory	May be characterized by one or more of the following: • Attempts to respond to prompt, but provides little or no coherent information; may only paraphrase the prompt. • Has no apparent organization OR consists of a single statement. • Minimal or no control over sentence boundaries and sentence structure; word choice may be inaccurate in much or all of the response. • A multiplicity of errors in grammar or usage (such as missing words or incorrect word use or word order), spelling, and punctuation severely impedes understanding across the response.

Grade 12 Informative

Score Point	Description
6 Excellent	• Information is presented effectively and consistently supported with well-chosen details. • Information is focused and well organized, with a sustained controlling idea and effective use of transitions. • Response consistently exhibits variety in sentence structure and precision in word choice. • Errors in grammar, spelling, and punctuation are few and do not interfere with understanding.
5 Skillful	• Information is presented clearly and supported with pertinent details in much of the response. • Response is well organized, but may lack some transitions. • Response exhibits some variety in sentence structure and uses good word choice; occasionally, words may be used inaccurately. • Errors in grammar, spelling, and punctuation do not interfere with understanding.
4 Sufficient	• Information is presented clearly and supported with some pertinent details. • Information is generally organized, but has few or no transitions among parts. • Sentence structure may be simple and unvaried; word choice is mostly accurate. • Errors in grammar, spelling, and punctuation do not interfere with understanding.
3 Uneven	May be characterized by one or more of the following: • Information is presented clearly in parts; other parts are undeveloped or repetitive OR is no more than a well-written beginning. • Organized in parts of the response; other parts are disjointed and/or lack transitions.

	• Exhibits uneven control over sentence boundaries and sentence structure; may exhibit some inaccurate word choices. • Errors in grammar, spelling, and punctuation sometimes interfere with understanding.
2 Insufficient	May be characterized by one or more of the following: • Information is presented clearly in parts; other parts are undeveloped or repetitive OR is no more than a well-written beginning. • Organized in parts of the response; other parts are disjointed and/or lack transitions. • Exhibits uneven control over sentence boundaries and sentence structure; may exhibit some inaccurate word choices. • Errors in grammar, spelling, and punctuation sometimes interfere with understanding.
1 Unsatisfactory	May be characterized by one or more of the following: • Responds to prompt but may be incoherent OR provides very minimal information OR merely paraphrases the prompt. • Little or no apparent organization. • Minimal or no control over sentence boundaries and sentence structure; word choice may be inaccurate in much or all of the response. • Errors in grammar, spelling, and punctuation severely impede understanding across the response.

Grade 4 Narrative

Score Point	Description
6 Excellent	• Tells a well-developed story with relevant descriptive details across the response. • Events are well connected and tie the story together with transitions across the response. • Sustains varied sentence structure and exhibits specific word choices. • Exhibits control over sentence boundaries; errors in grammar, spelling, and mechanics do not interfere with understanding.
5 Skillful	• Tells a clear story with some development, including some relevant descriptive details. • Events are connected in much of the response; may lack some transitions. • Exhibits some variety in sentence structure and exhibits some specific word choices. • Generally exhibits control over sentence boundaries; errors in grammar, spelling, and mechanics are minor and do not interfere with understanding.
4 Sufficient	• Tells a clear story with little development; has few details. • Events are generally related; may contain brief digressions or inconsistencies. • Generally has simple sentences and simple word choice; may exhibit uneven control over sentence boundaries. • Has sentences that consist mostly of complete, clear, distinct thoughts; errors in grammar, spelling, and mechanics generally do not interfere with understanding.
3 Uneven	• Attempts to tell a story but tells only part of a story, gives a plan for a story, or is list-like. • Lacks a clear progression of events; elements may not fit together or be in

	sequence. • Exhibits uneven control over sentence boundaries and may have some inaccurate word choices. • Errors in grammar, spelling, and mechanics sometimes interfere with understanding.
2 Insufficient	May be characterized by one or more of the following: • Attempts a response, but is no more than a fragment or the beginning of a story OR is very repetitive. • Very disorganized or too brief to detect organization. • Exhibits little control over sentence boundaries and sentence formation; word choice is inaccurate in much of the response. • Characterized by misspellings, missing words, incorrect word order; errors in grammar, spelling, and mechanics are severe enough to make understanding very difficult in much of the response.
1 Unsatisfactory	May be characterized by one or more of the following: • Attempts a response, but may only paraphrase the prompt or be extremely brief. • Exhibits no control over organization. • Exhibits no control over sentence formation; word choice is inaccurate across the response. • Characterized by misspellings, missing words, incorrect word order; errors in grammar, spelling, and mechanics severely impede understanding across the response.

Grade 8 Narrative

Score Point	Description
6 Excellent	• Tells a clear story that is well-developed and shaped with well-chosen details across the response. • The story is well organized with strong transitions. • Sustains variety in sentence structure and exhibits good word choice. • Errors in grammar, spelling, and punctuation are few and do not interfere with understanding.
5 Skillful	• Tells a clear story that is developed and shaped with details in parts of the response. • The story is clearly organized, but may lack some transitions and/or have occasional lapses in continuity. • Exhibits some variety in sentence structure and some good word choices. • Errors in grammar, spelling, and punctuation do not interfere with understanding.
4 Sufficient	• Tells a clear story that is developed with some details. • The parts of the story are generally related, but there are few or no transitions. • Exhibits control over sentence boundaries and sentence structure, but sentences and word choice may be simple and unvaried. • Errors in grammar, spelling, and punctuation do not interfere with understanding.
3 Uneven	May be characterized by one or more of the following: • Attempts to tell a story, but parts of the story are unclear, undeveloped, list-like, or repetitive OR offers no more than a well-written beginning. • Unevenly organized; parts of the story may be unrelated to one another.

	• Exhibits uneven control over sentence boundaries and sentence structure; may have some inaccurate word choices. • Errors in grammar, spelling, and punctuation sometimes interfere with understanding.
2 Insufficient	May be characterized by one or more of the following: • Attempts to tell a story, but the attempt may be a fragment and/or very undeveloped. • Very disorganized throughout the response OR too brief to detect organization. • Minimal control over sentence boundaries and sentence structure; word choice may often be inaccurate. • Errors in grammar or usage (such as missing words or incorrect word use or word order), spelling, and punctuation interfere with understanding in much of the response.
1 Unsatisfactory	May be characterized by one or more of the following: • Responds to prompt, but provides little or no coherent content OR merely paraphrases the prompt. • Has no apparent organization OR consists of a single statement. • Minimal or no control over sentence boundaries and sentence structure; word choice may be inaccurate in much or all of the response. • A multiplicity of errors in grammar or usage (such as missing words or incorrect word use or word order), spelling, and punctuation severely impedes understanding across the response.

Grade 12 Narrative

Score Point	Description
6 Excellent	• Tells a clear story that is consistently well developed and detailed; details enhance story being told. • Is well organized; integrates narrative events into a smooth telling; effective transitions move the story forward. • Consistently exhibits variety in sentence structure and precision in word choice. • Errors in grammar, spelling, and punctuation are few and do not interfere with understanding.
5 Skillful	• Tells a clear story that is well developed and elaborated with details in much of the response. • Is well organized with story elements that are connected across most of the response; may have occasional lapses in transitions. • Exhibits some variety in sentence structure and uses good word choice; occasionally, words may be used inaccurately. • Errors in grammar, spelling, and punctuation do not interfere with understanding
4 Sufficient	• Tells a clear story that is developed with some pertinent details. • Is generally organized, but transitions among parts of the story may be lacking. • Sentence structure may be simple and unvaried; word choice is mostly accurate. • Errors in grammar, spelling, and punctuation do not interfere with understanding.
3 Uneven	May be characterized by one or more of the following: • Tells a story that may be clear and developed in parts; other parts are unfocused, repetitive, or minimally developed OR response is no more than a well-written beginning.

