ABSTRACT

Title of Document: THE EFFECTS OF PROMPT CONDITION
                            AND GENRE ON THE WRITING
                            PERFORMANCE OF STUDENTS IN 3RD AND
                            5TH GRADE.

                            Kelly Marie Worland, Doctor of Philosophy,
                            2014

Directed By: Associate Professor Susan De La Paz
                            Department of Counseling, Higher Education
                            and Special Education

Research on writing at the elementary level has shown that prior knowledge and task environment can have a significant effect on writing performance. In addition, there is preliminary evidence that suggests that children's development in writing may vary by genre favoring the narrative over expository genres (e.g., Olinghouse & Wilson, 2011; and Camp, 1993). One way to mitigate the effects of prior knowledge on topic and genre is through varying prompting conditions. The new Common Core writing assessment context requires students to write in response to texts read. To date, however, there are no studies that simultaneously examine the effects of genre and prompt condition in the elementary grades.

This study examines the effect of two prompting conditions (supported and unsupported) on students' writing performance in multiple genres (narrative, persuasive and informational report) in order to assess the potential impact of the read aloud accommodation on these new types of writing assessments along with the effect of genre

simultaneously. Findings show that at the 3rd grade level, students write best in the informational report genre over the narrative and persuasive genre, and that the read aloud accommodation positively affects writing quality. At the 5th grade level, the read aloud accommodation does not have a significant effect on writing quality.

Based on the findings above, there are a number of implications for current testing policy and instruction. First, features of the prompt condition such as providing a common text, and audience and genre cues resulted in higher style, organization, conventions and mechanic scores for the informational report genre in the younger grades. This contradicts earlier findings that suggest at the elementary level, the narrative genre is more accessible. Accordingly, an effort should be made by test-makers and educators to provide students with these beneficial supports when designing tests and assignments particularly for expository genres. Second, given the observed benefit of the read aloud accommodation for students in the younger grades, test makers should consider designing tests that vary the degree of supports students are provided on the read and write response tasks as they progress through the grade levels.

THE EFFECTS OF PROMPT CONDITION AND GENRE ON THE WRITING

PERFORMANCE OF STUDENTS IN 3RD AND 5TH GRADE

By

Kelly Marie Worland

Dissertation submitted to the Faculty of the Graduate School of the

University of Maryland, College Park, in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

2014

Advisory Committee:

Dr. Susan De La Paz, Chair

Dr. Leslie V. Page-Voth

Dr. Rebecca Silverman

Dr. Wayne Slater

Dr. Jade Wexler

Dedicated to my parents, John and Mamiya Worland, who were my first, and continue to be my best teachers

And to my husband, Brian, my sturdy oak

ACKNOWLEDGEMENTS

I would like to thank the many people who supported me as I worked on my dissertation. First, I would like to thank the four teachers who welcomed me into their classrooms during the 2012-13 school year. All four were flexible, gracious and wonderful to work with. Thank you for your contribution to educational research and to the profession of teaching.

I am also deeply appreciative of the time, energy and support that the members of my committee from the University of Maryland provided to me as a graduate student, and throughout the dissertation process. Dr. Rebecca Silverman, I appreciated the opportunities you provided me to participate in some of your research projects and publications. Dr. Jade Wexler, thank you for sharing your insights into research design across a multitude of settings and topics. Dr. Tori Page-Voth, thank you for sharing your insights not only into my research, but also into balancing the life of an academic with that of family. Dr. Wayne Slater, thank you for helping to deepen my understanding of genre study and writing research. Finally, thank you to Dr. Susan De La Paz, my advisor. Your support and guidance throughout my time at the University of Maryland was invaluable. I could not have asked for a better or more supportive advisor to help me through this process.

Lastly, I would like to thank my parents and my husband. Thank you to all of you for your unconditional love, and support of all of my personal and academic endeavors.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1: INTRODUCTION

Problem Statement

Writing is a fundamental skill and an important part of school and post-secondary life. Failure to acquire competent writing skills has the potential to limit an individual's opportunities for future success in both education and employment (Graham, 2006). Given the recognized importance of writing skill development, current testing outcome data is worrisome.

According to the 2002 National Assessment of Education Progress (NAEP) 72% of students were performing below the proficient level in the 4th-grade. Furthermore, children with disabilities and English language learners performed on average 22% and 18% below their general education peers, respectively. In essence, many students, particularly students with disabilities (LD), and who are English language learners (EL) still struggle to meet basic levels of writing competency and as such are at-risk for school failure.

One way to improve students' writing ability may be through assessment. Writing assessments provide teachers and stakeholders with valuable information on students' strengths and weaknesses to help individualize instruction to meet student needs (Graham & Hebert, 2010). Writing assessments, however, can come in many forms. Standardized writing assessments are typically developed by districts, states, and other constituent groups, and are used for a number of reasons (e.g., to identify a student as having particularly weak writing skills, or for accountability purposes). It is possible to examine many of these standardized tests, and to then determine what types of writing are currently valued by educators and policy makers. One recent movement of relevance

includes the Common Core Curriculum (www.corestandards.org), which has been

adopted by 45 states and 3 United States territories. The Common Core Curriculum

includes English language arts learning standards specifically targeted towards narrative,

persuasive, and informational report writing.

Twenty-two states and the District of Columbia will begin participating in the

Partnership for Assessment of Readiness for College and Careers (PARCC;

www.parcconline.org) to assess the Common Core curriculum. The PARCC consortium

has reported that the grades 3-8 tests will include performance-based writing assessments

that include both a research simulation task and one focused on analyzing literature.

These tasks will involve asking students to read multiple texts and write several pieces to

"demonstrate the ability to read and comprehend a range of sufficiently complex texts

independently, to write effectively when using and analyzing sources, and to build and

present knowledge through integration, comparison, and synthesis of ideas."

(http://www.parcconline.org). Another assessment consortium is the Smarter Balanced

(SBAC) group (http://www.smarterbalanced.org). Twenty-one states have agreed to

participate in SBAC, which will have similar writing assessment formats to the PARCC.

There are a number of advantages to the newest form of standardized assessment,

proposed by PARCC and SBAC over previous models. Until recently, one of the most

commonly used forms of standardized state and district wide writing tests were one-time

assessments in a single genre once or twice over the course of a child's entire primary

and secondary education (Murphy, 2008). Within the elementary school context, the

genres most commonly assessed include narrative, persuasive and informational reports

(http://nces.ed.gov/nationsreportcard/writing/).

Prior assessment models have been criticized for three primary reasons. First, previous standardized writing assessments often relied on single samples of one form of writing (e.g., narrative or persuasive). Research has shown that one form of writing has not been found to be consistently representative of how a student might perform on all forms of writing (Olinghouse & Wilson, 2012; Hebert, Graham & Harris, 2011). Second, writing skills are not the same depending upon the genre (Hebert, Graham & Harris, 2011; Olinghouse & Wilson, 2012; and Camp, 1993). In response to both sets of findings, researchers suggested that assessment in a single genre may not be appropriate for the identification of students with writing difficulties (Hebert, et al., 2011, and Olinghouse & Wilson, 2012). A third criticism is that the infrequency of these standardized assessments is likely to be insufficient in providing teachers and school systems enough information on student writing progress. As such, the PARCC and SBAC's proposed assessments seem to offer a good alternative to the previous state and district wide assessments by: (a) assessing writing in multiple genres, (b) assessing writing in an applied context, and (c) assessing writing at multiple grade levels.

Unfortunately, there may be problems even with the proposed PARCC and SBAC writing assessments. To illustrate, the new form of assessment requires students to read a topic, and then write a response to what they read. The problem with the new form of writing assessment is that it may conflate reading ability and prior knowledge with writing ability. In addition, there is research to suggest that an assessment's task environment (e.g., prompting condition and demands), students' prior knowledge, and the genre of the assessment all can have the potential to impact a student's writing performance (e.g., Huot, 1999; Myhill, 2005; Olinghouse & Graham, 2009; and Scott &

Windsor, 2000). Less is known, however, on the relationship between how these components of writing assessment vary given individual student differences.

On November 12, 2013 the PARCC consortium adopted the PARCC Accessibility Features and accommodations manual for students with disabilities (including students with LD and students who are hard of hearing, blind or have other cognitive or physical disabilities), students who are EL, and students with 504 plans (http://www.parcconline.org/parcc-releases-accessibility-features-and-accommodations-manual). The accommodations manual provides a description of three layers of supports offered to students. The first two supports are available to all students and include embedded supports and accessibility features. These include accommodations such as directions read aloud and repeated, blank paper, highlighting tools, spell checker, writing tools (cut, paste, copy, underline) on the computer-based assessment format, and text-to-speech read aloud of all content on the mathematics assessment. The third layer of support includes testing accommodations that must be determined specifically for students with disabilities, students who are EL, and students with 504s. For this layer of support, the PARCC accommodations manual (2013) specifies that the text-to-speech or the read aloud accommodation on the literacy assessments including items, response options, and passages,

"is intended to provide access to printed or written texts in the PARCC ELA/Literacy assessments to a very small number of students with disabilities who would otherwise be unable to participate in the assessment because their disability severely limits or prevents them from decoding printed text." (p. 27).

The policy for the read aloud accommodation is even more stringent for the

SBAC assessment. In a paper published on March 11, 2014

(http://www.smarterbalanced.org/wordpress/wp-content/uploads/2014/03/Read-Aloud-

Guidelines.pdf), the SBAC "Guidelines for Read Aloud, Test Readers" state "test readers

are allowable across all grades as a designated support for mathematics and ELA items as

appropriate (not ELA reading passages). Test readers are allowable for ELA reading

passages as a documented accommodation in grades 6 - 8, and 11." (p. 1). The published

SBAC policy for the read aloud accommodation has been made in spite of a "Literature

Review of Testing Accommodations and Accessibility Tools for Students with

Disabilities" published by Laitusis, Buzick, Stone, Hansen and Hakkinen (2012) that

suggests that "none of the studies specifically evaluated the impact of audio presentation

on other aspects of ELA such as writing prompts." (p. 27). In effect, the current testing

accommodations policies for both the PARCC and SBAC writing assessments lack a

research-base to determine whether or not the read aloud accommodation in the read and

respond writing assessment context would have an effect on overall writing performance.

In the absence of reading accommodations for this portion of the ELA assessment, there

is the possibility that the SBAC and PARCC writing assessments may be conflating the

assessment of the skill of reading with the skill of writing.

Given the new developments in standardized writing assessments at the state and

district level, further research on how students' writing performance may be affected by

their individual differences reading ability, as well as varying prompt conditions and

genres of writing assessments are warranted.

Theoretical Framework.

The theoretical framework that guides this study is the Cognitive Process Theory of Writing first forwarded by the seminal work of Hayes and Flower (1980), which was subsequently updated by Hayes (1996; 2006). It includes underlying elements of adult or mastered writing ability, and how cognitive demands in writing such as task environment, and memory affect children's writing performance (Bereiter & Scardamalia, 1987, & McCutchen, 1996).

Cognitive Process of Writing

Hayes and Flower observed that writing is goal-directed and that it involves three major components: (a) task environment, (b) cognitive processes and (c) a writer's long-term memory. In this model of writing, task environment involves elements such as attributes of the writing assignment (e.g., topic, audience, and motivating cues) as well as the text produced thus far (Graham, 2006). The cognitive processes in this model refer to the acts of planning, composing, and revising. Finally, a writer's long-term memory refers to the role that an author's prior knowledge on the topic and audience influence the plan or approach the author takes when composing. Collectively, all of the components of the Hayes and Flower (1980) model yield an approach to writing that is demanding in a number of ways. First, it requires that a writer employ a number of mental operations to meet the goals of the task. Second, these mental operations occur in a complex iterative and non-linear fashion. As a result, the writer must deal with a number of cognitive demands at once.

Figure 1. The cognitive process model of the composing process (Flower & Hayes, 1981)

Task Environment

Writing Assignment
- Topic
- Audience
- Motivating cues

External Storage
- Text produced
- Other resources

Long-Term Memory
- Knowledge of topic
- Knowledge of audience
- Stored writing plans
  - (writing skills)

Generating

Planning
- Organizing
- Goal-setting

Translating

Reviewing
- Reading
- Editing

Monitoring

Working Memory
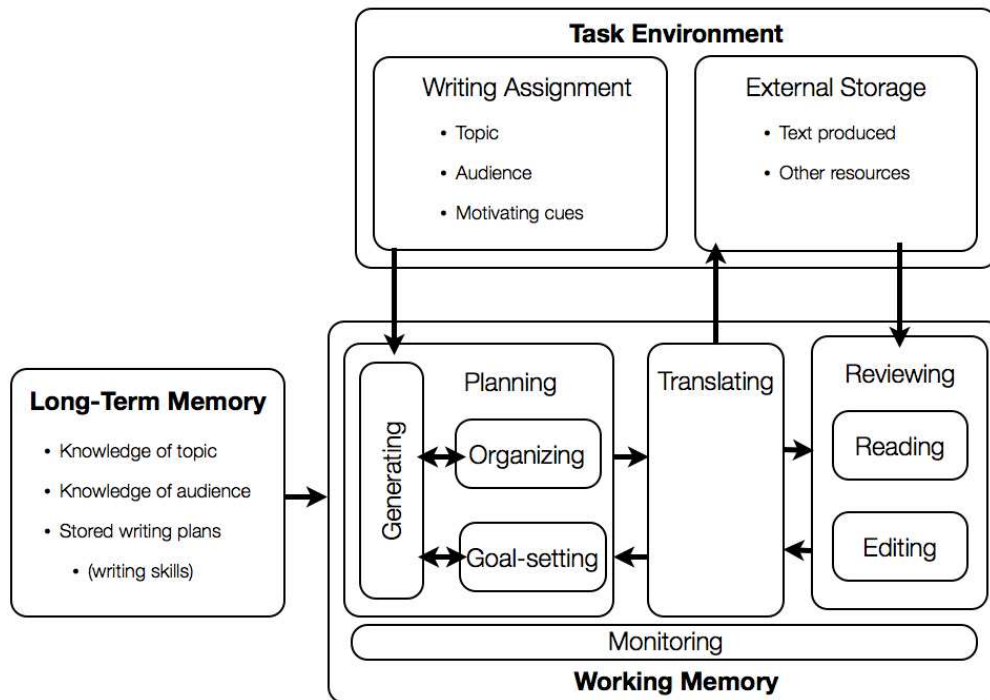
While the contributions of Hayes and Flower (1980) are substantial, some researchers did not consider this model of the writing process to be an accurate representation of how a young or novice writer might write. Bereiter and Scardamalia (1987) developed a model for writing that differentiates between novice and expert writers. More specifically, Bereiter and Scardamalia suggest that novice writers engage in a form of writing referred to as knowledge telling while expert writers engage in a form of writing called knowledge transformation. The primary distinction between these two modes of discourse lies in the role that prior knowledge has on how writers with varying maturity interpret the writing task and how they approach the overarching process of composing. Bereiter and Scardamalia contend that for children, limits on their prior knowledge and their simplification of the writing process have the potential to greatly constrain their ability to compose effectively.

Figure 2. Structure of the knowledge-transforming process (Bereiter & Scardamalia, 1987)
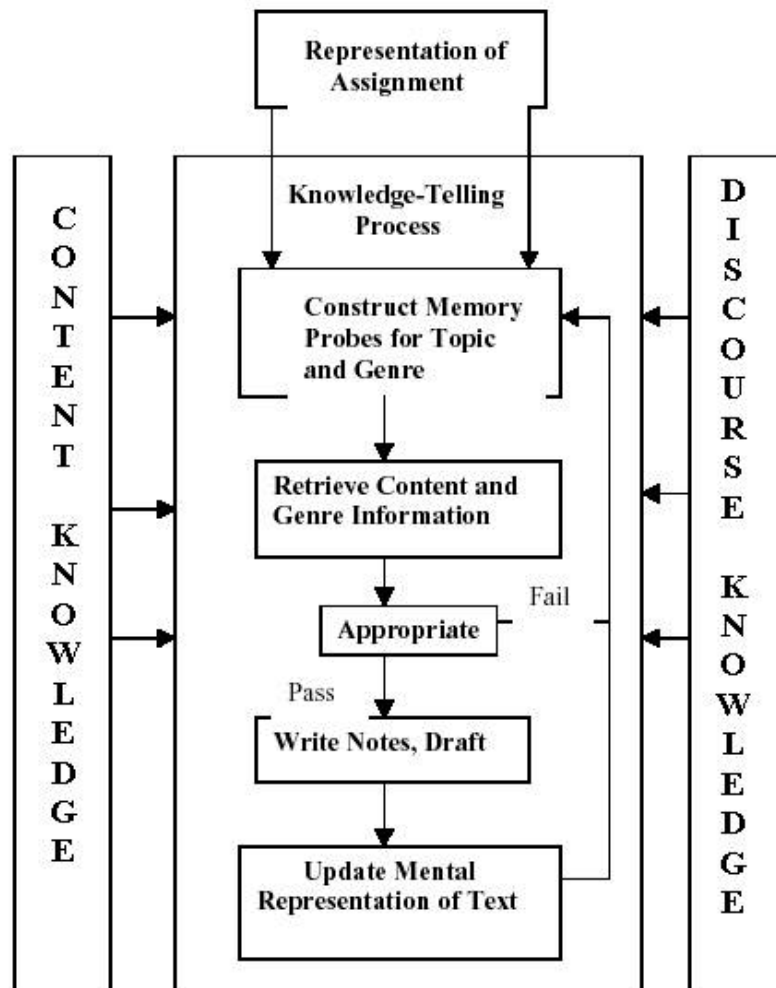
In contrast, adult writers have not only a greater fund of information on a wider array of topics, but also more experience; thus, they are generally are better prepared to plan text and engage in problem analysis and goal setting (Bereiter & Scardamalia 1987). In this context, problem analysis and goal setting refers to how the writer addresses concerns with content knowledge, and the text being written (i.e., a rhetorical space). Bereiter and Scardamalia suggest that experienced writers are not only better able to plan what they write; they are also better prepared to determine how to consider their audience

as they work to express their ideas. As such, experienced writers compose through a back-and-forth process of solving rhetorical and knowledge-related problems.

The distinction between novice and experienced writers holds significance for writing assessment. It clearly indicates that a child's prior knowledge on the writing topic and the writing genre can have an impact on their writing output. Given this finding, the topic and the writing task students are being asked to respond to should clearly be taken into consideration when generating assessments.

The importance of writing task was further elaborated in Hayes' (1996) later revised model of writing. In this model, Hayes places particular emphasis on the role that the task environment can have on a writer's composition. Here factors such as audience, texts read while writing, and the writing medium (e.g., hand-written vs. word processor) play a greater role. Further, in Hayes' new model, factors such as motivation and affect as well as the writer's schema on genre and linguistics have a more explicit link to text production. Hayes (1996) was also among the first writing theorists to integrate *working memory* into the writing process model. He suggests that working memory is comprised of three properties for processing information: phonological memory, visual and spatial information, and semantic memory (cited in Graham, 2006). All of these serve to work as a tandem link between motivation and affect, long-term memory, and the overarching cognitive processes involved in composing.

The role that *working memory* has on a child's writing process has been further elaborated in a review by McCutchen (1996). McCutchen contends that information from the environment and from long-term memory are stored during processing in working memory. She suggests that due to overall resource limitations, a trade-off exists such that

as more resources are used for processing information, fewer resources are available for storing information. Given this view on working memory, the implications for the complex process of writing become clearer. McCutchen argues that in order to compose a text, writers must, "coordinate within working memory planning goals (e.g., plans for content, audience, overall tone) and product goals (e.g., requirements of grammaticality, plan, fulfillment) while language generation processes retrieve words to express content to organize those words into appropriate text" (McCutchen, 1996, p. 301). Given these demands, it follows that if processing or storage capacities are compromised, one's overall writing performance will be negatively affected.

The interaction between working memory and the translating or transcription and text generation process particularly highlights the effect of working memory on children's writing. McCutchen (2000) found that inefficient processes in fluent text production such as the coordination of fine-motor skills for handwriting, or spelling can hinder a child's ability to compose, plan, and revise. Studies by Daiute (1984) and Fayol, Largy and Lemaire (1994) support this theory. In his study, Daiute was able to find negative correlations between short-term memory capacity and the frequency of errors in children's written texts. Fayol et al., found that by increasing working memory demands in a writing task, they could increase subject-verb agreement errors.

Findings related to the relationship between working memory and transcription in young children support Bereiter and Scardamalia's (1987) model of writing development. Children focus more energy on knowledge telling, and less energy on planning and revising when composing. Further, limits on prior knowledge with respect to the task environment (e.g., topic and genre) as well as working memory capacity, motivation and

10

affect all seem to have implications for students' overall text production. Given these potential limitations on young children's writing ability, it holds that any developmentally appropriate writing assessment should make an effort to minimize the effects of prior knowledge and working memory demands through task environments that support students' ability to write. Writing assessments that consider the above factors will likely help to obtain the best representation of a child's overall composing ability.

Study Context and Rationale

This quasi-experimental study examines the writing of 63, 3rd- and 5th-grade students at a public charter school (PK-6) in an urban school district in the Mid Atlantic. The participants in this study belong to a culturally diverse school of 350 students: typically 47% of the students are African American, 44% are Hispanic, 9% are Caucasian, and 1% are Asian American. Approximately 11% of the population receives special education services and 83.5% of students are eligible for free or reduced-price lunches. In addition, 45% of the students come from homes in which a language other than English is spoken (including Spanish, French, Amharic, Woolof, Arabic, Chinese, and Yoruba). Another contextual factor for this school is that all students, including monolingual/native English speakers, participate in either French or Spanish language immersion for half their school week. In addition, all students participate in their English language arts classes in English, with written language comprising a large component of instruction. The sample in this study is approximately representative of the school population with a sample comprised of 65.1% general education students, 11.1% students with LD, 15.9% students who are EL, and 7.9% of the students are both students with LD and who are EL. In contrast to the general population at this school, the sample is

comprised of mostly African American students (66.7%), and some Hispanic (25.4%) and Caucasian students (7.9%) with no Asian students.

Recent demographic information from within the past two years indicated that the majority of students from this school were from low-income families (84%). Information from the school district's Assessment and Accountability Data reported that 90% of the students tested at the school were considered "economically disadvantaged." These figures are considerably higher than the entire school district (66%).

Regardless of their socio-economic status, the school district's standardized reading assessment revealed that during the 2010-2011 academic year 58% of students at this charter school were identified as proficient readers, as compared to only 44% in the entire school district. Even more interesting was the difference in scores for African American students as 71% of students were identified as proficient readers compared to 39% in the district. This percentage is second only to a highly academic charter school where 77% of students were identified as proficient readers. In contrast, only 39 % of students identified as English Language learners (EL) were found to be proficient readers. Moreover, only 20% of students with disabilities at this school were identified as proficient readers.

The purpose of this study is to address gaps in the extant literature on writing assessment. In particular, the goal of this study is to evaluate the relationship between task environment and prior knowledge on academically, culturally, and linguistically diverse elementary children's performance across the three genres of writing they are most commonly asked to produce in schools (narrative, informational, and persuasive). Current research on how and why various groups of students compose differently across

these genres is limited due to inconsistent measurement and assessment procedures. In addition, there is limited research that simultaneously examines the interaction of prompt condition and students' performance within and across genres across grade levels.

Writing is a complex iterative process that requires the integration of multiple skills and cognitive abilities that develop in a non-linear fashion (Bereiter & Scardamalia, 1986). As a result, writing can be a difficult academic area to assess. This study seeks to find ways to improve the construct and content validity of writing assessments through the exploration of diverse groups of students' performance on two forms of prompting conditions: supported and unsupported. In both conditions, I provided students with topic content, audience, and genre goals. This information was provided to students in an attempt to moderate potential prior knowledge effects of topic and genre, and to isolate the potential effect of prompt condition on students' writing performance. In this study, prompt condition refers to whether or not students receive the read aloud accommodation for all texts and passages that are part of the writing prompt. In the supported condition, all students were provided a read aloud accommodation for all text in the prompt including reading passages and genre specific cues. In the unsupported condition, students were not provided the read aloud accommodation and were only read the directions. My initial hypothesis was that children in the supported condition would write better quality essays. More specifically, I hypothesized that all children, and in particular, younger students and students who are developing readers in the supported condition would write better quality essays than peers without the read aloud accommodation. Ultimately, this research addresses the following research questions:

1. What is the effect of genre on 3rd- and 5th-grade students' overall writing

quality and sentence level skills? Is the effect of genre similar or different across grade levels?

2. What is the effect of prompt condition (supported or unsupported) on 3rd- and 5th-grade students' overall writing quality and sentence level skills? Is the effect of prompt condition similar or different across grade levels?

Dissertation Organization

This dissertation is organized into five chapters. The current chapter includes the problem, theoretical framework, study context, rationale, and research questions that guide this study. In Chapter 2, I first present an overview of critical issues in writing assessment, which includes: (a) writing assessment measures, (b) writing assessment purposes, and (c) writing assessment forms. In this chapter I also introduce current research on students with LD and students who are EL, who are of particular interest given the population in the setting of my proposed study. This overview of critical issues in writing assessment will be used to situate the subsequent content and methodological review of empirical studies specific to elementary writers within the broader scope of writing assessment research. More specifically, I will present research that examines the effects of prior knowledge, prompt condition, and genre for elementary writers. I will also identify gaps in the extant research that provides the justification for my study. In Chapter 3, I review my research design, as well as methods, analysis, and procedures. Chapter 3 also includes an elaboration of the study's primary research questions as well as expected outcomes. In Chapter 4, I provide the results of my study including descriptive summaries and answers to each of my research questions using, descriptive statistics, paired sample t-tests, and MANOVA procedures. Finally, in Chapter 5 I discuss

the significance of my study, the limitations, as well as potential future directions for research.

Significance of Study

In a report from the Carnegie Corporation of New York, Graham, Harris, and Hebert (2011) suggest, "when teachers assess or monitor students' writing progress, it has a positive and statistically significant impact on students' overall writing performance" (p.19). They suggest that there is a need for the development of new "formative and summative assessments that are reliable, valid, and fair, as well as methods for determining how such assessments can best enhance writing instruction and students' writing development" (p. 31). In particular, they recommend that the field needs to develop a better understanding of how writing develops, and how writing assessments can minimize or eliminate factors that bias or invalidate such assessments.

This study is significant for several reasons. First, the findings in this study may help researchers gain a better understanding of children's writing development within and across genres. Second, this study seeks to minimize the effects of task environment in writing assessments and consider the effect of the read aloud accommodation on writing assessments that require reading. The findings from this study help to identify factors that may be biasing or invalidating results particularly for younger students and students who are developing readers. By examining these factors, findings from this study may provide stakeholders with assessment tools and standards that more clearly reflect the extent to which students' performance within and across genres may be a function of the assessment task, topic, or individual students' developmental growth and ability. Ultimately, this study forwards the overarching agenda of improving writing assessments

so that teachers can make informed decisions on how to deliver instruction that will positively impact struggling writers.

Definition of Key Terms

Cognitive processes: the acts of planning, composing, and revising (Hayes & Flower, 1980; Hayes, 1996; 2006).

Genre: different modes of discourse or rhetorical structure that have a distinctive style, form or content (Kamberlis, 1999)

Informational report: "writing is subject oriented. The focus of this kind of writing is on presenting information about the subject, rather than on the writer" (Prater & Padia, 1983, p. 129).

Knowledge-telling: a form of writing typically used by novice writers where the writer writes down everything they know about a topic with less concern for rhetorical and discourse goals (Bereiter & Scardamalia, 1987).

Knowledge-transforming: a form of writing typically used by experienced or adult writers that involves goal setting and problem solving while writing for a particular purpose (Bereiter & Scardamalia, 1987).

Long-term memory: the role that an author's prior knowledge on the topic and audience influence the plan or approach the author takes when composing (Graham, 2006).

Narrative or story: includes a setting (e.g., character, locale, time) and plot (e.g., initiating event, goal, direct consequence; Olinghouse & Santangelo, 2011).

Persuasive: "writing is audience oriented. The writer takes a position and supports it in an effort to convince an audience" (Prater & Padia, 1983, p. 129).

Prior knowledge: the whole of a person's actual knowledge that: (a) is available before a

certain learning task, (b) is structured in schemata, (c) is declarative and

procedural, (d) is partly explicit and partly tacit, (e) and is dynamic in nature and

stored in the knowledge base" (Dochy, 1994, p. 4699).

Prompt condition: involves the discourse mode, rhetorical specification, wording and

structure, and cues given to the writer on the writing assessment (Huot, 1999).

Task-environment: involves elements such as attributes of the writing assignment (e.g.,

topic, audience, and motivating cues) as well as the text produced thus far

(Graham, 2006).

T-unit: a dominant clause and its dependent clauses or one main clause with all

subordinate clauses attached to it (Hunt, 1965, p. 20).

Working-memory: primarily comprised of text generation (content selection, lexical

retrieval, and syntactic processes) and transcription (the cognitive and physical

acts of forming written representations of text (McCutchen, 2000).

CHAPTER 2: LITERATURE REVIEW

The following chapter is divided into two principal sections. The first section presents a broad conceptualization of the literature by discussing content that is relevant to the overarching purpose of this study, which is to explore critical factors related to writing assessment for elementary students. This section includes a broad base of material (such as literature reviews, position papers, as well as empirical studies) on current issues in writing assessment such as writing assessment measures, purposes, and forms.

In the second section, I begin by presenting the methods and search criteria I used to select the empirical studies that will serve as the basis for my own future research. Towards that end, I present a content and methodological review of empirical studies that examine the effect of prior knowledge, prompting conditions, and genre on elementary children's writing. I then provide conclusions, both by identifying limitations to the extant research and by summarizing what researchers know about the effects of prior knowledge, prompting conditions, and genre for elementary writers. Finally, I outline the rationale for my proposed study.

Overview of Critical Issues in Writing Assessment

Quellmallz et al. (1982) noted that the challenge in designing writing assessments arises from the requirement that assessments possess construct, content and ecological validity. Moreover, the researchers note that in order for tests to be valid, they not only need to be concerned with the form and content of the assessment, but also how these assessments are used as the basis for evaluating students' performance. Quellmallz et al. present some of the unique challenges to researching the topic of writing assessment.

Writing, as a construct, is a complex and multifaceted skill that includes the interaction of factors such as the learner's characteristics, and the writing task. Writing assessment not only needs to consider the role of the learner and the writing task, but also the methods and procedures for evaluating learners and the potential implications of the results. Given the comprehensive nature of writing assessment as a topic and the limits of what can be accomplished in a single study, it was important to first review writing assessment broadly to identify sub-topics that will serve as the primary foci for my future research.

Accordingly, the purpose of this overview on critical issues in writing assessment is two-fold. First, given the new SBAC and PARCC assessment context, it was important to situate the current study within the broader framework of issues related to writing assessment. This helped to identify factors with the greatest priority within an elementary school context. Second, a preliminary search of the literature using key phrases such as *writing assessment, writing tests, writing measures, writing assessment validity, writing assessment purposes, prior knowledge, background knowledge,* and *task environment,* resulted in literature that was primarily focused on secondary and post-secondary writers. While research at the secondary and post-secondary level does not directly apply to the context and rationale for the current study, this work provides valuable information on critical issues in writing assessment that informs this work.

Several literature reviews, position papers, published books, and articles that related to writing assessment were found as a result of this search. This body of work fell into three primary categories: (a) writing assessment measures, (b) writing assessment purposes, and (c) writing assessment forms. In each of these sections, seminal literature

reviews and position papers became the foundation for subsequent ancestral searches of articles relevant to these topics.

For example, in the section on writing assessment measures, Olinghouse and Santangelo (2010) and Huot's (1999) literature reviews on writing assessment serve as the starting points for subsequent searches for articles related to measures such as *Curriculum Based Measurement*, and *holistic, analytic*, and *primary trait rubrics*. In the section on writing assessment purposes, Olinghouse and Santangelo's review and Gebhard and Harman's (2011) position paper on writing assessment and English language learners were especially informative. In this section, I identify how writing assessments are used in the standardized test context to identify at-risk learners. In addition, I present a broad overview of important factors to consider when examining the writing of these populations of students, which is relevant when analyzing and interpreting writing assessment results. Finally, in the section on writing assessment forms, I use Huot (1990) and Dochy, Segers, and Buehl 's (1999) literature reviews as well as a position paper by Myhill (2005) as the basis for ancestral searches on topics related to the form and structure of writing assessments, and their potential interaction with a learner's prior knowledge.

After each subsection in this first portion of Chapter 2, I present implications for the current study. Finally, I close with a summary of findings that serve as the basis for narrowing the goals of my empirical review to studies and topics specifically related to elementary writers. This empirical review will be presented in the second half of Chapter 2.

Writing assessment measures

Writing involves the integration of multiple skills and sub-skills in order for children to generate a coherent and well-executed product. These skills and sub-skills include: (a) letter and word-level features like handwriting, spelling, and vocabulary; (b) sentence level features like punctuation, capitalization, and grammar or syntactical structure, and (c) text level features such as organization using paragraphs, structures specific to different genres of writing (e.g., descriptive, sequence), and coherence through the use of main ideas, details, and transitions (Olinghouse and Santangelo, 2010). Given the complex nature of writing, the assessment of writing requires an evaluation of multiple aspects of writing. Nevertheless, reliable and valid measures of growth and development in writing ability are often difficult to quantify. In the following sections I present four forms of writing measures that are frequently used to assess the writing of children in schools. These measures include Curriculum Based Measurements (CBM), and rubric-based measurements (e.g., holistic, analytic, and primary trait assessments). These measures are examined for both strengths and weaknesses, and will provide background on how the measures for the proposed study were developed.

CBM. One approach to writing assessment that has been successful is Curriculum Based Measurements Curriculum Based Measurements (CBM) involves the repeated sampling of student performance in the curriculum. Deno (1985) stated that CBM decreases "the separation between measurement and instruction—to make data on student achievement more integral to daily teacher decision making" (p. 221).

When CBM was first researched and developed, four design characteristics were specified. Measures should be: (a) reliable and valid to ensure that the results would be

accepted as evidence of student achievement and the basis for making instructional

decisions; (b) simple and efficient for teachers to frequently monitor student

achievement; (c) easily understood so that results could be clearly communicated to

parents, teachers, and students; and (d) inexpensive since multiple forms were to be

required for repeated measurement (Deno, 1985).

Deno (2003) later clarified the characteristics of curriculum-based assessments

relative to other standardized test measures. More specifically, in CBM (a) the curriculum

materials used for instruction are the test stimuli; (b) emphasis is placed on direct

observation and recording of student performance in response to selected curriculum

materials; (c) interobserver agreement is used to establish reliability of data collected; and

(d) social validity is the basis for justifying the use of information gathered. In effect,

curriculum materials, daily instruction, and assessment of performance are inextricably

linked.

To date, there is a large body of research that has established reliable and valid

measurements of students' technical writing ability using CBM (see, e.g., Benson &

Campbell, 2009; Espin et al., 2004; Jewell & Malecki, 2005; McMaster & Campbell,

2008; McMaster & Espin, 2007; Powell-Smith & Shinn, 2004, Weissenburger & Espin,

2005). Here, technical writing ability refers to the letter, word, and sentence level

features of writing described earlier in this section.

For example, Parker, Tindal, and Hasbrouch (1991a, 1991b; Tindal & Parker,

1989) demonstrated the effective use of CBM in their analysis of students' writing

samples using total words written (TWW), words spelled correctly (WSC), correct word

sequence (CWS), and words written legibly. All CBM measures were able to show

growth from fall to spring across grade levels; however, differences were smaller and less stable for older students (Espin et al., 2004). Subsequent studies such as Espin, De La Paz, Scierka, and Roelofs (2005) examined the relationship between curriculum-based measures (in this case CWS and correct minus incorrect word sequences or CWIS) relative to criterion measures such as the number of functional elements in and quality ratings of student essays. Results from this study revealed a strong relationship between curriculum-based and criterion measures. McMaster and Espin (2007) demonstrated that CWS and CWIS were reliable and valid measures of growth in narrative prompts in Grade 7 from fall to spring. On expository tests, CWIS was found to be a reliable and valid measure of growth for 7th-graders.

McMaster and Espin (2007) noted one significant limitation in their study as well as in previous studies that applied CBM to written expression. While CBM measures of written expression such as CWS, CWIS, TWW, and WSC have been established as reliable and valid measures of growth in letter, word, and sentence level features, they are not likely to capture all of the critical dimensions of writing. Towards that end, other measures of writing quality and text level features are necessary in order to evaluate the full construct of writing.