	• Is organized in parts of the response; other parts are disjointed and/or lack transitions. • Exhibits uneven control over sentence boundaries and sentence structure; may exhibit some inaccurate word choices. • Errors in grammar, spelling, and punctuation sometimes interfere with understanding.
2 Insufficient	May be characterized by one or more of the following: • Attempts to tell a story, but is very undeveloped, list-like, or fragmentary. • Is disorganized or unfocused in much of the response OR the response is too brief to detect organization. • Minimal control over sentence boundaries and sentence structure; word choice may often be inaccurate. • Errors in grammar, spelling, and punctuation interfere with understanding in much of the response.
1 Unsatisfactory	May be characterized by one or more of the following: • Responds to prompt but provides little or no coherent content OR merely paraphrases the prompt. • Has little or no apparent organization. • Minimal or no control over sentence boundaries and sentence structure; word choice may be inaccurate in much or all of the response. • Errors in grammar, spelling, and punctuation severely impede understanding across the response.

Grade 4 Persuasive

Score Point	Description
6 Excellent	• Takes a clear position and develops support with well-chosen details, reasons, or examples across the response. • Well organized; maintains focus. • Sustains varied sentence structure and exhibits specific word choices. • Exhibits control over sentence boundaries; errors in grammar, spelling, and mechanics do not interfere with understanding.
5 Skillful	• Takes a clear position and develops support with some specific details, reasons, or examples. • Provides some organization of ideas by, for example, using contrast or building to a point. • Exhibits some variety in sentence structure and exhibits some specific word choices. • Generally exhibits control over sentence boundaries; errors in grammar, spelling, and mechanics do not interfere with understanding.
4 Sufficient	• Takes a clear position with support that is clear and generally related to the issue. • Generally organized. • Generally has simple sentences and simple word choice; may exhibit uneven control over sentence boundaries. • Has sentences that consist mostly of complete, clear, distinct thoughts; errors in grammar, spelling, and mechanics generally do not interfere with understanding.
3 Uneven	May be characterized by one or more of the following: • Takes a position and offers limited or incomplete support; some reasons may not

	be clear or related to the issue. • Disorganized or provides a disjointed sequence of information. • Exhibits uneven control over sentence boundaries and may have some inaccurate word choices. • Errors in grammar, spelling, and mechanics sometimes interfere with understanding.
2 Insufficient	May be characterized by one or more of the following: • Takes a position, but provides only minimal support (generalizations or a specific reason or example); OR attempts to take a position but the position is unclear. • Very disorganized or too brief to detect organization. • May exhibit little control over sentence boundaries and sentence formation; word choice is inaccurate in much of the response. • Characterized by misspellings, missing words, incorrect word order. Errors in grammar, spelling, and mechanics may be severe enough to make understanding very difficult in much of the response.
1 Unsatisfactory	May be characterized by one or more of the following: • Takes a position but provides no support OR attempts to take a position (is on topic) but position is very unclear; may only paraphrase the prompt. • Exhibits no control over organization. • Exhibits no control over sentence formation; word choice is inaccurate across the response. • Characterized by misspellings, missing words, incorrect word order; errors in grammar, spelling, and mechanics severely impede understanding across the response.

Grade 8 Persuasive

Score Point	Description
6 Excellent	• Takes a clear position and develops it consistently with well-chosen reasons and/or examples across the response. • Well organized with strong transitions. • Sustains variety in sentence structure and exhibits good word choice. • Errors in grammar, spelling, and punctuation are few and do not interfere with understanding.
5 Skillful	• Takes a clear position and develops it with reasons and/or examples in parts of the response. • Clearly organized, but may lack some transitions and/or have occasional lapses in continuity. • Exhibits some variety in sentence structure and some good word choices. • Errors in grammar, spelling, and punctuation do not interfere with understanding.
4 Sufficient	• Takes a clear position and supports it with some reasons and/or examples. • Organized with ideas that are generally related, but there are few or no transitions. • Exhibits control over sentence boundaries and sentence structure, but sentences and word choice may be simple and unvaried. • Errors in grammar, spelling, and punctuation do not interfere with understanding.
3 Uneven	May be characterized by one or more of the following: • Takes a position and offers support, but may be unclear, repetitive, list-like, or undeveloped. • Unevenly organized; the response may be disjointed. • Exhibits uneven control over sentence boundaries and sentence structure; may

	have some inaccurate word choices. • Errors in grammar, spelling, and punctuation sometimes interfere with understanding.
2 Insufficient	May be characterized by one or more of the following: • Takes a position, but may be very unclear, very undeveloped, or very repetitive. • Very disorganized; thoughts are tenuously connected OR the response is too brief to detect organization. • Minimal control over sentence boundaries and sentence structure; word choice may often be inaccurate. • Errors in grammar or usage (such as missing words or incorrect word use or word order), spelling, and punctuation interfere with understanding in much of the response.
1 Unsatisfactory	May be characterized by one or more of the following: • Attempts to take a position (addresses topic) but is incoherent OR takes a position but provides no support; may only paraphrase the prompt. • Has no apparent organization OR consists of a single statement. • Minimal or no control over sentence boundaries and sentence structure; word choice may be inaccurate in much or all of the response. • A multiplicity of errors in grammar or usage (such as missing words or incorrect word use or word order), spelling, and punctuation severely impedes understanding across the response.

Grade 12 Persuasive

Score Point	Description
6 Excellent	• Takes a clear position and supports it consistently with well-chosen reasons and/or examples; may use persuasive strategy to convey an argument. • Focused and well organized, with effective use of transitions. • Consistently exhibits variety in sentence structure and precision in word choice. • Errors in grammar, spelling, and punctuation are few and do not interfere with understanding.
5 Skillful	• Takes a clear position and supports it with pertinent reasons and/or examples through much of the response. • Well organized, but may lack some transitions. • Exhibits some variety in sentence structure and uses good word choice; occasionally, words may be used inaccurately. • Errors in grammar, spelling, and punctuation do not interfere with understanding.
4 Sufficient	• Takes a clear position and supports it with some pertinent reasons and/or examples; there is some development. • Generally organized, but has few or no transitions among parts. • Sentence structure may be simple and unvaried; word choice is mostly accurate. • Errors in grammar, spelling, and punctuation do not interfere with understanding.
3 Uneven	May be characterized by one or more of the following: • Takes a position and provides uneven support; may lack development in parts or be repetitive OR is no more than a well-written beginning. • Organized in parts of the response; other parts are disjointed and/or lack transitions.

		• Exhibits uneven control over sentence boundaries and sentence structure; may exhibit some inaccurate word choices. • Errors in grammar, spelling, and punctuation sometimes interfere with understanding.
2	May be characterized by one or more of the following:	
Insufficient		• Takes a position but is very undeveloped. • Disorganized or unfocused in much of the response OR clear but very brief. • Minimal control over sentence boundaries and sentence structure; word choice may often be inaccurate. • Errors in grammar, spelling, and punctuation interfere with understanding in much of the response.
1	May be characterized by one or more of the following:	
Unsatisfactory		• Attempts to take a position (addresses topic) but position is very unclear OR takes a position but provides minimal or no support; may only paraphrase the prompt. • Little or no apparent organization. • Minimal or no control over sentence boundaries and sentence structure; word choice may be inaccurate in much or all of the response. • Errors in grammar, spelling, and punctuation severely impede understanding across the response.