Rubric Measures. There are three primary approaches for assessing quality of writing using rubrics: holistic, analytic, and primary trait. Holistic rubrics are used as a measure of overall writing quality. Scoring can be norm-referenced (compared to a group) or criterion referenced (compared to a pre-specified set of criteria) and are frequently used in large-scale writing assessments such as the National Assessment of Educational Progress (NAEP). For example, in the NAEP writing assessment, each rubric

includes a 6-point scale (1=low, 6=high) related to overall text structure, idea

development, sentence structure, and mechanics quality (see

http://nces.ed.gov/nationsreportcard/). According to Olinghouse and Santangelo (2010),

while holistic measures are considered economical and quick, they do not work as well

for progress monitoring purposes. They suggest that the 1-6 point scales usually

represents a significant jump from one point to the next and as such, incremental growth

is more difficult to measure. In addition, while holistic measures are useful tools for

identifying a child's overall writing ability, Olinghouse and Santangelo contend that they

are less useful in identifying specific areas for intervention as individual writing

processes are not isolated from each other. Holistic measures are also commonly used in

writing research because it is possible to capture a large range in writing performance

within a target population (e.g., third graders).

In contrast, analytic rubrics allow educators to evaluate specific areas of writing

separately (e.g., conventions, organization, word choice). The advantages of analytic

rubrics are that unlike holistic rubrics, individual areas for intervention can be isolated.

The disadvantage for such rubrics is that they often take longer to score taking

approximately 1 to 2 minutes to score each trait versus 1 to 2 minutes per paper for

holistic scoring (Spandel & Stiggins, 1980). Additionally, research has shown that

holistic scores correlate reasonably well with analytic scores (Freedman, 1984).

Primary trait rubrics, like analytic rubrics involve the evaluation of specific areas

of writing separately. Where primary trait rubrics differ from analytic rubrics are in the

specification of the traits under review. The basis behind primary trait assessments is that

the genre or specific discourse goals create the criteria for writing quality (Lloyd-Jones,

1977). Like analytic rubrics, however, cost and time is a consideration. According to Spandel and Stiggins (1980), primary trait rubrics also require approximately 1 to 2 minutes per trait. In addition, Veal and Hudson (1983) noted that primary trait rubrics often have lower correlations with other measures of writing quality. The combination of lower correlations and higher cost and time considerations can make this form of writing rubric less preferable to holistic and analytic rubrics.

Important considerations when evaluating the use of rubric-based measures of writing quality are the factors that influence how these measures are found to be valid and reliable. One way to establish the reliability and validity of rubric measures is through establishing adequate interrater reliability. Interrater reliability refers to the extent that two raters can establish agreement on a score. In a literature review by Huot (1990), he notes that research on interrater reliability suggests that at times, the goal of establishing interrater reliability may work against a raters' natural response to a students' writing (Barritt, Stock, & Clark,1986), and that conflicting responses might be treated as inaccurate in an effort to reach agreement (Stock & Robinson, 1987). In addition, Huot found that raters who were charged with using rubric measures were "more sensitive to content and organization than to sentence structure and mechanics" (p. 256). The above findings related to the role that interrater reliability and rater judgments have on evaluations of writing quality indicate the need to carefully train raters to attend carefully to benchmarks and guidelines for accurate use in scoring student writing.

Implications for the study. Based on the broad overview of literature related to writing assessment measures, there are specific implications for the current study. First, it is clear that while rubric measures are valuable tools in evaluating overall writing quality,

these measures have certain limitations. For example, primary trait rubrics are useful tools for evaluating specific discourse criteria, however, the additional time it takes to use these types of measures coupled with their low level of correlation with other assessment procedures may not be as useful for the purposes of the current study. Holistic measures can be used to compare students' overall writing quality to their peers, however, the published SBAC and PARCC rubrics all fall into the category of analytic measures. The benefits of analytic measures are that they can serve to isolate specific writing skills for the purposes of determining potential areas for intervention. Special concern for the effect that rater judgment may have on writing scores, particularly with respect to a potential preferential bias towards content and organization above and beyond sentence structure and mechanics are considered. Additionally, both the PARCC and the SBAC writing rubrics use analytic measures that include categories for organization, use of evidence from texts, and conventions on 4-point scales. On the PARCC assessment, an additional category dedicated to the raters' assessment of a student's demonstration of reading comprehension through students' essay is included as well. Accordingly, for the purposes of this study, an analytic rubric measure was selected over a holistic measure to ensure alignment with the writing assessment context.

Finally, the use of curriculum-based measures such as correct-word sequence or total words written may be useful in ensuring that these word and sentence-level features of writing are captured in the evaluation process as well. Given the overlapping nature of many CBM measures, an effort should be made to ensure that measures are independent.

Writing Assessment Purposes

Olinghouse and Santangelo (2010) suggest that there are four primary purposes for assessing students' writing. These include assessing: (a) to identify children who are at-risk for school failure, (b) to inform instructional planning and modification, (c) to monitor students' progress, and (d) to identify students for eligibility for special education services. Since the reauthorization of the Elementary and Secondary Education Act (ESEA) and the implementation of No Child Left Behind (NCLB) in 2001, a fifth possible purpose of assessment has been to determine the allocation of federal funding to schools.

Given the high-stakes nature of state and district-wide standardized assessments, stakeholders (i.e. policy makers, schools, parents, and teachers) need assessments that accurately reflect and measure how children's writing develops. The need for accuracy and equity is particularly true for students with LD and students who are EL. For students with LD, one concern is that children may be incorrectly determined eligible or ineligible for special education services based on these types of assessments. For students who are EL, Gebhard and Harman (2011) argue that children who are identified as struggling English language learners are often relegated to EL classrooms where they are exposed to "truncated, inauthentic reading material and little practice composing extended texts beyond the word or sentence level" (p. 46). Darling-Hammond (2006) further contends that the consequences for students who are EL and their teachers in the NCLB context is that it may incentivize schools to allow or even encourage their struggling students to leave.

According to the 2002 National Assessment of Education Progress (NAEP) students with LD and students who are EL from the school district in this study's setting had only 7% and 1% of students performing at the proficient level in writing respectively. In essence, most students with LD and students who are EL still struggle to meet basic levels of writing competency in the targeted school district. In addition, nearly 23.8% of the population in the current study includes students who are EL and 11.1% of the population includes students with disabilities. Given the above concerns regarding the potential marginalization of at-risk students with LD and students who are EL, a concerted effort to ensure fair and equitable practices in writing assessment are warranted.

The broader purpose of this study is to evaluate the effect that genre and prompt condition may have on elementary aged students including culturally, linguistically, and academically diverse students. Given the context and rationale for this study, it is important to specifically review what researchers know about the writing development and abilities of students with LD and students who are EL in order to be able to effectively collect and analyze assessment data for these specific populations. The following brief overview will examine the effects of disability and the effects of English language learner characteristics on students' performance in writing. This information will provide context for the broader goals of the current study, which will be presented in further detail in the next section of this chapter.

Students with learning disabilities. Writing can be especially difficult for children with learning disabilities. Limits on comprehension, working memory, phonological awareness, spelling, transcription skills and overall executive control are

often times amplified in a child with a LD. Compared to peers without disabilities, children with disabilities often produce texts that are shorter, poorly organized, are often incomplete, and are generally weaker in quality (Troia, 2006). Studies have also shown that the writing of students with disabilities often contain more mechanical and grammatical errors (Graham, 1990; Graham, Harris, MacArthur & Schwartz, 1991; Graham, Harris & Fink, 2000). Researchers suggest that these problems may be attributed to difficulties in some of the underlying aspects of the writing process and children's writing development (Englert, Rapahel, Fear, & Anderson, 1988; Graham, 1997; and De La Paz, Swanson, & Graham; 1998). This is particularly true for aspects of writing such as planning, content generation, revising, and general text transcription.

Houck and Billingsley (1989) found that students with LD demonstrated a number of areas of weakness relative to normally achieving peers. They wrote fewer words and sentences, produced fewer words with seven letters or more, wrote more sentence fragments, and had a higher percentage of capitalization and spelling errors. One interesting finding they noted was that children with disabilities produced more words per sentence. While at first glance this seems counter-intuitive, Houck and Billingsley suspect that this finding might be due to the likelihood that students with LD wrote more run-on sentences. They also found that there were no overall group differences between normally achieving students and students with LD in the number of t-units and mean morphemes per t-unit. This finding is aligned with findings from other studies that have suggested that t-units were not an effective measure of overall writing development (Stewart & Grobe, 1979, Nelson, 2011).

One explanation for the finding that children with disabilities often struggle with written production could be related to deficits in letter and word level aspects of writing such as handwriting accuracy, legibility and spelling (Graham & Weintraub, 1996). Students with challenges in these areas of transcription may avoid writing words they don't know how to spell, or write less than they might be able to express orally in order to shorten the writing process (Olinghouse & Santangelo, 2010). De La Paz and Graham (1995) and McCutchen (1996) also suggest that having to attend to conventions such as spelling can cause novice writers to forget their ideas or plans for writing. As such, students may only choose to include words they know how to spell. This can often result in writing with less diverse vocabulary and vocabulary that is below the writer's grade level (Olinghouse & Santangelo, 2010). Poor spelling in particular can cause a reader to perceive a students' writing to be poor quality regardless of its content (Chase, 1986).

Students with LD often have more grammatical errors such as incorrect use of verb tenses and articles at the sentence level (Scott & Windsor, 2000). Olinghouse and Santangelo (2010) suggest that limited prior knowledge regarding sentence conventions (e.g., correct word order and subject-verb agreement) can hinder the ability to write complex sentences. Perhaps as a result of many of these difficulties, struggling writers' compositions may primarily contain simple or repetitive sentence structures (Olinghouse & Santangelo, 2010).

Planning and organizing lengthier more complex texts is another common challenge for novices, and in particular, for students with LD. According to Hayes and his colleagues (Hayes & Flower, 1987; Hayes, 2006), planning involves: (a) formulating, prioritizing, and modifying goals and subgoals to address task and genre demands, and

perceived audience needs, (b) generating ideas, and (c) selecting and organizing content for meeting established goals. Moreover, novice writers and students with LD often start to write immediately after an assignment has been given, and generate text in an associative manner (Donovan, 2001). As they write, young and struggling writers often overlook considerations for text organization, rhetorical structures, genre and audience, as novices place a primary focus on the demands of the transcription process (Bereiter & Scardamlia, 1987; Graham, 1990; McCutchen, 1988). Olinghouse and Santangelo (2010) suggest that novice and struggling writers frequently produce texts in the form of one long paragraph for every composition. Alternatively, they may write multiple paragraphs that lack coherence and organization and jump from topic to topic. In addition, Olinghouse and Santangelo suggest that struggling writers may not be able to change the structure of their writing depending on the expected text genre.

   Implications for the current study. In essence, there are a number of barriers that can impede a student with LD's ability to write effectively. Challenges with the writing process including planning and text transcription all appear to be more difficult and they may be particularly vulnerable to these obstacles due to limitations in some of the underlying cognitive processes that are necessary for effective writing.

   It seems reasonable to believe that information about cognitive processing writing theories and information about how novice and struggling writers should be used to inform the design and interpretation of writing assessments. More specifically, it would be helpful if writing assessments were designed to reveal the absence or presence of surface level (e.g., transcription) difficulties or higher level (e.g., planning) problems in students, particularly at the elementary school level. This may, in turn, allow for the

identification of struggling writers for the purposes of intervention. Nevertheless, while identifying struggling writers is an important goal of assessment, assessments can also be used to find students eligible or ineligible for special education services. As such, a concerted effort should be made to create valid assessments that eliminate all factors that may be potentially biasing the assessments against students who may simply struggle with transcription or in level of prior knowledge. The elimination of potential bias will ensure that the assessment is an authentic representation of a student's overall writing ability, and not simply their performance on handwriting or their performance in a specific context that would not generalize given another topic or task environment.

Students who are English language learners. The intersection of writing and second language literacy presents a unique challenge to students who are EL (Strickland & Alvermann, 2004). As identified by the work of Sapir (1921), language is a cultural construct or artifact. As such, students' cultural background is likely to have a significant effect on the use of language both in oral and written formats. Second, students' ability to engage in more complex language structures and formats is related to social background and language proficiency (Loban, 1976). This is in line with Scarcella's (2003) findings that suggest that while in recent years teachers have improved their ability to teach skill-based literacy components to students who are EL, many students' literacy problems revolve more around a failure to acquire knowledge of academic English rather than their ability to perform discrete literacy skills such as spelling and simple sentence writing. She contends that as students move into the upper grades, increased attention on the distinctions between conversational and academic English must be made to provide students better access to the curriculum. Given these findings, it is no surprise that

language background has the potential to significantly affect students' performance in writing.

According to Hooper and Enright (2011) much of the research on the writing of students who are EL has focused on either the emergent literacy or the transition of these students from high school to higher education. Nevertheless, research on students who are EL is sparse and very little is known about how their writing develops, particularly at the elementary level. What is known from the research; however, is that students who are EL are vulnerable to school failure relative to their native-English speaking peers (Strickland & Alvermann, 2004).

Crosson, Matsumura, Correnti, and Arlotta-Guerrero (2012) suggest that much of the difficulty that students who are EL have with English-language writing tasks is due to "lack of familiarity with the lexical, grammatical, and discursive features that are associated with academic language" (p. 470). One underlying premise that exists in learning English as a second language is that first language (L1) competencies transfer to similarly cognitively demanding tasks in the second language (L2; Cummins, 1979). Furthermore, scholars (e.g., Colombi & Schleppegrell, 2002; Scarcell, 2002; Schleppergrell, 2004) and practitioners alike (c.f., Teachers of English to Speakers of Other Languages [TESOL] 2009) have asserted that competency in L1 has the potential to contribute to the development of competency in L2.

Crosson et al. (2012) explored this phenomenon in their study of 4th and 5th-grade writers in Spanish-English bilingual classrooms in the Southwest. Over the course of a two-year period, the researchers collected 224 writing samples or 4 samples for each writing task assigned from 12 different schools. They found that the quality of the writing

tasks assigned to students who are EL was generally of low cognitive demand involving tasks such as surface-level summaries or recalling fragmented information from reading assignments. They concluded that students were rarely asked to use features of academic language when writing in English or their native language. This was in-line with the findings of Colombi and Roca (2003) who found that school systems rarely give students the opportunity to develop advanced levels in their native language, which would theoretically support their ability to write more effectively in the language demanded by their school.

Crosson et al. (2012) also found that the cognitive demand of tasks was a significant predictor of students' use of features such as academic vocabulary, embedded clauses, temporal and causal connectives, and use of a variety of connectives. In particular, Crosson et al. found that low-cognitive demand tasks yielded writing samples with little to no evidence of the lexical and grammatical features of academic vocabulary. Instead, these types of writing tasks produced written products that contained basic and nonspecific word choice, and incomplete sentences. In addition, they found that cognitive demand of the task also predicted the overall quality of students' writing.

In a qualitative study of grade 3-5 students who are EL's writing, Brisk (2012) examined the relationship between students' understanding of first, second and third person grammatical markers and their ability to write in multiple genres. She found that students made successful attempts to write in multiple genres, but there use of incorrect grammatical person in certain genre contexts indicated a lack of understanding of academic English. For example, Brisk found that it was common for students who are EL to use the pronoun *we* in all the genres where American academic writing expects the

first person singular (e.g., narrative writing) for the narrator of the text. In addition, she

noted that many students who are EL have only an emerging understanding of a

connection to a particular audience for a particular genre. For example, in academic

writing, frequently students are expected to write in the third person to establish distance

from the audience and place emphasis on the topic. Brisk noted that many students who

are EL inappropriately use *I* and *you* in more formal expository texts due to their

misunderstanding of genre and audience.

While these findings are significant, Brisk (2012) and Crosson et al.'s (2012)

studies represent two of the few studies that specifically examine the development of

writers who are EL at the elementary level. At the middle-school level, students who are

EL have been shown to struggle with grammatical features of writing such as the

appropriate use of past tenses, prepositions, and second-person pronouns (Reynolds,

2005). In addition, in another study of middle school writers, students who are EL were

found to write shorter texts when writing in their second language (Stevenson, Schoonen,

& de Glopper, 2006). These findings at the middle school level broadly support the work

of Crosson et al. in that they reinforce the notion that learning a new language has a

significant effect on students' writing performance in the target/transfer language.

Implications for current study. There are multiple implications for writing

assessment given this literature on the effects of second language development on

writing. First, it is clear that if a student has limited knowledge and understanding of the

cultural expectations of academic language, the quality of their writing may be negatively

affected. This limited understanding of academic English suggests that measures that

examine factors such as syntax, vocabulary, and word usage would be useful to isolate

potential areas for instruction. Second, Crosson et al. (2012) suggest that the cognitive

demands of the writing task (high vs. low) can play a significant role in students' writing

performance both in overall writing output and in their use of appropriate grammatical

structures. It holds then that an effort should be made to present tasks that are sufficiently

cognitively demanding, but that do not overwhelm the writer who is learning a second

language with topics and tasks that are unfamiliar.

A final consideration with respect to writing assessment and instruction for

students who are EL comes from Hyland (2003). He suggests that in order for students to

gain a better understanding of how to write effectively in their second language, they

require not only direct instruction in features of writing such as grammar, spelling, and

composition, they also benefit from direct instruction in genres that are constructed

directly from the social context. Therefore, in essence, by focusing assessment on the

writing development of students across genres, a culturally contextualized view of

writing may be forwarded that would significantly benefit students who are EL.

Writing Assessment Forms

Another important aspect of writing assessment is the manner by which one can

solicit students' writing. There have been a number of different studies, literature

reviews, and position papers that have suggested that variability in the way that writing is

solicited from students can have a significant impact on their writing output and in turn

the evaluation of their abilities. In particular, research has suggested that factors of

writing prompt such as the task and topic may have an impact the types of scores given

by raters (Hoetker, 1982).

In Huot's (1990) literature review on concerns and prevailing trends on direct writing assessments, he highlights many of the issues that researchers and educators have determined are still important in writing assessment today. His primary foci include the following components of the writing prompt condition: genre mode, rhetorical specification (e.g.. audience), and the wording and structure of the writing prompt itself. Huot reviews a number of studies that are of interest in relationship to the above topics, which are synthesized in the following sections.

Another area of concern with respect to the form of writing assessments is the relationship between the topic of the prompt and the writers' prior knowledge. To explore the relationship between writing prompt topics and prior knowledge I present a literature review by Dochy, Segers, and Buehl (1999) and a position paper by Myhill (2005).

Prompt conditions. Within the broader theoretical framework of Cognitive Process Theory (Hayes, 1996; 2000), elements of the written assignment including both content and discourse goals can have a significant impact on children's writing performance. Elements of the prompting condition that may influence content and discourse goals can include the genre mode, rhetorical specifications, and the wording and structure of the prompt itself.

Genre mode. Huot (1990) reviewed five studies that specifically examined the role of genre or discourse mode on the writing performance of secondary and post-secondary writers. Of the five studies he reviewed, three of the studies found syntactical differences (Crowhurst, 1980; Nold & Freedman, 1977; and Rosen, 1969) between

students' writing in various genres, and one study found differences in number of T-units and overall length (Rosen, 1969).

In Quellmalz, Capel, and Chou's (1982) study of 11th and 12th-grade students, the researchers examined the effects of genre on raters' perceptions of writing quality. Students were asked to write a narrative and expository response and were evaluated using five different criteria: general impression, focus, organization, support, and mechanics. Quellmalz et al., found that the levels of performance varied based on the different genre modes. As a result, they warned that educators should be hesitant to judge a student's writing ability based on one writing sample in a single genre or form.

Rhetorical specification. According to Huot (1990), rhetorical specification effect refers to the relationship between a specified audience or writing purpose in a writing prompt and the quality of scores given by raters. In his literature review on rhetorical specification, Huot identifies six studies that explore this topic (Brossell, 1983; Hult, 1987, Leu, Keech, Murphy & Kinzer, 1982; McAndrew, 1982; Puma, 1986; and Redd-Boyd & Slater, 1989). It is important to note that all studies focused on secondary and post-secondary students.

For example, Brosell (1983) found that when comparing three levels of rhetorical specification (low, moderate, and high), highly-specified writing prompts with elaborated audience and writing purpose goals yielded lower mean scores on writing quality and shortest mean length of essays relative to the low and medium specificity prompts. In contrast, Brosell found that moderately specified writing prompts yielded the highest mean quality scores and longest mean length when compared to the low and highly specified prompting conditions. He theorized that it may be possible that within a timed

testing situation, writing prompts that were highly specified may have been too demanding, but that some specificity on the writing purpose and audience may be beneficial for writers.

Puma (1986), explored the effect of audience specification differently. In his study, Puma examined the effect of audience intimacy or a writer's sense of distance from the audience (peer vs. superior), on their subsequent writing quality. A finding of note in this study was that Puma found an inverse relationship between a writer's sense of intimacy to their writing quality. More specifically, Puma found that the closer a writer felt to their audience (i.e. a peer rather than a superior) the more likely their writing would resemble spoken discourse, which yielded poorer quality essays.

Findings from these six studies were mixed. Of the studies Huot reviewed, four of the studies did not find significant differences in the quality ratings of essays based on varying levels of rhetorical specification (Hult, 1987; Leu et al., 1982; McAndrew, 1982; and Redd-Boyd & Slater, 1989). For the two studies with findings indicating differences in quality ratings relative to rhetorical specifications, the results were difficult to interpret.

Wording and structure of prompt. Another area that Huot (1990) explored in his literature review was studies that examined the effect of the wording and structure of the prompt on students' writing quality. Based on his review, Huot concluded that while structure, wording, and the presentation of writing prompts can at times have an effect on students' writing, the nature of this effect was still unclear. He cites four studies that support this conclusion. It is important to note, however, that none of these studies examined the writing of elementary aged students.

Three of the studies found no statistically significant differences in writing quality ratings based on variations in the wording and structure of the prompt (Brosell & Ash, 1984; Greenberg, 1981; and Hoetker & Brosell, 1989). Variations in the wording and structure of the prompt included combinations such as (a) personal questions and commands versus neutral questions and commands (Brosell & Ash, 1984); (b) varying levels of writing demand based on the level of structure requirements (low vs. high) in the writing prompt and the level of personal experience (low vs. high) required to respond to the prompt (Greenber, 1981); and (c) personal vs. impersonal prompts paired with a brief or extensive wording of the prompt related to paragraph length, voice, purpose, and audience (Hoetker & Brosell, 1989).

One study that did report differences in the effect of varying wording and structures of prompts was Smith et al.'s (1985) work. To examine the effects of wording and structure of prompts, these researchers presented students with three different prompting conditions: (a) writing a response to an open ended prompt, (b) reading and writing a response to one text, and (c) reading and writing a response to three texts. Another variable of interest in this study was the level of the writers' ability (basic=low, general=average, and advanced). When examining the students' writing quality under all three conditions, the advanced writers consistently outperformed the general and basic writers particularly with respect to the open-ended structure condition. It was interesting to note, however, that in the one-text response topic condition, general and basic students were closer in their writing quality performance and advanced writers were only slightly superior to basic writers. Furthermore, in the multiple-test response structure, general and advanced writers performed significantly better than basic writers. Smith et al concluded

40

that the wording and structures of prompts may have differential effects for varying levels of writing ability.

Implications for current study. Huot's (1990) literature review highlights many key issues to consider when designing writing assessments. Of particular interest is how variations in the writing assessment prompting condition can affect the task environment or overall demands for the writer. In Hayes' (1996) revised Cognitive Process Theory of writing, task environment plays a significant role in the writer's output. More specifically, Hayes notes that factors such as audience specification, texts read while writing, and the genre mode have an explicit link to text production. The findings from Huot's literature review supports Hayes' theoretical framework particularly with respect to the potential effect of genre, audience specification and texts read while writing. First, based on Huot's review, it is clear that genre mode has an impact on secondary and post-secondary writers' syntax and length (Rosen, 1969) as well as quality ratings (Quellmalz et al., 1982). In addition, it appears that moderately specified audience and structure goals have the potential to increase writing quality relative to more open-ended prompting conditions (Brosell, 1983, and Smith et al, 1985). Furthermore, writing prompts that specify a superior as opposed to peer audience may yield higher quality essays (Puma, 1986). It is important to note, however, that too much specificity with respect to audience and purpose in the writing prompt can have a negative effect on writing quality (Brosell, 1983). In addition, tasks that require a response to multiple texts may disadvantage basic writers relative to general and advanced writers more so than in open-ended and single-text response writing tasks (Smith et al., 1985).

This last finding with respect to open versus multiple-text response requirements is of particular importance given that the current study is situated within the new PARCC assessment context. In this new standardized testing context, elementary students will be transitioning away from the open-ended prompt, to the response to text prompt condition. Huot's literature review focused on studies that examined the writing of secondary and post-secondary writers. While these studies provide preliminary evidence on how writing prompts might differentially affect basic, general and advanced writers at the secondary and post-secondary levels, less is known on how writing prompts may affect elementary writers. Furthermore no studies exist to date that consider the effect of providing the read aloud accommodation to writers in a text response prompt condition.

Prior Knowledge. In Dochy, Segers, and Buehl's (1999) literature review of the relation between assessment practices, prior knowledge, and writing outcomes, they begin by presenting an operational definition of prior knowledge. They note that various terms for prior knowledge are often used interchangeably such as experiential knowledge, background knowledge, and personal knowledge. For their purposes, these authors use Dochy's (1994) definition of prior knowledge which includes "the whole of a person's actual knowledge that: (a) is available before a certain learning task, (b) is structured in schemata, (c) is declarative and procedural, (d) is partly explicit and partly tacit, and (e) is dynamic in nature" (p. 4699).

Dochy et al.'s (1999) review is divided into two sections. The first section focuses on research related to assessment of prior knowledge or the methods used to assess prior knowledge. They identify six categories for assessing prior knowledge: (a) multiple-choice tests (e.g., Chiang & Dunkel, 1992), (b) open questions / cloze tests / completion

tests, (c) association tests, (d) recognition tests (e.g., Chiesi, Spilich, & Voss, 1979; Hasselhorn & Korkel, 1986) / matching tests, (e) free recall (e.g., Lambiotte & Dansereau, 1992; Sanbonmatsu, Sansone, & Kardes, 1991), and (f) experimenter judgment (Heit, 1994) and self-estimation (i.e. familiarity ratings; Afflerbach, 1986). The second section of the review focuses on the effect of prior knowledge on performance. It is important to note that the primary focus of Dochy et al.'s review was to examine the effect of prior knowledge on performance generally. Of the 183 empirical studies they reviewed, none of the studies used writing ability as the performance or outcome variable. Instead, performance variables typically included: (a) reading comprehension tasks (e.g., Afflerbach, 1986, 1990; Brittton & Tesser, 1982; and Johnston & Pearson, 1982, Matthews, 1982), (b) conceptual knowledge (e.g., Brynes & Guthrie, 1992), (c) memory (e.g., Chiang & Dunkel, 1992; Clifton & Slowiaczek, 1981; Willoughby, Walker, Wood & MacKinnon, 1993; Willoughby, Wood, & Kahn, 1994), and (d) content knowledge (e.g., Lavore, 1989). In addition, the majority of the studies reviewed focused on secondary and post-secondary students with one study examining the work of 4th-grade readers and in one study examining the work of 8th-grade readers.

Dochy et al.'s (1999) review is useful in that it presents a broad conceptualization of the role that prior knowledge can have on learning and academic performance. The researchers arrived at four primary conclusions. First, there is a strong relationship between prior knowledge and performance. Dochy et al. noted that 91.5% of the studies reviewed reported positive effects of prior knowledge on performance and that prior knowledge generally explained between 30-60% of the variance in performance. Second, other learning variables related to prior knowledge are essential to performance such as

interest and the learners' personal beliefs. Third, the method of assessment influences the observed effect of prior knowledge on performance. More specifically, these researchers noted that the positive effect of prior knowledge is most apparent when objective methods such as multiple-choice tests and cloze activities are employed. Other methods such as self-assessment and self-estimation are less reliable. Finally, flawed assessment methods such as familiarity ratings, self-estimation, and matching tests of prior knowledge can still yield informative results. For example, they noted that some studies that used flawed assessment methods of prior knowledge showed other interesting implications such as that having no prior knowledge may be more beneficial than flawed prior knowledge (Schiefele, 1990; Lipson, 1982; Ceci, Caves & Howes, 1981; Alvermann, Smith, and Readance; 1985; Marshall, 1985; Neuman, 1989) and that when subjects have little or no prior knowledge than interest of the topic may play a greater role (Alexander et al., 1990; Garner & Gillingham, 1992).

While the studies in Dochy et al's (1999) review did not use writing ability as the outcome variable for evaluating the effect of prior knowledge on students' performance, the intersection of the findings from Dochy et al's review with Myhill's (2005) position paper provide some interesting insights into the potential role that prior knowledge may play in writing performance. In Myhill's (2005) position paper on the impact of prior knowledge on written genres produced in examination settings, she presents the view that children "whose home background has socioculturally prepared them for production of written genres are advantaged over those with different cultural and meaning-making resources available to them" (p.289). Moreover, Myhill contends that the task demands in

examination questions frequently fail to acknowledge how prior knowledge might impact student responses to the question.

In her paper, Myhill (2005) primarily cites data from a study which analyzed children's writing produced for examination purposes in England. In the English context, examination items typically provide students with prompts that identify a purpose, an audience, and a form of writing that may simulate a real-life situation (e.g., imagine that you are a reporter for a local newspaper). In her summary of student responses, she noted that children's prior knowledge of narrative appeared to be stronger than their expository text knowledge, and as a result, many students reverted to writing narrative essays in response to expository prompts. In a similar example, Myhill illustrates how children prompted to write a radio advertisement instead drew upon their knowledge of print and television advertisements. She argued that the misalignment of children's prior knowledge and the discourse goals of a radio advertisement where brevity and voice are prioritized led to writing samples that were overly elaborate and detailed.

Implications for current study. Myhill's (2005) article presents preliminary evidence to suggest that children's prior knowledge on genre may have an influence on the quality of their writing output. Unfortunately, other features of prior knowledge highlighted in Dochy et al's (1999) review such as content knowledge are not discussed within the writing context in Myhill's position review. Nevertheless, Dochy et al. (1999) and Myhill (2005) make a strong case to suggest that a writer's prior knowledge may have a significant impact on their performance.

For the purposes of the current study, Dochy and colleagues's review presents a foundation for considering how to assess prior knowledge and how these forms of

assessment may differentially affect performance outcomes. More specifically, Dochy et al. noted that objective measures of prior knowledge (e.g., multiple choice tests) were more reliable predictors of performance outcomes than flawed procedures such as student self-assessment.

Additional considerations highlighted by Dochy et al. (1999) include the potential role that student interest in, and student's misconceptions around the topic can play in their performance. Namely, higher interest levels can mitigate difficulties associated with having low levels of content knowledge on a topic. Additionally, student's misconceptions and misunderstandings of a topic can have the potential to negatively impact their performance. While Dochy et al.'s review did not specifically examine students' writing performance; it holds that if interests and misconceptions on a topic can have an effect in reading and content area subjects, it should also hold for writing as well.

Therefore when designing writing assessments for elementary aged children, it may be beneficial to evaluate students' prior knowledge on the topic and the genre they are being asked to produce in writing assessments. Evaluations of student's prior knowledge should allow researchers to determine if differences in levels of prior knowledge may be having an effect on students' overall writing performance.

Summary

This overview on critical issues in writing assessment presents a broad conceptualization of three areas of interest: (a) writing assessment measures, (b) writing assessment purposes, and (c) writing assessment forms. Each topic provides insight relevant to issues in the current study's larger context and rationale. This study seeks to

examine the effect of prompt condition and genre on elementary aged students from culturally, linguistically, socially, and academically diverse backgrounds.

The section that reviewed writing assessment measures illustrated the importance of using various procedures for evaluating student work including both overall quality indicators and sentence level skills. For example, analytic rubrics should be used to identify strengths and weaknesses in specific quality features of writing such as content, organization, and style. In addition, Huot's (1990) finding that raters using writing quality rubrics tend to focus on features of writing such as content and organization above word and sentence level features such as mechanics and grammar is important to note. Given the potential for raters to overlook these equally important features of writing, additional CBM measures such as total word count and correct word sequence will likely be useful in obtaining a comprehensive view of elementary children's writing development.

In order to effectively analyze and apply the above measures of writing, a key consideration must be the overall purpose of these assessments. Olinghouse and Santangelo (2010), and Gebhard and Harman (2011) note that standardized tests are often used for identifying at-risk writers, and in particular students with LD and students who are EL. Towards that end, a critical understanding of how these potentially marginalized populations write is necessary in order to accurately interpret and contextualize findings. Features of writing for students with LD and students who are EL that should be highlighted include challenges with transcription skills such as spelling and handwriting, and challenges with task environment factors such as the prompt and prior knowledge. An attempt should be made when designing writing assessments to moderate the potential

effect of transcription and the task environment so that students with LD and students who are EL are not unfairly disadvantaged over their general education peers.

One way to mitigate or conversely amplify the effect of prompt and prior knowledge is through the form of the assessment itself. Factors such as genre mode, rhetorical specification, and wording and structure of the prompt have been proven to have a significant impact on students' writing performance at the secondary and post-secondary level (Huot, 1990). In particular, Huot found that moderately specified audience and structure goals have the potential to increase writing quality relative to more open-ended prompting conditions (Brosell, 1983, and Smith et al, 1985), while writing prompts that had too much specificity with respect to audience and purpose can have a negative effect on writing quality (Brosell, 1983). Another important consideration Huot uncovered was that tasks that require a response to multiple texts may disadvantage basic writers relative to general and advanced writers more so than in open-ended and single-text response writing tasks (Smith et al., 1985). Furthermore, writing prompts that specify a superior as opposed to peer audience may yield higher quality essays (Puma, 1986).

Another consideration is the potential effect of prior knowledge particularly with respect to content and genre knowledge as highlighted in Dochy et al's (1999) review and Myhill's (2005) position paper. Key points for consideration here are the role that prior knowledge of content can play on students' academic performance. More specifically, higher levels of content knowledge are associated with higher levels of performance in areas such as reading (Dochy et al., 1999). In addition, it appears that middle school writers may be negatively affected by a lack of prior knowledge on expository genres

relative to narrative genres (Myhill, 2005). Given the above findings with respect to prompt conditions and prior knowledge effects particularly for content and genre at the secondary and post-secondary level, it would be interesting to see if these same prompt and prior knowledge effects held at the elementary level.

The content outlined in this section's overview on critical issues in writing assessment will be used to situate the goals of the current study within the broader context of writing assessment research generally. Using the above information, the empirical content and methodological review of this chapter have been narrowed to three specific areas related to writing assessment research that are most relevant to elementary aged writers: (a) the effect of prior knowledge, (b) the effect of prompt condition, and (c) the effect of genre.

<div align="center">

Empirical Content and Methodological Review

</div>

Methods and Search Procedures

I conducted an automated database cross search of Education Research Complete (EBSCO), Educational Resources Information Center (ERIC), JSTOR, PsycINFO, and the Social Sciences Index for studies related to writing in multiple genres at the elementary level, task environment, prior knowledge, and students who are EL and students with LD. Descriptors used were *writ*, elem*, prompt, genre, prior knowledge, task environment, and prompting condition.* This yielded 55 articles. I then conducted an ancestral search to locate additional studies cited in the reference lists obtained from the database search. This yielded an additional three articles for a total of 58. These included literature reviews, dissertation abstracts, qualitative, descriptive, correlational, experimental and non-experimental studies.

From this body of work, my focus for inclusion in this review was based on several criteria. First, I limited selection to studies that had been published in a refereed journal. While it is commonly accepted that papers that have been vetted through a peer-reviewed process are considered more credible, Troia (1999) added a caveat that most published studies "do not fulfill the accepted conventions of methodological rigor" (p. 31). Thus, further evaluation of published work is needed.

Second, only studies that employed quantitative analysis with independent variables related to elementary grade levels combined with each of the following focus areas were included: (a) prior knowledge, (b) prompt conditions, and (c) genre. I did not include studies that were qualitative or descriptive in nature. Due to the limited number of studies on children's writing development at the elementary level, I did not exclude studies based on the date of publication. This yielded ten studies that fell into three primary categories: (a) effects of prior knowledge, (b) effects of prompt condition, and (c) effects of genre. I present a content matrix of each study by these focus areas in Appendix A.

It is important to note why it seemed best to separate studies that specifically targeted the effect of prior knowledge, which includes topic knowledge and genre knowledge from studies that solely focused on the effect of genre. My decision to create these two separate categories was based on the fact that current research has explored the effect of genre on elementary children's writing both as a function of prior knowledge and as a topic of potential developmental differences within and across genres. No studies to date have considered children's prior knowledge and performance in multiple genres simultaneously. Furthermore, no studies have simultaneously examined the potential

interaction of prior knowledge of topic and genre, with the prompt condition. This gap in the extant research is ultimately what provides the rationale for the current study.