APPENDIX C

Standardized and Unstandardized Estimates for the One-Factor Model

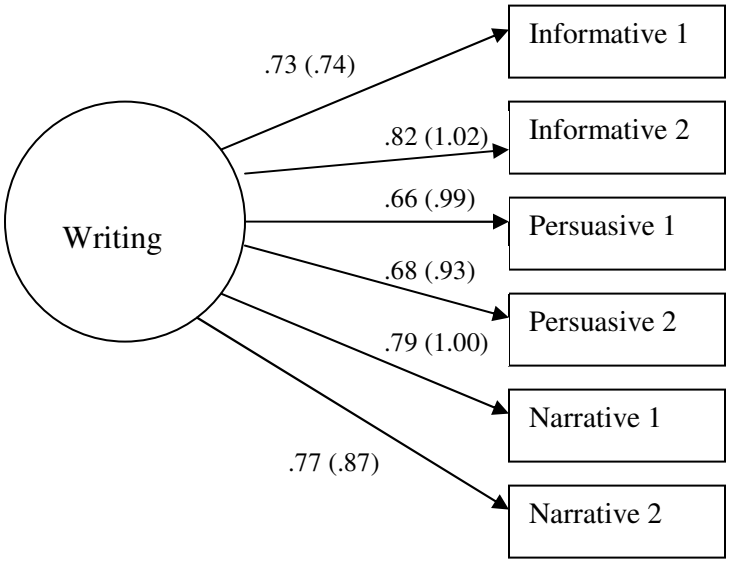

Figure C.1. Standardized and unstandardized estimates for grade 4, one-factor model.

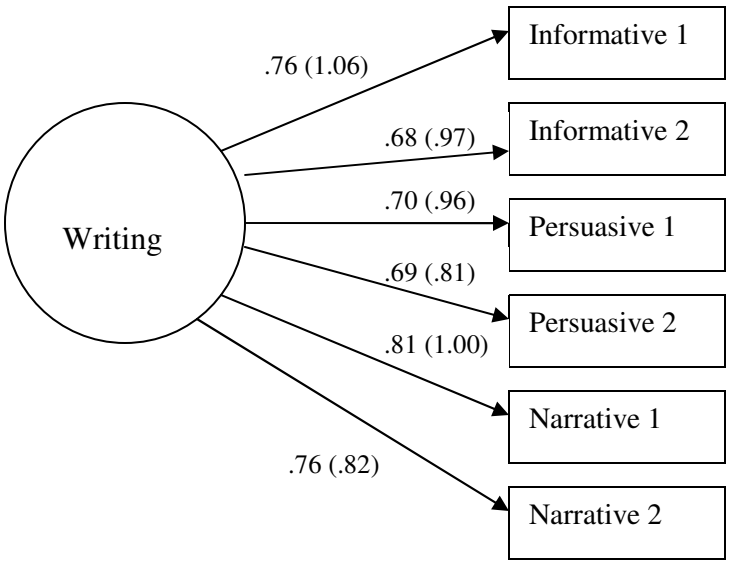

Figure C.2. Standardized and unstandardized estimates for grade 6, one-factor model.

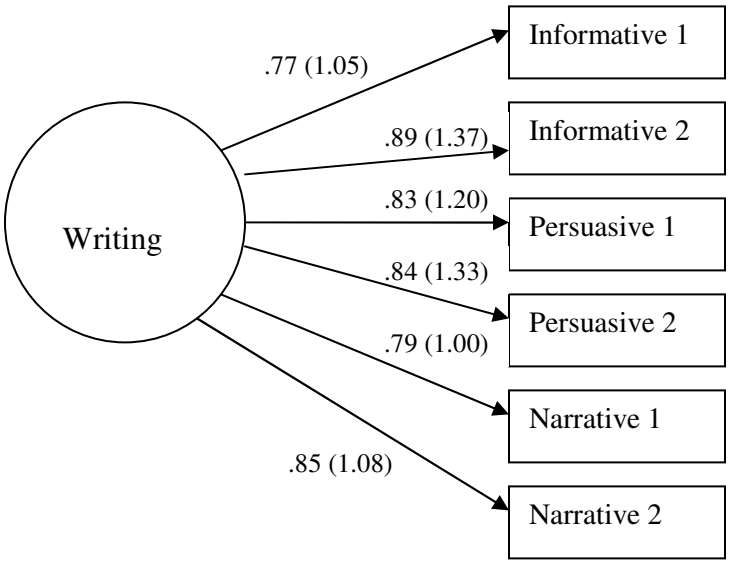

Figure C.3. Standardized and unstandardized estimates for grade 8, one-factor model.

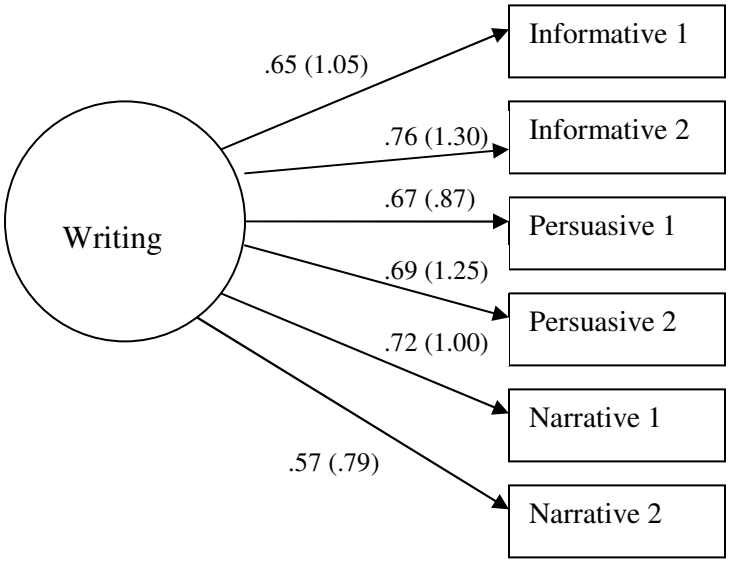

Figure C.4. Standardized and unstandardized estimates for grade 10, one-factor model.

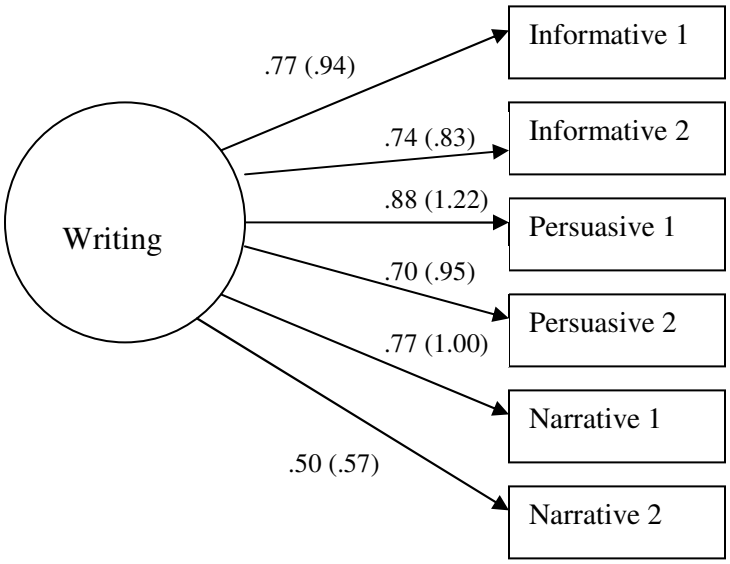

Figure C.5. Standardized and unstandardized estimates for grade12, one-factor model.