In addition to a content review, I also conducted a methodological review of these studies. According to Shadish, Cook and Campbell (2002), validity refers to the extent to which we can accept an inference to be true or correct based on evidence presented from a study. The extent to which researchers controlled for threats to validity influences the reliability and generalizability of their results and claims. I therefore evaluated the ten empirical studies in my review for: (a) internal validity; (b) construct validity; (c) statistical conclusion validity; and (d) external validity. After an initial review of the11studies, I chose to focus on 10 of the 37 validity threats highlighted by Shadish et al. (2002). The majority of the studies in this review were correlational and quasi-experimental. I selected these ten categories of validity criteria based on their relevance to this body of research. I operationally define each of the validity threats in Appendix B and provide a summary of methodological strengths and weaknesses by threat and study in Appendix C.

Results from the content and methodological review.

In the sections below, I will review these ten studies where the focus of assessment is on identifying students' performance in writing based on factors such as prior knowledge, prompt condition and genre. Of the ten studies reviewed, four of the studies placed primary emphasis on the effect of prior knowledge, three focused on the effects of prompt condition, and three examined the effects of genre on students' writing performance Within each of these studies, participants vary across elementary grade

levels and learner characteristics (e.g., EL and LD status). In addition, the types of genres used as the focus of assessment vary as well.

After presenting a review of both the content and methodological strengths and limitations of these studies' designs, I will briefly synthesize what researchers know about the writing of typically developing elementary students, students with LD and students who are EL. More specifically, I will summarize the role that prior knowledge, prompt condition, and genre plays in students' overall writing performance. I will also suggest ways in which assessments can be improved to better isolate the factors that may be contributing to students' differential performance both across and within the genres they are most frequently asked to produce in elementary school.

Effect of prior knowledge. In this portion of the literature review, I use a definition of prior knowledge that combines Dochy's (1994) definition described in the previous section of this chapter with elements of Hayes' (1996) cognitive process theory model. Here, prior knowledge refers to the students' content or topic, and discourse knowledge (including genre purpose and rhetorical structures) available to them before the writing task. Of the ten studies reviewed, four of the studies examined the effect of prior knowledge on the quality of students' writing performance.

One of the earliest studies to examine the effect of topic prior knowledge on the quality of students' written responses was DeGroff's (1987) study of 40 fourth grade students. Students were given a 49-item short answer test of baseball knowledge used in two previous studies (Mosenthal, 1984; Mosenthal et al., 1985). Of the 95 original students in the sample, 20 (17 boys and three girls) were identified as high knowledge students and 20 (15 girls and five boys) were identified as low knowledge students. In

addition to the prior knowledge assessment, children were administered the Comprehensive Tests of Basic Skills, 1981 (CTBS). The CTBS examined students' total language, mechanics and expression scores. No significant relationship was found between baseball knowledge and children's CTBS scores. This was done to ensure that students' writing would reflect their baseball knowledge rather than their writing ability.

Students were then asked to write stories about a baseball game. Stories were scored and analyzed using Voss, Vesonder, and Splich's (1980) baseball grammar for information about the goals of a baseball game. DeGroff (1987) found that students with higher prior knowledge on baseball included more goal related information while students with lower prior knowledge included more non-goal related information. In addition, high knowledge students' wrote longer pieces of writing that were more syntactically complex as measured y mean T-unit length relative to their low knowledge peers. Another measure of quality used by the researchers was a 4-point holistic score (1=very poor, 4=very good). These scores were generated by the 4th-grade students; in this study, high knowledge writers were more satisfied with their work than low knowledge writers.

This study represents one of the earliest attempts at assessing the effect of topic prior knowledge on elementary writers' compositions. Strengths of this study included the use of the CTBS score to control for factors such as students' individual writing abilities above and beyond their prior knowledge on the topic of baseball. In addition, the use of holistic, analytic (in the form of baseball grammar content), and sentence level (syntactic measure of mean t-unit length) writing measures strengthens the study's overall construct validity. Nevertheless, the lack of control for gender and other

individual factors introduces both selection bias and generalizability threats into the study. Furthermore, by using students' self-ratings on the holistic measures, statistical conclusion validity threats are not controlled for due to the unreliability of this type of self-assessment measurement.

Another area of prior knowledge that can have a significant effect on students' writing performance is their prior knowledge on the genre of the composition. Genre refers to the different modes of discourse that have a distinctive style, form or content (Kamberlis, 1999). The remaining three studies in this section focus on the effect of genre or discourse knowledge on students' performance in expository and narrative genres.

Englert, Raphael, Fear, and Anderson's (1988) study examined the role that students' metacognitive knowledge of expository writing played in 30, 4th and 5th-grade students' writing. Students were randomly selected from a pool of 260 students in the 4th and 5th-grades using a stratified random sampling procedure to select 10 students with LD, plus 10 high-achieving and 10 low-achieving students. Achievement levels were based on students' performance on the Stanford Achievement Test. To measure students' metacognitive knowledge about the writing process (i.e. students' knowledge of planning, editing, revising, and text organization), researchers conducted interviews using three vignettes centered on the writing problems of three hypothetical children where students were asked to give these children advice on how to write in each of four genres: (a) informational reports, (b) compare/contrast, (c) explanation, and (d) problem and solution. Interview responses were ranked as high, medium, low, and no knowledge (3-0 points).

Students were then asked to write an explanation and compare/contrast compositions in a counterbalanced design. These essays were scored using a primary trait score based on the extent to which the composition met the required organizational and text structures. For the explanation essays, there were four primary trait categories worth a total of 12 points and for the compare/contrast essay there were five categories worth 15 points. In addition to the primary trait analytic rubric, essays were scored using a holistic score from 0-3 points based on the degree to which the paper was interesting and effectively communicated (Englert et al., 1988).

Englert et al., (1988) found that students with LD were less aware than high-achieving students on the metacognitive process of expository writing. Furthermore, there was a strong positive relationship between students' writing performance and metacognitive strategies of organization, generating and monitoring expository writing. There were also strong positive correlations between the quality of students' writing and their awareness of writing strategies and what the researchers described as the students' internal awareness of when a paper was finished and why.

Strengths of this study included the random selection of participants, the counter-balancing of the explanation and compare/contrast genres, and the use of multiple text level measures of writing. Limitations of this study included a lack of sentence-level measures to represent the full spectrum of the construct of writing and the unreliability of the analytic primary trait measures that differed in points across the two genres of writing making them more difficult to compare. Nevertheless, this study represents one of the few studies that addresses the unique effect that prior knowledge can have on students

with LD, although it did not address the potential effects of task environment on students who are EL.

In contrast, Lee, Penfield, and Buxton's (2011) study specifically examined the effect of prior knowledge or schema on the writing of students who are EL. This study was part of a larger five-year intervention study that focused on English language development or English for speakers of other languages (ESOL) in a science curriculum. The goal of the intervention was to teach students who are EL science content with a focus on literacy strategies. Pre and post-test writing samples were collected for informational reports on the water cycle and students were asked to write in either English or their native language. These essays were then used to evaluate the relationship between students' knowledge of writing form and content relative to their EL background.

The research took place in a large urban school district over a three-year period in 3rd grade classrooms. There were 638, 661, and 676 participants in each of the three years respectively. Participants were evenly divided by gender and were predominantly Hispanic (46.4%-50.5%) and African American (45.9%-49.8%) with approximately 1% of students with disabilities in each of the three years. Students were additionally categorized based on their EL status in each of the three years with the least English proficient students representing 13.5% to 15% of the population, the moderately proficient students representing 38.7% to 53.3% of the population, and the proficient to native English speaking students representing 33.2% to 47.7% of the sample population across each of the three years of the study.

All student essays were scored with a form (conventions, organization, style/voice) and content (specific knowledge and understanding of science) rubric. Inter-rater agreement was established at 90%. Only a small number of students wrote in their home language of either Spanish or Haitian Creole, which were scored by research personnel that were fluent in these languages. Lee et al. (2011) used hierarchical linear modeling. Analyses were conducted on 2,020 students' essays. Approximately 37% of students were omitted from the sample because they were either missing a pre- or posttest. Omitted students had similar demographic variables for gender, ethnicity and EL status. The results indicated that there was a significant relationship between writing form and content and that this relationship was stronger for the post-test. Furthermore, this relationship was stronger for students who were not EL. The researchers suggested that this indicated that students with greater English proficiency learned science content and literacy skills simultaneously while students who are EL did not show simultaneous growth to the same degree. Lee et al. concluded that students who are EL required support not only in the form of writing they are being asked to produce, but also the content or the topic.

One of the greatest strengths of this study was its extremely large sample size. Based on this large sample size, Lee et al. (2011) was better able to control for selection bias and their results can certainly generalize to students who are EL and African American and Hispanic populations. In addition, their use of multiple raters and the establishment of inter-rater agreement at 90% strengthen the reliability of their measurements. Areas of their study that could have been improved upon are primarily in the measures they used. They used the same prompt for the pre and post-tests, which may

yield some testing effects. The use of the same prompt could also result in problems with respect to mono-operation bias. The prompt used in this study asked students to "imagine they were a water drop and part of the water cycle." This type of first person, personified approach to an informational report may have been a more challenging task environment. As such, it would have been nice to see if their study would have yielded similar results if students had been given multiple prompts on the same topic.

The last study in this section of the review focused on the relationship between 2nd and 4th-grade writers' discourse knowledge in the narrative genre and their writing performance. Olinghouse and Graham (2009) examined this relationship controlling for 4 writing (handwriting fluency, spelling, attitude toward writing, advanced planning), and 3 non-writing variables (grade, gender, and basic reading skills). In order to obtain a range of writing abilities in the study, participants were chosen using a stratified random sampling procedure using Verbal IQ on the Weschler Abbreviated Scale of Intelligence and the Test of Written Language III. As such, from an original sample of 32 second and 32 fourth-grade students a sample of 8 students below the 25th percentile, 16 students between the 25th and 75th percentile, and 8 students above the 75th percentile were selected.

Students were asked to write a story and then asked to respond to a series of questions on discourse and procedural knowledge of the narrative genre. Olinghouse and Graham (2009) found that five aspects of discourse knowledge made a unique and significant contribution to the prediction of story length, quality, and vocabulary diversity beyond the control variables. These included students' knowledge of substantiation, production, motivation, story elements, and irrelevant factors. In addition, they found that

4th-grade writers possessed more discourse knowledge than 2nd-grade writers and also wrote better quality essays. This supports the theoretical stance that greater knowledge of how to compose in a particular genre influences the quality of written output.

One important limitation was the sample size. In their regression analysis, Olinghouse and Graham (2009) used 12 predictor variables and because the sample size was so low, the reported values may be less than optimal population estimates. The authors also noted that they may have inadvertently created testing effects by biasing students' responses on the discourse knowledge questions immediately after administering the handwriting fluency test. They suggested that this may have accounted for students' focus on writing mechanics over broader text level features of writing.

Relative strengths of the study include the use of stratified random sampling and the controlling for variables such as handwriting fluency, which have proven to negatively bias scorers' quality ratings of students' texts (Graham, 1999). In addition, the use of multiple measures such as an analytic rubric, story length, and vocabulary diversity provide a fuller view of the construct of writing.

An additional strength of the study was that students' were provided three line drawing story prompts to select from. These prompts had been vetted in previous investigations as yielding similar writing performance with elementary students (Graham et al., 2005). In addition, advocates of examinee choice in writing assessment support this approach and suggest that choice allows students to select topics that are more familiar so that writers can develop a greater sense of ownership for their writing (Atwell, 1987). Nevertheless, findings on studies of examinee choice at the secondary level have been mixed. For example, Engelhard, Gordon, and Gabrielson (1992) found in a study of 8th-

grade students that writing tasks that required more personal responses (e.g., direct and imagined experiences) yielded essays that received higher quality ratings than writing tasks that required outside knowledge or were impersonal. In contrast, the same researchers in a study of 11th-grade students found that task choice had no substantive effect on the quality of essays (Gabrielson, et al, 1995). As such, it would be interesting to see if Olinghouse and Graham's (2009) study would have yielded similar findings if children had been provided a different task environment for writing.

   Implications for current study. The findings from the studies above present a number of important implications for my current research. First, it is clear that a student's prior knowledge on the topic and the genre of the writing task has an impact on their overall writing quality and sentence level features of writing (DeGroff, 1987; Englert et al., 1988; Lee et al, 2011; Olinghouse & Graham, 2009). In particular, it is important to note that the relationship between prior knowledge of both topic and genre relative to writing performance appeared to be stronger for students with LD (Englert et al., 1988) and students who are EL (Lee et al., 2011). Furthermore, Lee et al, concluded that students who are EL required support not only in the form of writing they are being asked to produce, but also the content and the topic. Nevertheless, none of the studies reviewed above examined the potential effect of varying the prompting condition with features such as content and rhetorical specification. Given the findings regarding the effect of genre and topic knowledge on students' performance in writing, it holds that additional supports in the prompting condition may help to moderate the disadvantages that struggling writers such as students with LD and students who are EL may have on writing tasks relative to their general education peers.

Effect of prompt condition. One way to mitigate the potential effects of prior knowledge on students' writing performance is through adaptations to the prompt condition. These adaptations can include audience, genre, and motivating cues. Of the ten studies reviewed, only three studies considered the effect of prompt condition on students' writing performance.

In Brodney, Reeves, and Kazelskis (1999) quasi-experimental study, 96, 5th-grade students in intact classrooms were assigned to one of four prompting conditions: (a) reading and pre-writing, (b) reading only, (c) pre-writing only, and (d) control. Students in all conditions first observed a video on volcanoes. They then wrote an expository essay on volcanoes for 30 minutes under one of the four prompting conditions listed above. For the first three prompting conditions, children were provided an additional 20 minutes of time for the reading and pre-writing, reading only, or pre-writing only conditions. In the control condition, children were not provided any additional time and were told to write using the information from the volcano video to help them. Outcome measures included a holistic measure, four analytic measures of ideas, style, organization, and mechanics, as well as total words per t-unit. Raw reading scores on the Stanford Achievement Test were used as a covariate measure.

Multivariate analysis of covariance showed that the type of pre-writing treatment significantly affected scores on students' expository compositions. More specifically, the reading paired with prewriting prompting task environment resulted in higher style, organization, and mechanics scores than the other prompting conditions. A finding of note, however, was that students in the comparison group wrote longer sentences that contained several ideas resulting in higher word per t-unit scores. Brodney et al. (1999)

noted that the comparison group sentences did not yield more cohesive essays. They found instead that students in the comparison group focused on sentence level ideation, whereas students in the reading and pre-writing condition focused on global paragraph level ideation. As such, the authors suggest that using words per t-units as a measure of quality may be restrictive and that "multiple assessments of student compositions, rather than the use of single indicators, provide a more comprehensive view of students' writing performance and achievement." (p. 18). Brodney et al. suggest that holistic and analytic measures provide useful information about different dimensions of writing performance, and that sentence level measures should not be used as the sole measure of writing performance.

The use of multiple measures was a relative strength of the study and helped to support the construct validity of their findings. Limitations that the authors noted included a focus on only one genre of writing. In addition, they noted that they did not include students' domain knowledge and interest in the topic, which could also be explanatory variables for students' performance. Another potential limitation is that participants represented one grade level in a single elementary school with a minority population of less than 10%. Adjusted group means were used and individual student ability based on the covariate of reading achievement as measured by the Stanford Achievement Test was not shared. This coupled with the use of intact classroom assignments to each of the treatment conditions presents both generalizability and selection bias concerns. Finally, there may also be instrumentation effects across the control and treatment conditions as the control students were not given 20 minutes of

additional preparation time before being asked to compose, which may give students in the treatment condition a performance advantage.

In Hudson, Lane, and Mercer's (2005) study, students were asked to write narrative stories under 6 different prompting conditions: (a) copying a story starter, (b) writing a dictated story starter, (c) discussing and then copying a story-starter, (d) discussing and then writing a dictation for a story starter, (e) discussing a story starter, and (f) no priming condition with just an assignment of topic. Results showed that writing produced under the discussion and topic conditions were lengthier than writing produced under the copying and discussion-copying conditions. Hudson et al., (2005) hypothesized that writing prompts may have hindered the students' writing fluency. The authors noted further limitations in their study. For example, students were not randomly assigned to conditions as they were in intact groups. Controls for prompt effects were not in place as the prompts changed with the priming condition. An additional problem may be that the handwriting demands in the copying conditions may have exhausted students prior to asking them to write independently.

Another approach to promoting the quality of students' essays is through a task environment that specifies audience awareness goals. According to Hayes and Flower (1980), awareness of audience and the intentional use of ideas to create written text for a specific audience is a hallmark of good writing. One study that did find conclusive differences in the writing quality of students' essays based on audience awareness prompts was Midgette, Haria and MacArthur's (2008) study of 5th and 8th-grade students. While this study evaluated the effects of a revising intervention, its findings have the potential to offer unique insight into another possible area of research with respect to

prompt effects. In their study, Midgette et al., asked students to revise their persuasive essays under three conditions: (a) a general goal, (b) a goal to improve content, and (c) a goal to improve content and communication with an audience. Under each condition, students were provided specific directions. General goals consisted of directives such as make any changes that you think would improve the essay. Content goals included prompts such as make sure that your opinion is clearly stated in your essay or think of other reasons to support your opinion, making sure to include at least three reasons. Audience goals included directions such as think about the people who might disagree with your opinion and think about how you would defend your opinion and show that they are wrong. Essays were then analyzed for overall persuasiveness using measures of persuasive discourse such as position, reason, and elaboration. Midgette et al., (2007) found that students in the audience goal group were more likely than both other groups to write more persuasive essays.

Relative strengths in the design of this study included random assignment to the three goal conditions using a matched-triads design, and identical writing prompts across the goal conditions. Limitations of their study however included mono-operation bias through the use of one-time assessment, and a narrow lens on the construct of writing due to their limited focus on holistic features of persuasive writing alone.

Implications for current study. Based on the three studies reviewed above, it is clear that variations in prompt conditions can have an impact on students' performance in expository (Brodney et al., 1999), narrative (Hudson et al., 2005), and persuasive (Midgette et al., 2007) genres. More specifically, features of the prompt condition such as varying the task (e.g., reading and pre-writing, reading only, pre-writing only) show

that combinations of pre-writing and reading tasks resulted in higher style, organization, and mechanics scores (Brodney et al., 1999). In addition, audience specification at the elementary and middle school level appear to improve students' writing performance (Midgette et al., 2007), which is in line with research at the secondary and post-secondary level (Huot, 1990). The above findings with respect to the potential benefits of including audience specification goals and reading and pre-writing tasks as part of the writing prompt condition were considered when designing the prompts for the current study.