APPENDIX D

Correlations, Means, and Standard Deviations for Indicator Variables by Grade

Subscale	Reading 1	Reading 2	Informative	Narrative	Persuasive
			Grade 4		
1. Reading 1	—	.84	.47	.60	.42
2. Reading 2		—	.43	.55	.37
3. Informational			—	.77	.70
4. Narrative				—	.68
5. Persuasive					—
Mean (SD)	1.24 (1.38)	1.19 (1.20)	0.44 (1.72)	0.49 (1.83)	0.06 (2.08)
			Grade 6		
1. Reading 1	—	.83	.44	.35	.32
2. Reading 2		—	.49	.45	.43
3. Informational			—	.73	.74
4. Narrative				—	.71
5. Persuasive					—
Mean (SD)	1.49 (1.36)	1.51 (1.24)	0.55 (1.71)	0.87 (1.47)	-0.36 (1.53)
			Grade 8		
1. Reading 1	—	.80	.65	.51	.63
2. Reading 2		—	.68	.56	.69
3. Informational			—	.80	.86
4. Narrative				—	.77
5. Persuasive					—

Subscale	Reading 1	Reading 2	Informative	Narrative	Persuasive
Mean (SD)	0.85 (1.31)	1.47 (1.53)	0.33 (2.08)	0.64 (1.85)	-0.29 (2.21)
			Grade 10		
1. Reading 1	—	.69	.56	.53	.48
2. Reading 2		—	.51	.60	.48
3. Informational			—	.51	.62
4. Narrative				—	.66
5. Persuasive					—
Mean (SD)	1.25 (1.31)	1.72 (1.51)	1.62 (1.82)	1.45 (1.47)	0.73 (1.63)
			Grade 12		
1. Reading 1	—	.65	.52	.57	.56
2. Reading 2		—	.35	.42	.46
3. Informational			—	.62	.77
4. Narrative				—	.63
5. Persuasive					—
Mean (SD)	1.68 (1.30)	2.34 (1.28)	0.31 (1.93)	1.34 (1.99)	0.73 (2.31)

Note. All correlations are significant at the $p < .01$ level.

APPENDIX E

Standardized and Unstandardized Coefficients for the Reading-Writing Model

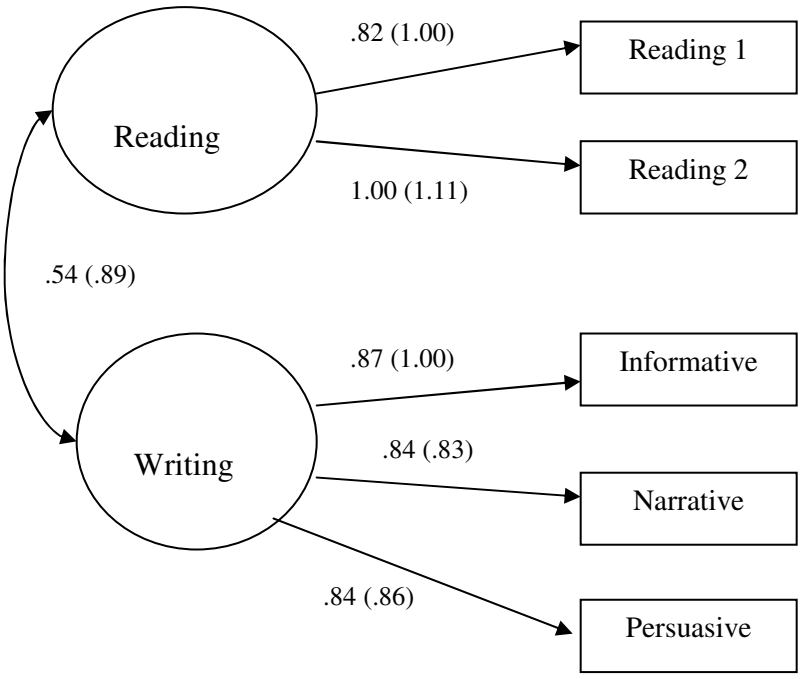

Figure E.1. Grade 4 standardized and unstandardized coefficients.

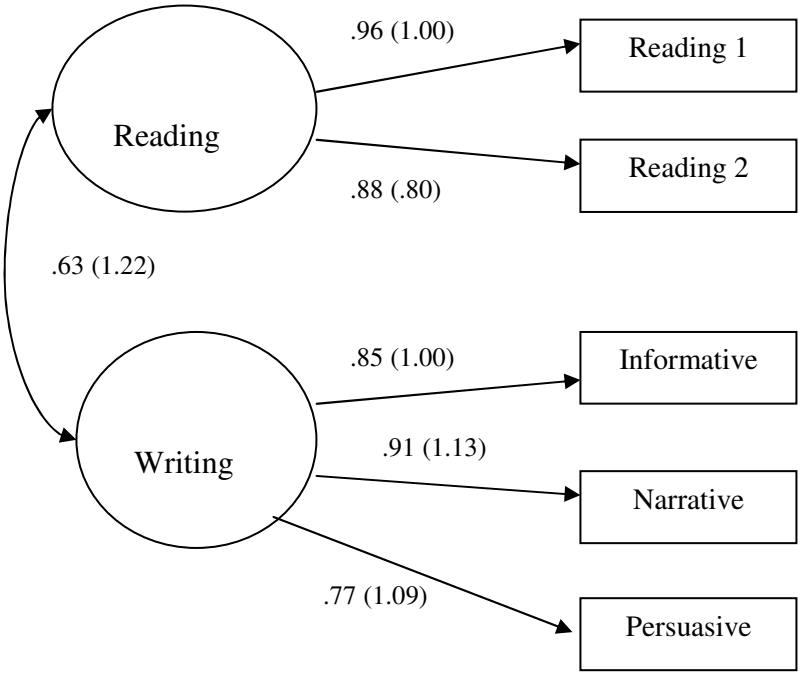

Figure E.2. Grade 6 standardized and unstandardized coefficients.

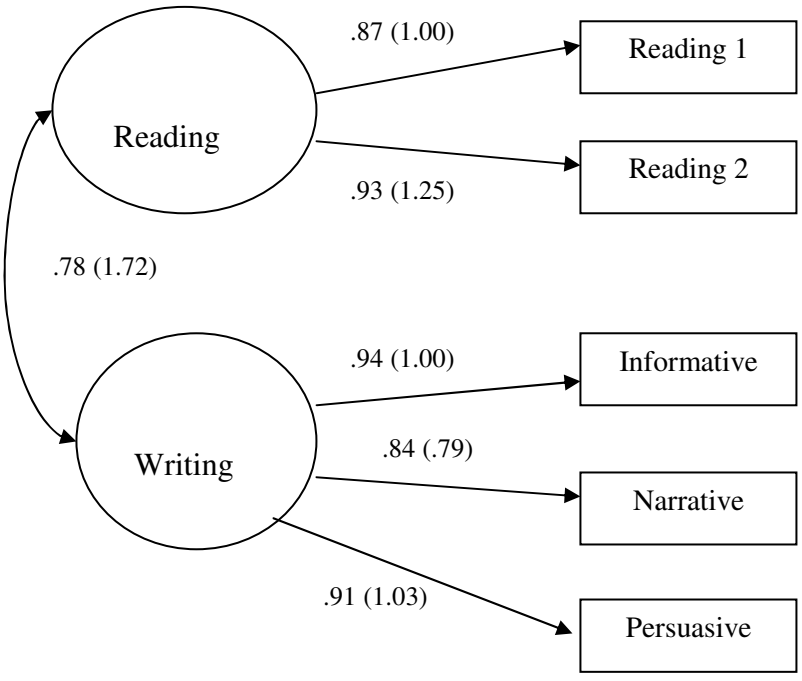

Figure E. 3. Grade 8 standardized and unstandardized coefficients.

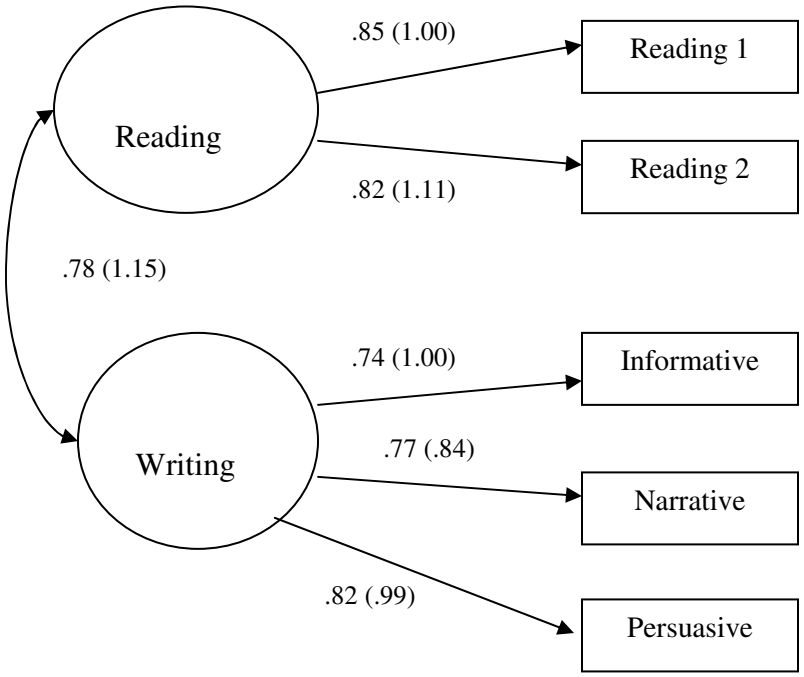

Figure E.4. Grade 10 standardized and unstandardized coefficients.

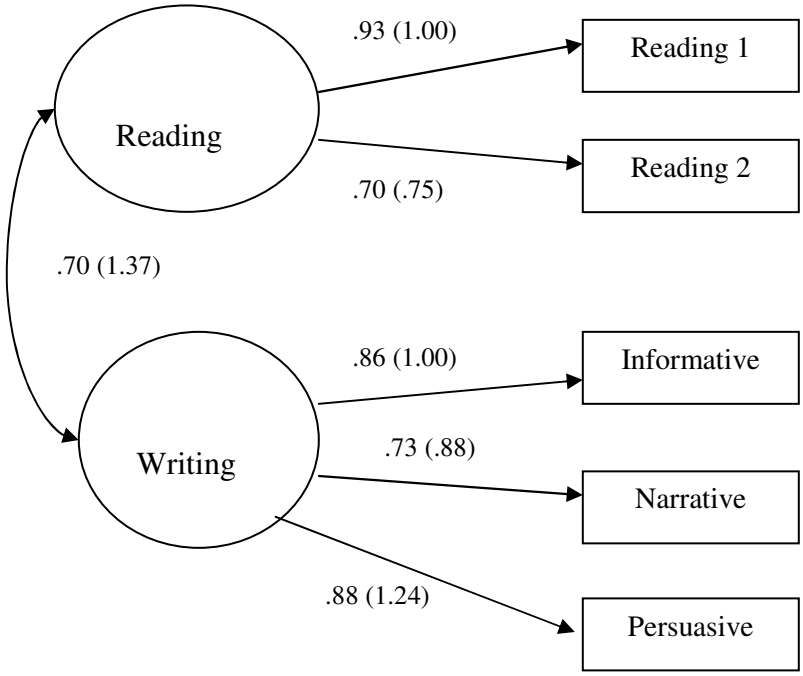

Figure E.5. Grade 12 standardized and unstandardized coefficients.

APPENDIX F

Correlations, Means, and Standard Deviations for Indicator Variables by Gender and Race

Subscale	Reading 1	Reading 2	Informative	Narrative	Persuasive
Male (n = 186)					
1. Reading 1	—	.73	.49	.46	.49
2. Reading 2		—	.44	.47	.51
3. Informational			—	.64	.75
4. Narrative				—	.74
5. Persuasive					—
Mean (SD)	1.30 (1.43)	1.74 (1.53)	0.31 (2.12)	0.87 (1.80)	-0.08 (2.13)
Female (n = 220)					
1. Reading 1	—	.76	.53	.52	.49
2. Reading 2		—	.56	.52	.58
3. Informational			—	.71	.72
4. Narrative				—	.66
5. Persuasive					—
Mean (SD)	1.33 (1.29)	1.73 (1.33)	0.96 (1.73)	1.18 (1.66)	0.32 (1.86)

Subscale	Reading 1	Reading 2	Informative	Narrative	Persuasive
Blacks (n = 165)					
1. Reading 1	—	.71	.40	.42	.33
2. Reading 2		—	.30	.35	.39
3. Informational			—	.62	.68
4. Narrative				—	.59
5. Persuasive					—
Mean (SD)	0.59 (1.00)	1.03 (1.05)	0.14 (1.68)	0.59 (1.60)	-0.49 (1.71)
White (n = 225)					
1. Reading 1	—	.66	.46	.43	.44
2. Reading 2		—	.49	.48	.51
3. Informational			—	.66	.74
4. Narrative				—	.72
5. Persuasive					—
Mean (SD)	1.89 (1.34)	2.30 (1.42)	1.09 (2.01)	1.43 (1.71)	0.66 (2.03)

APPENDIX G

Standardized and Unstandardized Coefficients for Gender and Race

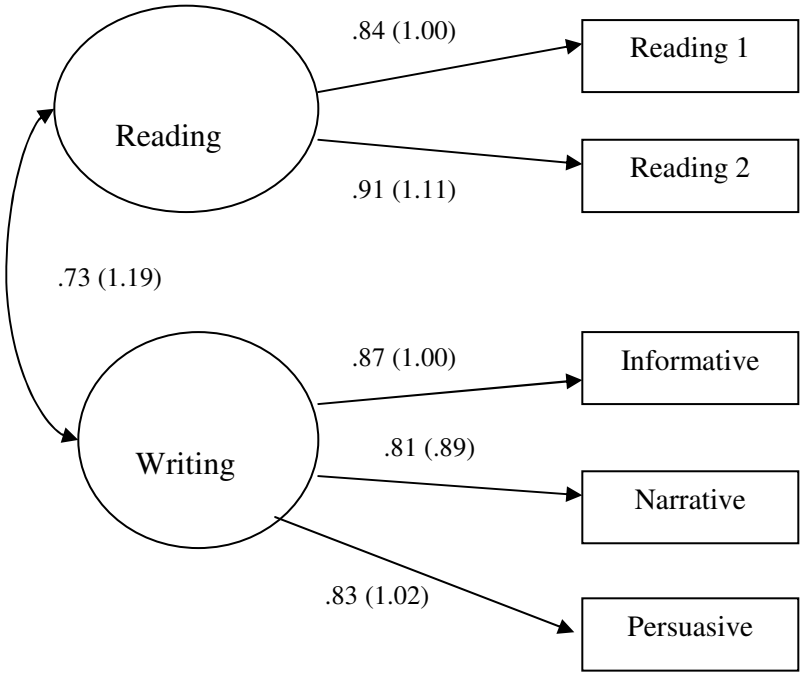

Figure G.1. Standardized and unstandardized coefficients for males.