  Effect of genre. Studies that focus on the development of writing abilities across genres take on many forms. What unifies theses studies are their independent variables. All studies included independent variables that are in some way related to learner characteristics such as grade level (K-7), age, gender, or disability. All studies also look at some form of genre (e.g., narrative, persuasive, informational reports, and poetry). Where these studies differ; however, is in what specific learner characteristics they examined, the types of genres that were the area of focus, and most importantly, their dependent variables. Dependent variables across studies varied from general language performance indicators (GLPM), to holistic scales, to measures of syntactic complexity. Given the degree of variability across these factors, it is difficult to gain a complete picture of how students' writing develops across the genres they are most frequently asked to produce in schools. Nevertheless, each of these studies does represent a part of the puzzle.

  In their studies of upper-elementary writers, both Beers and Nagy (2010) and Scott and Windsor (2000) focused on relatively narrow constructs of writing related to syntactic maturity and GLPM at the sentence level. While their focus on the construct of

writing was quite narrow, they did find that all students were affected by genre in similar ways.

Beers and Nagy's (2010) correlational study involved the collection of writing samples for the same two cohorts of students on two occasions, two years apart as part of a larger longitudinal study. Their participants included 83 students in grade three and then grade five, as well as 96 students in grades five and then seven. Students were asked to compose based on topic prompts related to geographic locations of the Pacific Northwest in four genres: narrative, descriptive, compare/contrast, and persuasive. Each text was then coded for length, clauses per t-unit or sentence, and words per clause.

Results showed significant effects for grade and genre on text length. Furthermore, they also found distinctions between syntactic complexities across the four different genres under review. More specifically, while children were able to differentiate between the genres broadly and were able to recognize the different purposes associated with each genre, they struggled to write effectively in them. This was particularly true for the compare/contrast and persuasive genres. For example, the compare/contrast text had a relatively low score for syntactic complexity in both clauses per T-unit and words per clause. In contrast, while children wrote persuasive essays with higher levels of clauses per unit relative to other genres, their number of words per clause were lower than descriptive texts. Beers and Nagy suggest that this finding was due to the simplified pattern that students often followed in persuasive texts where they would state their opinion and provide a justification for it.

One limitation of the study that Beers and Nagy (2010) offered was that while broad measures of syntactical maturity provided some information on distinctions across

genre, it did not necessarily provide insight into the extent to which students effectively employed genre-specific syntactical formats. As a result, the extent to which qualitative aspects of students' ability to effectively communicate their purpose in each genre from grade level to grade level was not captured.

A strength of Beers and Nagy's study was their ability to assess students on two occasions. This gives more credence to their findings on the effect of genre on performance as their findings were duplicated on both occasions of assessment. Even so, there are some threats to internal and construct validity that have the potential to weaken their claims. First, the same four assessments were used on both occasions, which could lead to testing effects as well as the threat of mono-operation bias. Second, grade level was used as a within subjects factor in the analysis where the samples from the two time points were analyzed within the individual cohorts as opposed to across cohorts. As such, factors such as history and maturation given the length of time between occasions were not adequately controlled. It would have been interesting to see if grade level differences would have been detected had this been a between subjects factor of analysis across the two groups. In addition, lack of random assignment due to their use of volunteer participants has the potential to lead to selection bias, and limits the ability for results to generalize to other individuals. Finally, while Beers and Nagy (2010) stated that they attempted to control for topic effects using the common theme of the Pacific Northwest where the study was conducted, the effect of topic and prompt cannot be isolated when the same four prompts were given to all students on both occasions.

The focus of Scott and Windsor's (2000) study was on the effects of narrative and informational prompts on the GLPI of students with and without disabilities for both

written and oral language tasks. Students were randomly assigned in matched-triads based on three distinct learner characteristics: learning disability, chronological age peers, and language-age peers. Students were then given parallel tasks and prompts in narrative and expository genres after watching videos and hearing similar directions. This was done to control for both background knowledge effects and prompt effects that may have influenced students' performance in each genre.

Scott and Windsor (2000) found that clauses per T-unit were not significant indicators of group differences. They did find that relative to chronological age peers, students with language based LD produced only 62% and 49% of the volume of narrative and expository summaries. In addition, they found significant group differences in the number of errors per T-unit. More specifically, they found that number of grammatical errors per T-unit for students with LD ranged from two to five times higher than chronological age peers in their study of 60 students between the ages of 8 and 12. Most importantly, Scott and Windsor also found main effects that favored narrative over expository contexts for GLPM.

Of the ten studies reviewed, Scott and Windsor (2000) met the requirements for 7 of 10 threats to validity. Moreover, there are a number of strengths to their study. First, Scott and Windsor were able to employ a matched-triad design that helped to control for some threats to selection bias. Second, by using the matched-triads design, Scott and Windsor model an effective method for meeting statistical assumptions of independence of observations for ANOVA designs. In addition, through their use of reliability checks, they moderately controlled for threats to statistical conclusion validity such as the reliability of their measures. An important strength of Scott and Windsor's (2000) study,

not found in many writing studies was the use of counterbalancing. The use of counterbalancing allowed for greater control of testing effects. In addition, through the use of the video priming conditions and parallel writing tasks, better control for background knowledge and prompt effects was exercised

By methodological standards, Scott and Windsor's (2000) results are likely the most reliable and generalizable. But their control over validity threats are countered with inevitable trade-offs in terms of other constructs such as generalizability. Scott and Windsor's study is very narrowly focused on the development of children's productivity or fluency and grammatical complexity across two genres. This narrow lens on the construct of writing does not sufficiently account for factors of writing in multiple genres such as content, organization, and style. Thus, while they were able to find main effects for genre that favored narrative over expository contexts, this finding is limited to GLPI and is not fully representative of how children would perform on all aspects of writing.

In Prater and Padia's (1983) study, gender, grade, and discourse topic served as the independent variables while the dependent variables included a four-point holistic scale. Participants from six elementary schools were purposefully selected to represent a mixture of urban/suburban settings and diversity in SES. This yielded 140 participants with half at the 4th-grade level and half at the 6th-grade level. Students were asked to write in response to three writing tasks within the same one-week period using a counter-balanced design across intact classrooms. Vocabulary for the task was controlled at no higher than third grade level using the EDL Core Vocabulary List. Writing tasks that specified purpose and audience, with several structuring sentences, were designed to elicit each of three types of writing: expressive, persuasive, and explanatory.

Prater and Padia (1983) found that for both genders, expressive writing tasks produced higher quality essays relative to explanatory and persuasive writing tasks. Persuasive writing tasks were found to be the most difficult type of writing for all subgroups with the exception of 4th-grade boys whose lowest performance was on the explanatory writing task. This finding is significant in that it is one of the few studies that examined genre and grade level differences that found an inconsistent trend. Expressive writing tasks involve the student writing about self, and according to Prater and Padia "may develop naturally with somewhat general instruction, (while) the other types of writing require direct, focused instruction" (p.150). The finding that 4th-grade writers performed better on explanatory writing than expressive writing somewhat contradicts this theory.

Prater and Padia's (1983) study has both strengths and limitations in their design. Their use of purposeful sampling allowed for better control over potential selection bias; however, their use of intact classrooms somewhat negated that. The use of 4-point holistic scales with multiple scorers to establish inter-rater reliability was helpful in establishing an effective means for comparing students' performance across genres and grade levels. In addition, by counter-balancing the sequence of the assessments across the six schools, some level of control was established for potential testing effects. Nevertheless, threats to validity that have been common across most studies exist: (a) mono-operation bias due to the use of one-time assessments; (b) too few constructs due to the lack of sentence-level measures; and (c) lack of control for background knowledge and topic effects.

Implications for the current study. The research summarized above provides a number of content and methodological insights into designing this study. First, while it is clear that genre has some effect on children's writing performance, the nature of that effect is not entirely clear due to variations in data collection procedures and dependent measures across each of the studies above. As such, future research on the effect of genre is warranted. Relative strengths in some of the studies above include the use of multiple assessments (Beers & Nagy, 2010), random assignment to conditions (Scott & Windsor, 2000), and counterbalancing the sequence of assessments in multiple genres (Prater & Padia, 1983). These strengths were taken into consideration when designing the current study.

Conclusions

Within the context of assessment, writing must be conceptualized as a broad integration of skills. While factors such as fluency and grammatical complexity should certainly be considered, other factors such as structure, content, mechanics and an understanding of audience and purpose are equally important within the context of children's writing development (Bereiter, 1980).

Based on the research reviewed above, there are a number of insights and implications that are relevant for the current study. First, it is clear that variations in prompt conditions can have an impact on students' performance in expository (Brodney et al., 1999), narrative (Hudson et al., 2005), and persuasive (Midgette et al., 2007) genres. More specifically, features of the prompt condition such as varying the task (e.g., reading and pre-writing, reading only, pre-writing only) show that combinations of pre-writing and reading tasks resulted in higher style, organization, and mechanics scores

(Brodney et al., 1999). In addition, audience specifications appear to improve students' writing performance (Midgette et al., 2007). I considered the above findings with respect to the potential benefits of including audience specification goals and reading and pre-writing tasks as part of the writing prompt condition when designing the prompts for the current study.

It is also clear that a student's prior knowledge on the topic and the genre of the writing task has an impact on their overall writing quality and sentence level features of writing (DeGroff, 1987; Englert et al., 1988; Lee et al, 2011; Olinghouse & Graham, 2009). In particular, the relationship between prior knowledge of both topic and genre relative to writing performance appeared to be stronger for students with LD (Englert et al., 1988) and students who are EL (Lee et al., 2011). Furthermore, Lee et al, concluded that students who are EL required support not only in the form of writing they are being asked to produce, but also the content and the topic. Given the findings regarding the effect of genre and topic knowledge on students' performance in writing, it holds that additional supports in the prompting condition may help to moderate the disadvantages that struggling writers such as students with LD and students who are EL may have on writing tasks relative to their general education peers.

Finally, relative strengths in some of the studies above include the use of multiple assessments (Beers & Nagy, 2010), random assignment to conditions (Scott & Windsor, 2000), and counterbalancing the sequence of assessments in multiple genres (Prater & Padia, 1983). These strengths were taken into consideration when designing the current study.

While the findings from the above studies provide preliminary evidence on how students' writing develops within and across genres, research is very limited with respect to how elementary aged students are affected by genre, prompt condition, and prior knowledge. In addition, findings to date are somewhat limited due to limits in control over internal and construct threats particularly in the areas of selection bias, lack of control for topic and prompt effects, mono-operation bias, and measures that inadequately explicate the entire construct of writing. These factors make it difficult to isolate the causes of the disparity between students' performance across the genres of writing they are most frequently asked to produce at the elementary level. For example, students' performance could be attributed to any number of factors including their individual abilities, differences in their knowledge of the form of the genre, students' background knowledge on the topic of the prompt, or the task demands of the prompt itself. These limitations present opportunities for future research that build on what has been learned thus far.

Discussion and Rationale for Study

Writing has long been considered a complex cognitive process rather than a body of knowledge or a specific set of skills (Hayes and Flower, 1980). As a result, it is no surprise that learning how to write is challenging for elementary aged children. Cognitive process theorists such as Hayes and Flower (1980), Bereiter and Scardamalia (1987), and McCutchen (1988) present a view on writing development that is contingent upon gaining competency in managing complex cognitive demands. These include the task environment, working memory, and general processes of writing such as transcription, planning, and composing. In this view on writing, multiple factors can have an impact on

a students' performance from their language background, to disability characteristics, to the form and the topic of the prompt.

Studies that examine students' writing development within genres most commonly associated with the culture and context of school, suggest that researchers' understanding of how students' writing develops across genres is incomplete. Preliminary evidence suggests that students' ability to write effectively in narrative genres develops more easily than expository genres (Beers & Nagy, 2010; and Scott & Windsor, 2000). Yet, Prater and Padia's (1983) study showed inconsistencies in students' performance across expository genres at the 4th-grade level. Anomalies such as this coupled with limitations in these studies' designs (e.g., mono-operation bias, lack of random assignment and selection bias, inadequate explications of construct, lack of control for topic effects, and generalizability of assessments) make it difficult to determine if differences in performances across genres can solely be associated with developmental features that make expository genres inherently more cognitively demanding than narrative genres.

To date, only three studies have empirically examined the effects of prompting conditions on students' performance within genres for elementary-aged writers (Brodney, Reeves, & Kazelskis, 1999; Hudson, Lane & Mercer, 2005; and Migette, Haria & MacArthur, 2007) yet none of these studies examined students' performance across genres. Based on the four studies that examined the effect of prior knowledge on students' writing performance it is clear that a students' prior knowledge on both the topic and the genre are an important area of consideration when designing assessments

and evaluating students' writing ability (DeGroff, 1987; Englert et al., 1988; Lee et al., 2011; Olinghouse & Graham, 2009).

      In order to gain a truly representative understanding of how students' perform across genres, an effort should be made to eliminate all known potentially biasing effects that may be invalidating results. Considering the following ideas may strengthen future writing assessment research. First, any future research should be grounded in a broader conceptual framework as it relates to children's writing development. This framework should include multiple constructs of writing and should not be limited to discrete skills. This will help researchers gain a better understanding of the integrative nature of the writing process and children's writing development at the elementary level. Second, in order to control for internal validity threats, one should attempt to employ randomization either through random assignment or through random sampling. In addition, if random or purposive sampling is available, external threats to validity such as generalizability to individuals may be achieved. Third, researchers should make a concerted effort to assess multiple genres using multiple methods of assessment. This will lend greater credibility to findings and help mitigate possible mono-operation bias and internal validity threats. In addition, a concerted effort should be made to control for topic and prompt effects through the varying prompt conditions across genres. While the form and topic of assessments in multiple genres will inherently have differences, the above recommendations will allow for greater control of any task environment effects.

      Decisions regarding measurement of writing development are important factors in this field of research. Multiple measures should be used to fully represent the construct of writing. This should include both word and sentence level measures (e.g., spelling,

grammar, word usage, vocabulary, punctuation) and overall quality measures (e.g., content, structure, and style). In particular, an effort should be made to analyze text level features individually through the use of analytic rubrics to help isolate potential developmental trends within quality features. To further promote measurement reliability, independent scorers, blind analysis, and inter-reliability checks should be employed. Finally, in order to better generalize to individuals who are most at-risk for school failure, studies should seek to examine the work of populations that include students with LD and students who are EL.

Through these methods, researchers may gain a better understanding of how children, and in particular how students with developing skills such as students with LD and students who are EL's overall writing develops. Furthermore, this research will enable educators to identify how children's writing develops uniquely across each of these sub-constructs in an integrated fashion. Information from assessment tools developed in studies that employ the above recommendations will give teachers what they need to provide targeted instruction for the purposes of intervention.

CHAPTER 3: METHODS

Overview

In this chapter, I discuss methodology relevant to this study. The methodological components are as follows: (a) design; (b) setting and participants; (c) independent, dependent, and control variables; (d) procedures; and (e) data analysis. This chapter also includes the study's expected outcomes, which are guided by the following research questions:

1. What is the effect of genre on 3rd- and 5th-grade students' overall writing quality and sentence level skills? Is the effect of genre similar or different across grade levels?

2. What is the effect of prompt condition (supported or unsupported) on 3rd- and 5th-grade students' overall writing quality and sentence level skills? Is the effect of prompt condition similar or different across grade levels?

Design

This study employed a quasi-experimental design. According to Kline (2009), "dealing with selection-related threats to internal validity in quasi-experimental designs is a major challenge." (p. 92). He suggests that in order to control for these threats, researchers should identify possible threats to the study as well as potential alternative explanations prior to enactment. Additionally, Kline suggests that design elements such as a pretest, additional measures, or group assignment can be used to better control for possible threats. He argues that with "sufficient controls, a quasi-experimental design can be a powerful tool for evaluating causal hypotheses" (p. 92). Accordingly, relevant design elements were used throughout this study.

In order to exercise sufficient controls over potential group differences, I first implemented a stratified random sampling procedure using relevant demographic characteristics to determine how the two independent variables (prompt condition and genre) under investigation vary with specific subgroups of students. After obtaining teacher and parent consent, I randomly assigned students into one of two prompting conditions: supported and unsupported, after matching for initial writing ability based on the Test of Early Written Language III (TEWL-3) and additional demographic data (EL status, disability status, reading proficiency level, and grade). Following random assignment, students were administered three writing prompts in narrative, persuasive, and informational report genres using a counter-balanced design. The following sections provide descriptive details on the setting and participants, independent, dependent, and control variables, procedures, and data analysis methods.

Setting and Participants

Sample

The final number of participants in this study included 63 3rd ($n=37$) and 5th grade ($n=26$) students from a public charter school (PK-6) in an urban school district in the mid-Atlantic.

Setting

The participants in this study belong to a culturally diverse school of 350 students: 47% of the students are African American, 44% are Hispanic, 9% are Caucasian, and 1% are Asian American. Approximately 11% of the population receives special education services and 83.5% of students are eligible for free or reduced-price lunches. In addition, at this particular school, 45% of the students came from homes in

which a language other than English (including Spanish, French, Amharic, Woolof, Arabic, Chinese, and Yoruba) was spoken. It is important to note, however that all students, regardless of being labeled EL, were considered by school personnel to be proficient in the English language.

At both the 3rd- and 5th-grade level, there are two classes each. All students at this school, including monolingual/native English speakers participate in either French or Spanish language immersion for half of their school week; in other words, parents chose for their children to attend the target school to take advantage of its dual language approach to instruction. It is important to note, however, that all students were expected to participate in their English language arts class in English, and that written language was a large component of instruction.

Recent demographic information from within the past two years indicated that the majority of students from this school were from low-income families (84%) and that information from the school district's Assessment and Accountability Data reported that 90% of the students tested at the school were considered "economically disadvantaged." These figures are considerably higher than the entire school district (66%).

Regardless of their socio-economic status, the school district's standardized reading assessment revealed that 58% of students at this charter school (during the 2010-2011 Academic Year) were identified as proficient readers compared to only 44% in the entire school district. Even more interesting was the difference in scores for African American students as 71% of students were identified as proficient readers compared to 39% in the district. This percentage is second only to a highly academic charter school where 77% of their students were identified as proficient readers.