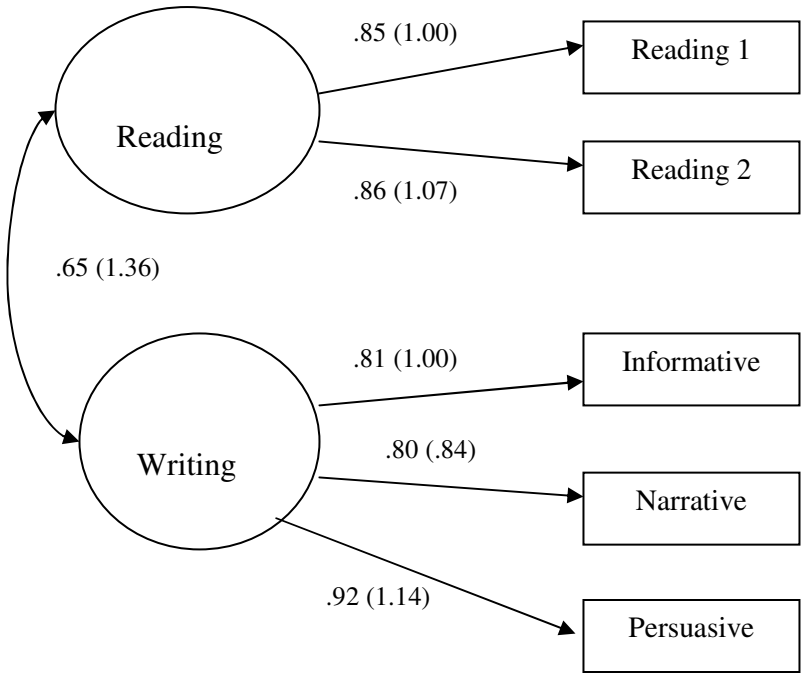

Figure G.2. Standardized and unstandardized coefficients for females.

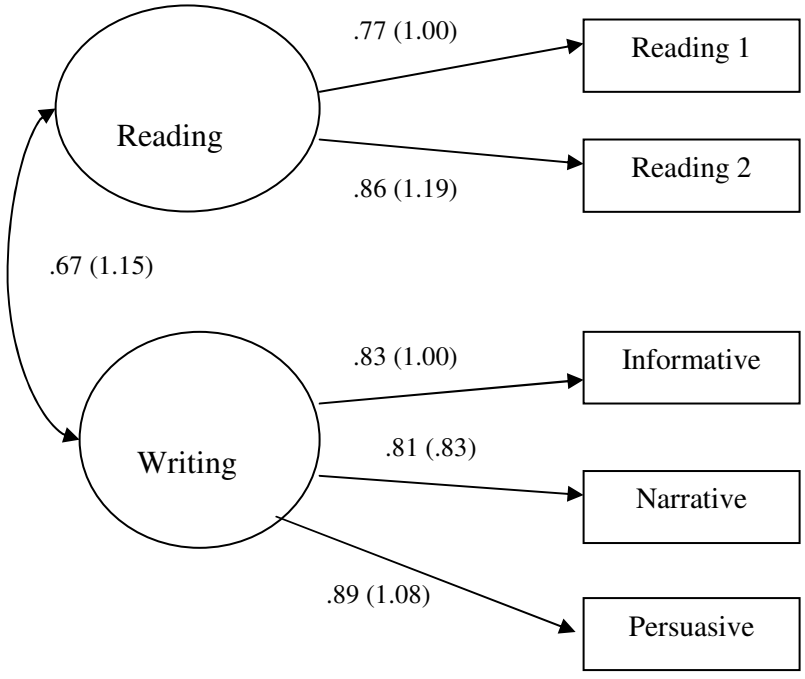

Figure G.3. Standardized and unstandardized coefficients for Blacks.

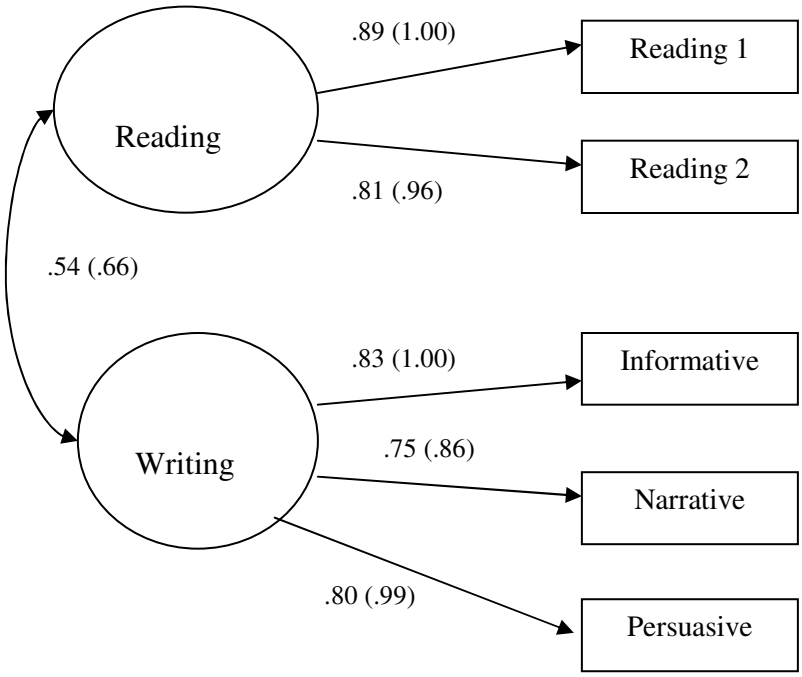

Figure G.4. Standardized and unstandardized coefficients for Whites.

REFERENCES

Abbott, R., & Berninger, V. W. (1993). Structural equation modeling of relationships among developmental skills and writing skills in primary- and intermediate-grade writers. *Journal of Educational Psychology, 85*, 478-508.

Anderson, J. C., & Gerbing, D. (1984). The effect of sampling error on convergence, improper solutions, and goodness-of-fit indices for maximum likelihood confirmatory factor analysis. *Psychometrika*, *49*, 155-173.

Anderson, R. C., & Pearson, P. D. (1984). A schema-theoretic view of basic processes in reading. In P. D. Pearon, R. Barr, M. L. Kamil, & P. Mosenthal (Eds.), *Handbook of reading research* (pp. 255-291). New York: Longman.

Andrich, D. (1988). *Rasch models for measurement*. Newbury Park, CA: Sage.

Arbuckle, J. L. (1999). *Amos 4.0 user's guide*. Chicago, IL: Smallwaters.

Arbuckle, J. L. (1995-2006). *Amos (Version 7.0).* [Computer software]. Spring Hill, PA. Amos Development Corporation.

Bangert-Drowns, R. L., Hurley, M. M., & Wilkinson, B. (2004). The effects of school-based writing-to-learn interventions on academic achievement: A meta-analysis. *Review of Educational Research, 74*, 29-58.

Berninger, V. W. (2000). Development of language by hand and its connections to language by ear, mouth, and eye. *Topics in Language Disorders, 20*, 65-84.

Berninger, V. W., Abbott, R. D., Abbott, S. P., Graham, S., & Richards, T. (2002). Writing and reading: Connections between language by hand and language by eye. *Journal of Learning Disabilities, 35*, 39-56.

Bond, T. G. (2007). Applying the Rasch model: Fundamental measurement in the human sciences (2nd ed.). Mahwah, NJ: Lawrence Erlbaum.

Boomsma, A. (1983). *On the robustness of LISREL (Maximum Likelihood Estimation) against small sample size and nonnormality*. Amsterdam: Sociometric Research Foundation.

Boomsma, A. (1985). Noncovergence, improper solutions, and starting values in LISREL maximum likelihood estimation. *Psychometrika,* 50, 229-42.

Bond, G. L., & Dykstra, R. (1967). The cooperative research program in first-grade reading instruction. *Reading Research Quarterly*, *2,* 5-142.

Brandt, D. (1994). Remembering writing, remembering reading. *College Composition and Communication, 45*, 459-479.

Brennan, R. (2001). *Generalizability theory*. New York: Springer-Verlag.

Carrell, P. L., & Connor, U. (1991). Reading and writing descriptive and persuasive texts. The *Modern Language Journal, 75*, 314-324.

Chen, F., Bollen, K. A., Paxton, P., Curran, P. J., & Kirby, J. (2001). Improper solutions in structural equation models: Causes, consequences, and strategies. *Sociological Methods and Research*, *29*, 468-508.

Crick, J. E., & Brennan, R. L. (1983). *Manual for GENOVA: A generalized analysis of variance system* (American College Testing Technical Bulletin No. 43). Iowa City, IA: ACT.