Moreover, 42% of the Hispanic students at this charter school were identified as proficient readers. The percentage of proficient readers from this charter school (42%) is considerably higher than a bilingual charter school (30%) where 100% of the students are Hispanic. It is also important to note, however, that only 39 % of students identified as English Language learners (EL) were found to be proficient readers in this school. In addition, only 20% of students with disabilities at this school were identified as proficient readers

In sum, these descriptions indicate that students at the participating school were in some ways more proficient in terms of literacy than other students in the overall school district, but in other ways somewhat less so. In addition, while the vast majority of students at the target school were English learners, with the exception of students who were Hispanic, most were proficient in reading, according to district standards. Certainly, the overall population of students was typical of many schools both in the participating school district as well as in neighboring school districts. Thus, this school appears to be an appropriate setting for the current study.

This district has recently begun to adopt the Common Core standards. These standards include a significant focus on reading and writing in multiple genres. At the upper elementary grades, particular emphasis is placed on reading expository texts and writing opinion or argumentative essays (www.corestandards.org). Given this new focus on writing in multiple genres, the current study will assess the same genres targeted within the Common Core curriculum.

Recruitment and selection

In the fall of 2012, parents and guardians of 3rd and 5th-grade students received a permission form to allow their child to participate in the current study. The permission form stated that four writing samples as well as survey information would be collected from their child over the course of the school year. In addition, I requested permission to obtain their child's district assessment and demographic data from the school.

To encourage student participation and to maximize sample size, participants were offered an incentive for returning the permission form (with either a "yes" or "no") in the form of goodie bags that were distributed by their classroom teacher. Goodie bags included pencils, erasers, and a small amount of candy that were provided to all students who returned their permission form regardless of whether or not their parents agreed to their participation in the study.

Upon receipt of permission forms, I obtained demographic and district assessment data from the school's data manager for those students who had received permission to participate in the study. Data provided included students' age, race, first semester reading grades (below basic, basic, proficient, advanced), previous year's district reading levels (below basic, basic, proficient, advanced), disability status as identified by the presence or absence of an Individualized Education Plan (IEP), English Learner status, home language, and language level as determined by a district assessment (1-entering, 2-beginning, 3-developing, 4-expanding, 5-bridging, and 6-reaching). Teachers reported that all students identified as having IEPs received special education services for language-based learning disabilities.

Resulting sample

All students who agreed to participate in the study were initially included in the sample. This sample included 37, 3rd grade students and 33, 5th-grade students for a total of 70 students. Of the original 70 participants, 7 students failed to complete the study due to their inability to do make-up assessments following absences from school on days that data was collected resulting in their removal from the study and the loss of 3 matched pairs from the sample. All remaining students were kept in the study yielding a final sample of 63 participants.

The sample in this study is fairly representative of the school population with 65.1% or 41 general education students, 11.1% or 7 students with LD, 15.9% or 10 students who are EL, and 7.9% or 5 students identified as students with LD who are also considered to be EL. In contrast to the general population at this school, the sample has fewer students who are identified as EL and is comprised of mostly African American students (66.7%), with fewer students who are Hispanic (25.4%) or Caucasian (7.9%) and no Asian students.

Students' performance on the district wide assessment from the previous school year show that 42 of the students were reading at the proficient and advanced level with 18 students reading at the basic level and 3 students reading at the below basic level. Of the 17 students who were identified as either an EL or as a student who had LD and was EL, 11 students spoke Spanish as their first language, 3 students spoke French, 2 students spoke Ibo, 1 student spoke German, and 1 student spoke Yoruba at home. District assessment criteria for EL included the following categories: 1- entering; 2-beginning; 3-developing; 4-expanding; 5-bridging; and 6- reaching. According to this criteria, most

students receiving EL services at the participating school were identified as having expanding English language skills with 9 of these students categorized as having oral and written language with minimal phonological, syntactic or semantic errors that do not impede the overall meaning of the communication when presented with oral or written connected discourse with sensory, graphic or interactive support. Five students were identified as having oral or written communication in English that was comparable to English-proficient peers. The remaining 5 students in the sample were identified as performing at a developing level on the district English language assessment. This suggests that for these students, oral or written language may include phonological, syntactic or semantic errors that impede communication, but retain much of its meaning.

Independent, Dependent, and Control Variables

Based on the diverse population of the school in my study, care was taken to study the results in evaluating the effects of prompt condition and genre by considering a number of demographic factors including: (a) English-language learner status; (b) special education status; (c) reading proficiency level as determined by the most recent district assessment; and (d) grade. In addition, to control for initial writing ability, the Contextual Writing subtest of the Test of Early Written Language-III (TEWL-3) was administered to all students and used as a matching variable in addition to the demographic variables listed above. The TEWL-3 is a norm-referenced writing assessment that requires a student to "spontaneously produce a narrative story based on a picture stimulus" (TEWL-3, 2012, p. 3). The story is then evaluated based on its theme, dialogue, elaboration of detail, use of characters, vocabulary selection, and sentence structure.

There are three independent variables of central importance in this study. The first

independent variable of interest is grade level. The inclusion of grade level as a variable in this study was driven by a desire to see if there may be developmental differences in students' writing performance based on prompting conditions and genre by grade.

The second independent variable is genre, as students were asked to write narrative, persuasive, and informational report compositions. These genres were selected due to their alignment to the Common Core standards, and were used to determine if genre has an impact on students' writing performance.

The third independent variable is the prompting condition: supported and unsupported. After stratification and matching, students were randomly assigned to write compositions in either a supported or unsupported condition. Both conditions seek to mirror the task environment proposed in the new SBAC and PARCC assessments. Students were asked to read a passage, plan, and then write a response to the passage. In the unsupported condition students were asked to write in response to a passage and a genre specific checklist they were asked to read with minimal guidance and support (see Appendix D for examples). In contrast, in the supported condition all students were provided the read aloud accommodation for all texts in the prompt including the passage and the genre-specific planning checklist in order to minimize knowledge effects related to level of prior knowledge on topic and genre as well as to decrease the overall cognitive demands of the task environment. The use of the two different prompting conditions in the overall design will help determine the potential influence of the reading accommodation on a writing assessment given the new SBAC and PARCC writing assessment context.

Writing outcomes

Given the small sample size in my study, I selected two dependent measures of students' writing performance to represent students' writing ability in each of the genres and prompt conditions. Prior to obtaining my final sample size, I initially intended to use four outcome measures: (a) a holistic measure of writing quality, (b) an analytic rubric of writing quality; (c) a CBM measure that evaluates the number of correct-word-sequences (CWS) or grammatical errors in a sentence, and (d) a total word count measure. However, based on a G*Power 3.1 analysis for MANOVA models (Faul, Erdfelder, Buchner, & Lang, 2009), in order to get a medium effect size with 0.90 power and 4 outcome variables (i.e. holistic, analytic, CWS, and word count) a minimum sample size of 124 participants would have been required. Based on my goal of aligning to the new writing assessment context, I prioritized the analytic rubric over the holistic rubric as it best matched with the rubrics provided by the PARCC and SBAC consortiums. In addition, the holistic rubric was highly correlated with the analytic rubric ($r=0.68$ to 0.79, $p < 0.00$) suggesting a lack of independence, or failure to capture different aspects of writing quality. Of the two CBM measures that could be used to assess sentence level features of writing development (e.g., fluency) I chose the total word count measure over the CWS measure as it was more widely used in writing research to date (e.g., Beers & Nagy, 2010; Olinghouse & Graham, 2009, Hudson, Lane & Mercer, 2005; Scott & Windsor, 2000, etc.).

My final measures included: (a) an analytic rubric to assess quality features of writing individually, and (b) total word count as a measure of sentence level development. The four point analytic rubric was used for each of the following four text

level features: organization, content focus, style, and conventions yielding a final 16-point total (see Appendix F). In this context, organization refers to a logical progression of ideas or events. Content focus refers to the writer's ability to maintain focus on the topic or subject throughout the composition. Style consists of specific, developed details and skillful use of vocabulary that is precise and purposeful. Style also includes skillful sentence fluency such as varied length, good flow and rhythm, and varied structure. Lastly, convention refers to the writer's control over grammatical conventions appropriate to the writing task such as sentence formation, standard usage including agreement, tense, and case, and mechanics including use of capitalization, punctuation, and spelling.

In addition to a measure of analytic quality, I used one sentence level CBM measure of writing production: students' total word count. The use of both dependent measures provided a more comprehensive assessment of students' overall writing ability. It also provides researchers and teachers valuable information on students' strengths and needs. Based on the findings from these writing assessments, areas for intervention and instruction may be determined.

Procedures

I collected data during the fall and winter of the 2012-2013 school year. In the fall, following IRB approval, I distributed permission forms for all students and teachers at the 3rd and 5th-grade level. All writing assessments including the TEWL-3 and the narrative, persuasive, and informational report prompts were administered to all students as part of their regular classroom curriculum; however, I only collected and analyzed data from students for whom I had received permission.

Stratified Random Sampling

This study included four participating classrooms (i.e., two classrooms each at the 3rd and 5th grade levels). I randomly assigned students in matched pairs by classroom and grade level to either the supported or unsupported condition. Students were initially assigned in matched-pairs based on their performance on the TEWL-3, district reading assessment proficiency levels, as well as demographic characteristics, which included disability status, English learner status, and gender. See Table 1 for further information on students' performance on the TEWL-3 based on learner characteristics.

Table 1
Mean TEWL-3 Percentile Score (with standard deviations) by Age

	N	Mean
General Education	41	78.90 (23.56)
SLD	7	54.57 (25.07)
EL	10	70.50 (22.35)
Dually Exceptional	5	26.40 (28.62)
Total	63	70.70 (27.87)

Race was not used as a stratifying variable as the majority of students were African American and Latino. Unfortunately, due to attrition, four of the matched-pairs were lost in the final sample. See Table 2 for frequencies of demographic subgroups across each condition in the final sample of the study.

Table 2
Demographic Characteristics by Condition

	Supported	Unsupported
General Education Students	24	17
Students with LD	3	4
Students who are EL	4	6
Students who are EL with LD	2	3
Female	18	14
Male	15	16
Total	33	30

Nevertheless, while the original matched pairs across conditions changed in the final sample (*n*=63), there were no statistically significant differences in students' performance on the TEWL-3 (*p*=.211) with the students in the supported condition performing on average .15 SD above the mean (SD=.99) and students in the unsupported condition performing on average .16 SD below the mean (SD=1.00). Furthermore, the students' average reading performance level as determined by the district assessments were equivalent at .03 SD below the mean (SD=1.00) for both students in the supported and unsupported reading condition. This is significant in that the differences between the two conditions are that in the supported condition, students were read the prompt, while in the unsupported condition, students had to read the prompt independently.

Finally, because the school could not provide students' individual socio-economic status, this demographic variable was not under consideration in this study.

Prompts

I administered three writing prompts to the participants: narrative, informational, and persuasive. When designing these prompts, I considered factors that have a significant impact on students' writing performance. These factors included the genre mode, rhetorical specification, and wording and structure of the prompts as Huot (1990) found that moderately specified audience and structure goals have the potential to increase writing quality relative to more open-ended prompting conditions and because Brosell (1983) and Smith et al. (1985) found that writing prompts with too much specificity with respect to audience and purpose can have a negative effect on writing quality. In addition, Puma (1986) found that writing prompts with a specified superior as opposed to peer audience might yield higher quality essays.

In accordance with the above findings, all writing prompts included a specified audience and genre-specific structure goals (see Appendix D). For the narrative prompt, I directed students to write a story for a story-writing contest to be published in the "Mini Pages." a weekly kids insert that the children have seen in their local newspaper. For the informational report prompt, I directed students to write a report on pandas to be published by "Time for Kids" magazine. Finally, for the persuasive prompt, I directed students to write a persuasive article on whether or not kids should eat at Wendy's to be published by the "Kid's Post."

Based on findings from Huot (1990), I provided students genre specific structure goals in the form of checklists on pages for planning. Included on the planning pages in each prompt was a generic graphic organizer that followed an introduction, supporting paragraphs, and conclusion format from the Oregon State Department of Education's Writing Performance Assessment.

http://www.ode.state.or.us/teachlearn/subjects/elarts/writing/resources/wrgraphicorganizer1.pdf   These types of generic graphic organizers are made available to all students on most state and district assessments, and have been found to have a supportive effect in the writing process (Marzano, Pickering & Pollock, 2001). The specified audience cues were read to all students as part of the primary directions of the writing prompt. For the planning portion of the assessment, the general directions included a prompt to plan for five minutes. While all students received the written checklists in their assessment packets, only students in the supported condition were read the checklists. See Appendix E for scripted prompt directions for both the supported and unsupported writing conditions.

Huot (1990) also suggested that writing prompts that require a response to multiple texts may disadvantage basic writers relative to general and advanced writers more so than in open-ended and single-text response writing tasks (Smith et al., 1985). Dochy et al. (1999) and Myhill (2005) note that another consideration when designing writing prompts is the potential effect of prior knowledge of content and genre on students' performance. These authors noted that higher levels of content knowledge are associated with higher levels of performance in areas such as reading (Dochy et al., 1999). In addition, it appears that middle school writers may be negatively affected by a lack of prior knowledge on expository genres relative to narrative genres (Myhill, 2005). Given the above findings with respect to prompt conditions and prior knowledge effects, an effort was made to provide all students additional content and genre knowledge supports when writing in the expository genres tested in this study. This included the use of a single-text writing response task for both the informational report and persuasive writing prompts and the genre specific planning checklists. For the informational report prompt, I provided students with a text that included photos, a map, and a diagram about pandas. For the persuasive prompt, I provided students with a sample of the Wendy's menu as well as text that described possible reasons for and against eating at Wendy's from the perspective of the character, Wendy, who was for Wendy's, and from a doctor who was against Wendy's.

I differentiated texts for the 3rd and 5th grade levels using a Lexile analyzer ensuring that all texts fell into the Common Core specified Lexile range for each grade level (Grade 3=420L-820L and Grade 5=740L-1010L) http://www.lexile.com/using-lexile/lexile-measures-and-the-ccssi/text-complexity-grade-bands-and-lexile-ranges/. In

the supported condition, students were read all of the content on these pages of texts as well as the genre specific checklists in the planning section. In the unsupported condition, students independently read all text in the prompts.

In the narrative prompt condition, students were not provided a text to read, as prior content knowledge effects are less of a concern for narrative genres (Myhill, 2005). Instead, all students were provided a picture to write a response to (see Appendix D). In an effort to ensure that students wrote narrative stories that included story elements such as characters and a plot, I selected a picture that focused on a single character facing a problem; in this case a set of mystery paw prints.

For every writing prompt, students were provided 5 minutes to either examine the picture in the case of the narrative prompt, or read the passage in the case of the persuasive and informational report prompts. Students were then given 5 minutes to plan their writing. After the 5 minutes were up, students were given 15 minutes to write their narrative, persuasive or informational report pieces. The timing of the writing assessment was structured to mirror that of the TEWL-3, which was used as a matching and control variable at the start of the study. Furthermore, I presented all writing prompts and supporting materials to teachers for review and feedback prior to the start of the study. This allowed teachers the opportunity to give feedback on the topic of the prompts to ensure that the topics were not familiar, but also appropriately challenging.

Prompt Administration and Counter-balanced Design

There were a total of four participating classrooms in this study consisting of two 3rd grade and two 5th-grade classrooms. Matched pairs were created by classroom and across the grade level using the TEWL-3 and demographic variables. Students were

randomly assigned to one of two classrooms where the classroom teacher or I administered either the supported or unsupported prompt condition to the class on a rotating schedule to minimize any potential testing effects based on the administrator of the assessments. One assessment was given each week to each class over the course of a three to four week time period in the month of February depending on the class (See Table 3). While there were slight differences in the length of time between assessments for each class, these differences were minimal given that at this school, students receive instruction in English and writing composition for only half of the week as the other half of the week is spent in either a Spanish or French immersion setting.

All students were assessed in each of the three genres under review (narrative, persuasive, and informational report) in a counter-balanced design across grade levels (See Table 3). The use of counterbalancing allowed for greater control of any potential testing effects based on the type and sequence of genre that students were asked to write.

Table 3
Counterbalanced Test Administration by Grade and Condition

Genre Type and Administration Date			
3rd Grade Class I (n=20)	Persuasive 2-6-13	Informational 2-20-13	Narrative 2-28-13
3rd grade Class II (n=17)	Informational 2-5-13	Narrative 2-12-13	Persuasive 2-26-13
5th grade Class III (n=9)	Persuasive 2-7-13	Informational 2-21-13	Narrative 2-28-13
5th grade Class IV (n=17)	Informational 2-5-13	Narrative 2-12-13	Persuasive 2-19-13

Given unequal sample sizes, the variability in the time between assessments, and the counterbalanced sequence of the different genres across classes, I ran a MANOVA to assess the homoscedasticity of the dependent variables or the assumption that the dependent variables have equal variances across classes. This assumption is tested using Box's M Test of Equality of Covariances (*p>.001*). Box's M Test showed that for measures of writing quality and word count, the assumption that variances across classes were equal was met: (a) holistic scores (*p=.28);* (b) analytic scores (*p=.21*); and (c) word count scores (*p=.35*).

Data Collection and Scoring

After all assessments were collected, I generated student identification numbers to blind student essays to control for potential scorer bias. Handwriting has been shown to have a significant effect on scorers' rating of quality (Graham, 1999). To control for this effect, I typed all essays without corrections and used Microsoft Word's word count feature to identify the total number of words each student wrote for each essay.

In addition to length of essay, there was one other dependent variable of interest: an analytic quality measure using a16-point total rubric. I hired two outside raters to score the analytic rubric measures to eliminate any potential researcher bias from this measure in the study. The raters worked in pairs of two score the analytic measure. One of the raters was a doctoral student in special education, and the other was a master's level student in the same program.

Training. Raters were provided anchor and practice essays for training purposes from students whose parents permitted participation, but who did not complete the study. One of the 3rd grade teachers who participated in this study and I scored anchor essays for

the analytic measure to ensure that the analytic measure was aligned with grade level

expectations and construct and content validity threats were minimized.

Scoring procedures. Upon completion of the training, the pair of raters worked

together to score the student essays using the analytic rubric. There were a total of 63

participants in the study who each produced three essays for a total of 189 essays to be

scored.