Crowhurst, M. (1991). Interrelationships between reading and writing persuasive discourse. *Research in the Teaching of English, 25*, 315-338.

Cunningham, A. E., & Stanovich, K. E. (1991). Tracking the unique effects of print exposure in children: Associations with vocabulary, general knowledge and spelling. *Journal of Educational Psychology, 83*, 264-274.

DeVries, T. (1970). Reading, writing frequency, and expository writing. *Reading Improvement, 7*, 14-15.

Digest of Education Statistics. (2005). National Center for Education Statistics. Retrieved November 15, 2007, from http://nces.ed.gov/programs/digest/d07/tables/dt07_040.asp

Duke, N. K., Bennett-Armistead, S., & Roberts, E. M. (2003). Filling the great void: Why we should bring nonfiction into the early-grade classroom. *American Educator, 27*, 30-34, 46.

Duke, N., & Pressley, M. (2002). Effective practices for developing reading comprehension. In A. J. Farstrup & S. J. Samuels (Eds.), *What research has to say about reading instruction* (pp. 205-242). Newark, DE: International Reading Association.

Durkin, D. (1988). *Teaching them to read* (5th ed.). Boston: Allyn & Bacon.

Dynes, J. J. (1932). Comparison of two methods of studying history. *Journal of Experimental Education, 1*, 42-45.

Eckhoff, B. (1984). How reading affects children's writing. In J. Jensen (Ed.), *Composing and comprehending* (pp. 105-114). Urbana, IL: National Conference on Research in English. ERIC Clearinghouse on Reading and Communication Skills.

Elley, W. B., Barham, I. H., Lamb, H., & Wyllie, M. (1976). The role of grammar in a secondary school English curriculum. *Research in the Teaching of English, 10*, 5-21.

Emig, J. (1977). Connecting reading and writing: College learners' metacognitive awareness. *Journal of Developmental Education, 21*, 10-19.

Engelhard, G. (1992). The measurement of writing ability with a many-faceted Rasch model. *Applied Measurement in Education, 5,* 171-191.

Engelhard, G. (1994). Examining rater errors in the assessment of written composition with a many-faceted Rasch model. *Journal of Educational Measurement, 31*(2), 93-112.

Engelhard, G. (2000). Monitoring raters in performance assessments. In G. Tindal & T. Haladyna (Eds.), *Large-scale assessment programs for ALL students: Development, implementation and analysis* (pp. 261-287). Mahwah, NJ: Lawrence Erlbaum.

Engelhard, G. (2001). Historical view of the influences of measurement and reading theories on the assessment of reading. *Journal of Applied Measurement, 2*, 1-26.

Engelhard, G., Walker, E. S., Gordon, B., & Gabrielson, S. (1994). Writing tasks and gender: Influences on writing quality of black and white students. *Journal of Educational Research, 87*, 197-209.

Fitzgerald, J. (1990). Reading and writing as "mind meeting." In T. Shanahan (Ed.), *Reading and writing together: New perspectives for the classroom* (pp. 81-97). Norwood, MA: Christopher-Gordon.

Fitzgerald, J., & Shanahan, T. (2000). Reading and writing relations and their development. *Educational Psychologist, 35,* 39-50.

Fitzgerald, J., & Teasley, A. B. (1986). Effects of instruction in narrative structure on children's writing. *Journal of Educational Psychology, 78*, 424-432.

Gardner, E. F., Rudman, H. C., Carlsen, B., & Merwin, J. C. (1985). *The Stanford achievement test* (8th ed.) San Antonio, TX: The Psychological Corporation.

Graham, S., & Perin, D. (2007). *Writing next: Effective strategies to improve writing of adolescents in middle and high schools*. A report to Carnegie Corporation of New York. New York: Alliance for Excellent Education.

Gyadenda, I. S., & Engelhard, G. (1998, April). *Applying the Rasch model to explore rater influences on the assessed quality of students' writing ability*. Paper presented at the Annual Meeting of the American Educational Research Association, San Diego, CA.

Hambelton, R. K., Swaminathan, H., & Rogers, H. J. (1991). *Fundamentals of item response theory*. Newbury Park, CA: Sage.

Hart, B., & Risley, T. R. (1992). American parenting of language-learning children: Persisting differences in family-child interactions observed in natural home environments. *Developmental Psychology*, *28*, 1096-1105.

Hartman, D. K. (2004). Deconstructing the reader, the text, the author, and the context: Intertextuality and reading from a "cognitive " perspective. In N. Shuart-Faris & D. Bloome (Eds.), *Uses of intertextuality in classroom and educational research* (pp. 353-372). Greenwich, CT: Information Age.

Heck, R. H., & Crislip, M. (2001). Direct and indirect writing assessments: Examining issues of equity and utility. *Educational Evaluation and Policy Analysis,* 23, 275-291.

Heller, M. F. (Ed.). (1995). *Reading-writing connections: From theory to practice*. White Plains, NY: Longman.

Hiebert, E. H., Englert, C. S., & Brennan, S. (1983). Awareness of text structure in recognition and production of expository discourse. *Journal of Reading Behavior, 15*(4), 63-79.

Hiebert, E. H., Pearson, P. D., & Taylor, B. M. (1998). *Every child a reader: Applying reading research in the classroom*. Center for the Improvement of Early Reading Achievement. Ann Arbor: University of Michigan.

Hollenbeck, K., Tindal, G., & Almond, P. (1999). Reliability and decision consistency: An analysis of writing mode at two times on a statewide test. *Educational Assessment, 6*, 23-40.

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, *6*, 1-31.

Hunt, K. W. (1965). *Grammatical structures written at three grade levels*. Champaign, IL: National Council of Teachers of English.

Irwin, J. W., & Doyle, M. A. (Eds.). (1992). *Reading/writing connections: Learning from research*. Newark, NJ: International Reading Association.

Jenkins, J. R., Johnson, E., & Hileman, J. (2004). When reading is also writing: Sources of individual differences on the new reading performance tests. *Scientific Studies of Reading, 8*, 125-151.

Juel, C. (1988). Learning to read and write: A longitudinal study of 54 children from first through fourth grades. *Journal of Educational Psychology, 80*, 437-447.

Kamberelis, G. (1999). Genre development and learning: Children writing stories, science reports, and poems. *Research in the Teaching of English, 33*, 403-459.

Kamberelis, G., & Bovino, T. D. (1999). Cultural artifacts as scaffolds for genre development. *Reading Research Quarterly, 34*, 138-170.

Kegley, P. (1986). The effect of mode discourse on student writing performance: Implications for policy. *Educational Evaluation and Policy Analysis, 8*, 147-154.

Kellogg, R. T. (1994). *The psychology of writing*. New York: Oxford University Press.

Kline, R. B. (1998). *Principles and practice of structural equation modeling*. New York: Guilford.

Kuhlemeier, H., van den Bergh, H., & Wijnstra, J. (1995, April). *Multilevel factor analysis applied to national assessment data*. Paper presented at the annual meeting of the American Educational Research Association, San Francisco.

Korat, O., & Schiff, R. (2005). Do children who read more books know "What is good writing" better than children who read less? A comparison between grade levels and SES groups. *Journal of Literacy Research*, *37*(3), 289-334.

Langer, J. A. (1986). *Children reading and writing: Structures and strategies*. Norwood, NJ: Ablex.

Linacre, J. M. (1988) *FACETS: Computer program for many-faceted Rasch model*. Chicago: MESA Press.

Linacre, J. M. (2007). *A user's guide to Winsteps, Ministep Rasch-model computer programs*. Chicago: John M. Linacre.

Loban, W. D. (1963). *The language of elementary school children*. Urbana, IL: National Council of Teachers of English.

Loban, W. D. (1967). *Language ability—grades 10, 11, and 12: Final report*. Urbana, IL: National Council of Teachers of English.

Mehta, P. D., Foorman, B. R., Branum-Martin, L, & Taylor, W. P. (2005). Literacy as a unidimensional multilevel construct: Validation, sources of influence, and implications in a longitudinal study in grades 1 to 4. *Scientific Studies of Reading,* 9. 85-116.

Mills, E. B. (1974). Children's literature and teaching written composition. *Elementary English, 51*, 971-973.

Mississippi Department of Education. (2007). Mississippi Assessment and Accountability Reporting System. Retrieved March 15, 2008 from: http://orsap.mde.k12.ms.us:8080/MAARS/index.jsp?userSessionId=191

Moffett, J., & Wagner, B. J. (1983). *Student-centered language arts and reading, K-13* (3rd ed.). Boston: Houghton Mifflin.

National Center for Education Statistics. (2003). *The Nation's Report Card: Writing Highlights 2002,* NCES 2003-531, Washington, DC: National Center for Education Statistics.

National Assessment of Educational Progress. (2005). National Assessment of Educational Progress: Reading 2005. Retrieved November 5, 2007, from The Nation's Report Card: http://nationsreportcard.gov/reading_math_grade12_2005/s0201.asp

National Assessment of Educational Progress. (2007). National Assessment of Educational Progress: Reading 2007. Retrieved November 5, 2007, from The Nation's Report Card: http://nationsreportcard.gov/reading_2007/r0002.asp

Nelson, N., & Calfee, R. C. (Eds.). (1998). *The reading-writing connection. 97th yearbook of the National Society for the Study of Education*. Chicago: National Society for the Study of Education.

Newell, G. E. (1984). Learning from writing in two content areas: A case study/protocol analysis. *Research in the Teaching of English*, 18, 265-285.

No Child Left Behind Act of 2001 (NCLB). (2002). P.L. 107-110.

Perin, D. (1998). Assessing the reading-writing relationship in adult literacy students. *Reading Psychology: An International Quarterly, 19*, 141-183.

Pressley, M. (2000). What should comprehension instruction be the instruction of? In M. L. Kamil, P. B. Mosenthal, P. D. Pearson, & R. Barr (Eds.) *Handbook of reading research* (Vol. 3, pp. 545-562). Mahwah, NJ: Lawrence Erlbaum.

Quellmalz, E., Capell, F. & Chou, C. (1982). Effects of discourse and response mode on the measurement of writing competence. *Journal of Educational Measurement, 19*, 241-258.

Rasch, G. (1980). *Probabilistic models for some intelligence and attainment tests*. Chicago: University of Chicago Press.

Rosenblatt, L. M. (1978). *The reader, the text, the poem: The transactional theory of the literary work*. Carbondale, IL: Southern Illinois University Press.

Ryan, S. M. (1985). An examination of reading and writing strategies of selected fifth grade students. In J. Niles & R. Lalik (Eds.) *Issues in literacy: A research perspective* (pp. 386-390). Rochester, NY: National Reading Conference.

Sachse, P. (1984). Writing assessment in Texas: Practices and problems. *Educational Measurement: Issues and Practice, 3*, 21-23.

Scholastic, Inc. (1999). *Scholastic Reading Inventory*. New York: Scholastic.

Scholastic, Inc. (2003). *Scholastic Reading Inventory educator's guide*. New York: Scholastic.

Scholastic, Inc. (2007). *Scholastic Reading Inventory technical guide: Working with student performance data*. New York: Scholastic.

Shanahan, T. (1984). Nature of the reading-writing relation: An exploratory multivariate analysis, *Journal of Educational Psychology, 76,* 466-477.

Shanahan, T., Ed. (1990). *Reading and writing together: New perspectives for the classroom*. Norwood, MA: Christopher-Gordon.

Shanahan, T. (2006). Relations among oral language, reading, and writing development. In MacArthur, C. A., Graham, S. & Fitzgerald, J. (Eds.). *Handbook of writing research* (pp. 171-183). New York: Guilford.

Shanahan, T., & Lomax, R. G. (1986). An analysis and comparison of theoretical models of the reading-writing relationship. *Journal of Educational Psychology, 78*, 116-123.

Shanahan, T., & Lomax, R. G. (1988). A developmental comparison of three theoretical models of reading-writing relationship. *Research in the Teaching of English, 22*, 196-212.

Shavelson, R. J., & Webb, N. M. (1991). *Generalizability theory: A primer*. Newbury Park, ca: Sage.

Spearman, C. (1904). The proof and measurement of association between two things. *American Journal of Psychology, 15*, 72-101.

Squire, J. R. (1984). Composing and comprehending: Two sides of the same basic process. In J. Jensen (Ed.), *Composing and comprehending* (pp. 22-32). Urbana, IL: National Conference on Research in English.

Stahl, S. A., Pagnucco, J. R., & Sutters, C. W. (1996). First graders reading and writing instruction in traditional and process-oriented classes. *Journal of Educational Research, 89*, 131-144.

Stanford Achievement Test. (1983). *Stanford writing assessment program guide*. San Antonio, TX: Harcourt Brace Jovanovich.

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. (2000). *Progress in understanding reading*. New York: Guilford.

Stauffer, R. G. (1980). The language-experience approach to the teaching of reading (2nd ed.). New York: Harper & Row.

Stotsky, S. (1983). Research on reading/writing relationships: A synthesis and suggested directions. *Language Arts, 60*, 568-580.

Thissen, D., Nelson, L., Rosa, K., & McLeod, L. D. (2001). Item response theory for items scored in more than two categories (pp. 141-186). In H. Wainer & D. Thissen (Eds.), *Test scoring*. Mahwah, NJ: Lawrence Erlbaum.

Thompson, R., Jones, L., Haynes, J., & Rucker, S. (2000). *Schools and parents partnering for student success: A parent's guide to the Mississippi curriculum framework for reading, language arts, and mathematics*. Jackson: Mississippi Department of Education.

Tierney, R. J., & Pearson, D. (1984). Toward a composing model of reading. In J. M. Jensen (Ed.), *Composing and comprehending* (pp. 33-45). Urbana, IL: National Conference on Research in English. ERIC Clearinghouse on Reading and Communication Skills.

Tierney, R. J., & Shanahan, T. (1996). Research on the reading-writing relationship: Interactions, transactions, and outcomes. In R. Barr, M. L. Kamil, P. Mosenthal, & P. D. Pearson (Eds.), *Handbook of reading research* (Vol. 2, pp. 246-280). Mahwah, NJ: Lawrence Erlbaum.

Tierney, R. J., Soter, A., O'Flahavan, J. F., & McGinley, W. (1989). The effects of reading and writing upon thinking critically. *Reading Research Quarterly, 24*(2), 134-173.

Troia, G. A., & Graham, S. (2002). The effectiveness of a highly explicit, teacher-directed strategy instruction routine: Changing the writing performance of students with disabilities. *Journal of Learning Disabilities, 35*, 290-305.

Van Driel, O. P. (1978). On various causes of improper solutions in maximum likelihood factor analysis. *Psychometrika*, *43*, 225-43.

Wittrock, M. C. (1984). Writing and the teaching of reading. In J. M. Jensen (Ed.), *Composing and comprehending*. (pp. 77-83). Urbana, IL: National Conference on Research in English. ERIC Clearinghouse on Reading and Communication Skills.

Woodcock, R. (1987). *Woodcock reading mastery test-revised*. Circle Pines, MN: American Guidance Service.

Woodfin, M. J. (1968). Correlations among certain factors and the written expression of third grade children. *Educational and Psychological Measurement*, 28, 1237-1242.

Wright, B. D., & Masters, G. N. (1982). *Rating scale analysis.* Chicago: MESA Press.