I randomly selected 25% of the student essays for the pairs of raters to score

together to establish inter-rater reliability. I provided the analytic raters the same set of

scoring directions (see Appendix I). In general, pairs of raters were asked to

independently score essays in sets of five. After each set of five essays, raters were asked

to share their scores to identify inter-rater reliability. Before proceeding to the next set of

five essays, raters were asked to review their scores for any essays in which their scores

differed to help pairs of raters calibrate their scoring procedures so that they could

consistently rate the remaining essays according to procedures outlined by De La Paz

(1999). After establishing inter-rater reliability, all remaining essays were divided

between the pairs of scorers and scored separately. Inter-rater reliability for the analytic

essays were *r=.86.*

Data Analysis

The data analysis includes descriptive statistics, correlations, and inferential statistics

using Repeated Measures ANOVA, paired sample t-tests, and Multivariate Analysis of

Variance (MANOVA) procedures. The descriptive statistics indicate means and the

standard deviations of students' performance on the two dependent variables of interest in

this study including total word count and analytic rubric scores by each genre under

review (narrative, informational and persuasive). The initial relations among all measures

are shown through bivariate correlations. To address the first research question regarding

the effect of genre on 3rd and 5th-grade students, a series of Repeated Measures ANOVAS

were conducted to see if there are statistically significant mean differences in students'

performance on each of the genres. To determine the nature of these differences, I

conducted post-hoc contrasts using paired-sample t-tests. A Bonferroni correction to

alpha (a /15 = .003) was applied to control the Type-I error rate.

Finally, three MANOVA analyses are shared to explore the effects of prompt

condition, and grade level, on students' performance on the two dependent variables

mentioned above. The following section includes: (a) expected outcomes and

significance; (b) a description and rationale for the Repeated Measures ANOVA and

paired sample t-tests; (c) a description and rationale for the MANOVA methodology; (d)

MANOVA models; and (e) a description of anticipated outcomes.

Expected Outcomes and Significance

There are a number of hypotheses that have guided my research. My first

hypothesis is that in both conditions, students' performance in expository genres will not

be significantly lower and may even be higher than their performance in narrative genres

due to the use of supports such as genre specific cues and topic content in the form of

reading passages in both prompt conditions. Previous research suggests that children in

the elementary grades are more likely to perform better in narrative genres because

narrative genres are deemed less challenging and more familiar in the elementary grades

than expository genres (DeGroff, 1987; Englert et al.,1988; Lee et al., 2011; Olingouse &

Graham, 2009; Brodney et al., 1999; Hudson et al., 2005; Midgette et al., 2007). This

suggests that there may be developmental differences in students' performance across

genres. There is, however, competing evidence to suggest that the task environment and

the amount of prior knowledge students have on the topic and the genre may influence

the effect of genre on students' writing performance (Huot, 1999, and Myhill, 2005).

If my hypothesis proves true and children do not inherently perform better in

narrative genres than expository genres, this may suggest that given appropriate supports,

expository genres are equally accessible and manageable for elementary aged-students. If

students continue to perform significantly better in narrative genres despite the inclusion

of audience and genre specific cues as well as topic content supports in informational

report and persuasive writing, then my study will further support the belief that narrative

genres are developmentally more accessible to elementary-aged students. Alternatively, if

there are grade level differences in children's performance across genres, this might

imply that children's abilities to write across genres may be influenced by their overall

development or that it may be a matter of increased instruction and exposure to different

genres.

My second hypothesis is related to the effect of the reading accommodation or

supported versus unsupported condition. There is no research to date that examines the

effect of the reading accommodation on students' performance on writing assessments.

Smith et al.'s (1985) work at the college level, however, suggests that there may be a

relationship between a read and respond writing task environment and students' writing

performance based on different learner characteristics. More specifically, the authors

found that advanced writers consistently outperformed the general and basic writers

particularly with respect to the open-ended structure condition, but that general and basic

students were closer in their writing quality performance and advanced writers were only slightly superior to basic writers when given a one-text response condition. This finding supports my hypothesis that the use of a text to support topic knowledge for informational report and persuasive genres may be beneficial to all students, but specifically basic or developing writers such as those in grade 3.

Nevertheless, Smith et al.'s (1985) study looked at college level writers, which provides less insight into how developing readers' abilities to respond to a one-text writing response might differ at the elementary level. Furthermore, Smith et al. found that in the multiple-text response structure, basic college level students performed significantly worse than their general and advanced peers suggesting that as the cognitive demands for reading increased, students' performance in writing decreased specifically for struggling learners. Accordingly, my third hypothesis is that based on the above finding. It might hold that 3rd grade children or younger readers at the elementary level, even with the one-text response writing condition, may still struggle significantly more so than their 5th grade peers.

Repeated Measures ANOVA and Paired Sample *t-test*

To answer the first research question in this study related to the effect of genre on 3rd and 5th grade students' performance on overall writing quality and sentence level features I used a Repeated Measures ANOVA, combined with post-hoc paired sample *t*-tests. Repeated measures ANOVA is used when measuring an individual two or more times on the same dependent variables (Hinkle, Wiersma, and Jurs, 2003). In this study, I assessed students using the analytic and total word count measures for each of the three focus genres: narrative, informational, and persuasive. I also assessed the effect of genre

for each outcome measure by grade level. Using a standard ANOVA in this case is not appropriate because it fails to model the correlation between the repeated measures thereby violating the ANOVA assumption of independence. Repeated measures ANOVAs help to reduce potential error variance in cases where there is a great deal of variation between sample members where error variance estimates in traditional ANOVAs may be large.

There are several assumptions that underlie the use of repeated measures ANOVA (Hinkle, Wiersma, and Jurs, 2003). These include: (a) the sample was randomly selected from the population, (b) the dependent variable is normally distributed in the population, (c) the population variances for the test occasions are equal, and (d) the population correlation coefficients between pairs of test occasions scores are equal. In particular, if the last two assumptions are violated the Type I error rate can be seriously affected. Accordingly, I used Mauchly's test of sphericity (a measure that evaluates the variances of the differences between all possible pairs of groups) for each of the models to ensure that the population variances for the test occasions are equal.

In this study, I conducted repeated measures ANOVAs for each of the outcome variables to determine if the means for each outcome variable (analytic and total word count) differed significantly across each genre. Upon completion of this analysis, I conducted post-hoc contrasts using paired sample t-tests to determine the nature of these differences. A Bonferroni correction to alpha (a /15 = .003) was applied to control the Type-I error rate. Upon completion of the initial analysis of the complete data set, I separately ran repeated measures ANOVAs and post-hoc contrasts using paired sample t-tests by grade level.

Multivariate Analysis of Variance (MANOVA)

To answer the remaining research question in this study related to the effect of condition, and grade level on overall writing quality and sentence level skills, I used MANOVA. According to Mertler and Vannata (2005) "determining the appropriate statistical technique relies upon the identification of the type of variables (categorical or quantitative) and the number IVs (independent variables) and DVs (dependent variables) all of which influence the nature of the research questions being posed" (p. 20).

The general purpose of MANOVA models is to determine whether multiple levels of independent variables (IVs) on their own, or in combination with one another have an effect on multiple dependent variables (DVs). In contrast to Analysis of Variance (ANOVA) procedures, MANOVA is used to simultaneously study two or more related DVs while controlling for the correlations among the DVs.

MANOVA models are guided by certain assumptions that were examined in this study (Mertler & Vannata, 2005, p. 123). These assumptions are: (1) the observations within each sample must be randomly sampled and must be independent of each other; (2) The observations on all dependent variables must follow a multivariate normal distribution in each group; (3) The population covariance matrices for the dependent variables in each group must be equal (this assumption is often referred to as the homogeneity of covariance matrices assumption or the assumption of homoscedasticity); and (4) The relationships among all pairs of DVs for each cell in the data matrix must be linear.

In lieu of the homogeneity of variance used in ANOVA and *t*-test procedures, a MANOVA model examines the homogeneity of covariance or homostedacity. It is

assumed that the amount of variance within each group is comparable and that these results can be pooled to produce an error value that is representative of the population from which the groups in the sample are derived. Therefore it is also assumed that if there are large differences in error variance within each group, then the inferences generated from the estimated error measure for the model may be misleading.

The statistical power of any test is limited by a small sample size as a greater amount of variance will be attributed to error in smaller sample sizes, reducing the chances of significant findings. A value known as Box's M, given by most statistical programs, can be examined to determine whether the sample size is too small. Box's M determines whether the covariance in different groups is significantly different. If the difference in groups is significantly different, then it is presumed that the sample sizes in each cell are inadequate to make statistical inferences (Ho, 2006).

Generally speaking, MANOVA procedures are robust to moderate violations of normality and unequal sample sizes provided the model uses only a few DVs and a sample size of at least 20 in the smallest cell in the model (Mertler & Vannata, 2005). Given the relatively small sample size in this study (n=63), the limited number of participants in each sub-group under review (see Table 4), and the number of IVs and DVs of interest in this study, a MANOVA procedure was selected to minimize the potential for Type I error and incorrect statistical inferences. In addition, there are limited and uneven numbers in each learner characteristic subgroup (e.g., general education, students with LD, students who are EL, and students who are dually exceptional). As a result, I was unable to model these subgroup differences with any statistical significance.

Table 4
Number of Participants by Subgroup

Supported	33
Unsupported	30
Grade 3	37
Grade 5	26
General Education	41
SLD	7
EL	10
Dual	5

Alternative Procedures Considered and Rejected

Finally, it should be noted that one alternative analysis procedure was under consideration for this portion of the study. Multiple Regression Analysis (MRA) can also allow investigators to examine main effects and interaction effects by group for a single outcome variable. A limitation of MRA in this context is that it does not allow for the assessment of group differences. The independent variables of primary interest in this study were the effect of grade level and the effect of prompt condition on writing quality as measured by both sentence level (total word count) and overall writing quality (analytic quality) variables. Accordingly, a MANOVA procedure was deemed the best fitting method of analysis for this portion of the study relative to the targeted research questions of interest.

Post-hoc MRA analysis was attempted to determine if there might be statistically significant effects for condition and grade for non-general education students. Unfortunately, even when including all the outcomes for a collapsed learner characteristics group that included all students with LD, students who are EL, and students who are dually exceptional in the study, the sample size was still only 22. As a result, there was insufficient power to generate statistically significant effects using this model. Based on a G*Power 3.1 analysis (Faul, et al., 2009), in order to get a medium

effect size with 0.90 power and 3 predictor variables (i.e. grade, condition, and learner characteristic) a minimum sample size of 99 participants would have been required.

MANOVA Models

Variables. Like ANOVA, MANOVA examines the degree of variance within the independent variables and determines whether it is smaller than the degree of variance between the independent variables. If the within subjects variance is smaller than the between subjects variance it suggests that the independent variables have had a significant effect on the dependent variables. In this study, there were three independent variables of interest. These included grade level (3rd and 5th), and condition (supported and unsupported).

There were two categories of dependent variables of interest in this study: analytic quality, and total word count. It should be noted that initial writing ability and reading ability were controlled for in the stratified random sampling procedure through the use of the TEWL-3 and district reading assessment proficiency levels as matching variables.

My primary interests in this portion of the analysis were to determine if there were main effects of condition, and grade level and if there were additional interaction effects. To address these questions, I created three MANOVA models. Each model focused on a different genre, and all of the independent variables yielding three 3 x (genre) x 2 (outcome variables) x 2 (condition, grade) MANOVAs.

CHAPTER 4: RESULTS

Overview

This chapter presents the results of the data analysis conducted to assess the independent variables (genre, grade and condition) and dependent variables (analytic and word count measures) of this study. The two primary research questions that guide this study are:

1. What is the effect of genre on 3rd- and 5th-grade students' overall writing quality and sentence level skills? Is the effect of genre similar or different across grade levels?

2. What is the effect of prompt condition (supported or unsupported) on 3rd- and 5th-grade students' overall writing quality and sentence level skills? Is the effect of prompt condition similar or different across grade levels?

The results of the analysis in this study are organized into the following sections: (a) descriptive statistics and correlations, (b) effect of genre, (c) effect of condition and grade, and (d) conclusions. The first section presents descriptive statistics (i.e., means and standard deviations) of the raw scores for each of the writing quality outcome measures by independent variable subgroups as well as correlations to present an overview of the data and sample. The second section provides the results of the Repeated Measures ANOVA and paired sample *t*-tests that seek to answer the first research question in the study. The third section shares the results for each of the three MANOVA models presented in the previous chapter that seek to answer research question 2. In addition, this section presents the results of post-hoc analyses to determine if the effects found in the MANOVA models were a function of small sample size or if these effects were

statistically unique. This chapter concludes with a summary of results.

Descriptive Statistics and Correlations

Description of the Variables

The variables in this study include two categorical variables: (a) condition, and (b) grade. The dependent variables in this study include one ordinal measure of quality (the analytic rubric) and one continuous dependent measure (total word count). The former measure provides a total score up to 16 points that is based on four, 4-point rubric measures, assessing organization, style, conventions, and content. All outcome measures were used to assess each of the three genres of focus in this study: narrative, informational and persuasive.

Descriptive Statistics of Outcome Variables

The variables of interest in this study are the impact of genre, grade, and condition on students' performance in writing. Tables 5, 6, and 7 display means, and standard deviations for the outcome variables by condition and grade level and learner characteristics for each of the genres.

Table 5

Narrative Genre Means (and Standard Deviations) for Writing Outcome Measures by Condition, Grade, and Learner

Outcome	Condition	Grade	Learner Characteristic	Means(Standard Deviations)
Analytic	Supported	3	GenEd	9.75 (0.87)
			SLD	9.50 (3.53)
			EL	9.00(1.73)
			Dual	11.00 (--)
		5	GenEd	12.92 (2.07)
			SLD	10.00 (--)
			EL	10.00 (--)
			Dual	8.00 (--)
	Unsupported	3	GenEd	8.23 (3.00)
			SLD	7.50 (2.12)
			EL	7.67 (2.08)
			Dual	4.00 (--)
		5	GenEd	13.75 (2.23)
			SLD	10.00 (0.00)
			EL	10.66 (2.08)
			Dual	7.00 (0.00)
Word Count	Supported	3	GenEd	85.08 (33.02)
			SLD	63.00 (49.50)
			EL	96.67 (25.58)
			Dual	108.00 (--)
		5	GenEd	161.66 (72.07)
			SLD	97.00 (--)
			EL	126.00 (--)
			Dual	27.00 (--)
	Unsupported	3	GenEd	75.69 (44.50)
			SLD	96.50 (85.56)
			EL	62.00 (25.63)
			Dual	26.00 (--)
		5	GenEd	167.75 (27.26)
			SLD	52.50 (14.85)
			EL	183.00 (12.17)
			Dual	27.00 (9.90)

GenEd = general education students, SLD = students with LD, EL = students who are EL
Dual = students who are EL with LD

Table 6

Informational Genre Means (and Standard Deviations) for Writing Measures by Condition, Grade and Learner

Outcome	Condition	Grade	Learner Characteristic	Means(Standard Deviations)
Analytic	Supported	3	GenEd	11.58 (1.38)
			SLD	9.00 (4.24)
			EL	10.67 (2.08)
			Dual	13.00 (--)
		5	GenEd	13.25 (1.48)
			SLD	15.00 (--)
			EL	13.00(--)
			Dual	4.00(--)
	Unsupported	3	GenEd	9.15 (3.31)
			SLD	8.00 (2.83)
			EL	9.00 (4.36)
			Dual	5.00(--)
		5	GenEd	14.75 (1.50)
			SLD	12.00 (1.41)
			EL	12.67 (0.58)
			Dual	11.00 (5.66)
Word Count	Supported	3	GenEd	10.47 (3.83)
			SLD	10.00 (2.94)
			EL	10.83 (3.43)
			Dual	9.00 (5.29)
		5	GenEd	66.83 (22.98)
			SLD	50.50 (33.23)
			EL	43.00 (14.73)
			Dual	47.00(--)
	Unsupported	3	GenEd	127.83 (50.51)
			SLD	166.00(--)
			EL	165.00(--)
			Dual	18.00(--)
		5	GenEd	58.92 (29.80)
			SLD	54.00 (26.87)
			EL	60.00 (50.11)
			Dual	37.00(--)

GenEd = general education students, SLD = students with LD, EL = students who are EL
Dual = students who are EL with LD

Table 7

Persuasive Genre Means (and Standard Deviations) for Writing Outcome Measures by Condition, Grade, and Learner

Outcome	Condition	Grade	Learner Characteristic	Means(Standard Deviations)
Analytic	Supported	3	GenEd	9.83 (1.99)
			SLD	6.00 (1.41)
			EL	10.33 (0.58)
			Dual	12.00 (--)
		5	GenEd	12.17 (1.28)
			SLD	13.00 (--)
			EL	9.00 (--)
			Dual	6.00 (--)
	Unsupported	3	GenEd	9.38 (2.72)
			SLD	8.00 (1.41)
			EL	9.33 (0.58)
			Dual	7.00 (--)
		5	GenEd	14.00 (1.15)
			SLD	10.50 (0.71)
			EL	12.33 (1.53)
			Dual	11.50 (0.71)
Word Count	Supported	3	GenEd	73.08 (24.78)
			SLD	52.00 (49.50)
			EL	65.67 (40.00)
			Dual	52.00 (--)
		5	GenEd	125.8333
			SLD	143.00 (--)
			EL	111.00 (--)
			Dual	30.00 (--)
	Unsupported	3	GenEd	61.23 (39.23)
			SLD	47.50 (6.36)
			EL	41.67 (24.79)
			Dual	63.00 (--)
		5	GenEd	158.25 (53.01)
			SLD	77.00 (29.70)
			EL	141.67 (25.32)
			Dual	53.50 (2.12)

GenEd = general education students, SLD = students with LD, EL = students who are EL
Dual = students who are EL with LD

Initially, I was interested in examining whether the effect of genre and condition

was similar or different for students across grade levels and by learner characteristics

(e.g., students with LD, students who are EL, and students who are dually exceptional relative to their general education peers). Unfortunately, due to the small sample sizes for each of the subgroups that include these populations (students with LD = 7, students who are EL = 10, and students who are dually exceptional = 5) I am unable to use inferential statistics to model any potential learner characteristic group differences. Furthermore, as demonstrated in the tables above, in many instances, there was only one participant in each of these learner categories for each grade level and condition. Based on this, the variable of learner characteristic was removed from subsequent analyses, as any significant results would have little educational significance.

Correlations

Bivariate correlations among the MANOVA variables (condition, grade level, total word count and analytic writing quality) are presented in Table 8.

Table 8
Correlations among MANOVA variables

	1	2	3	4	5	6	7	8
1.ANANARR	-							
2.ANAINFO	$.575^{**}$	-						
3.ANAPERS	$.653^{**}$	$.562^{**}$	-					
4.WCNARR	$.602^{**}$	$.454^{**}$	$.589^{**}$	-				
5.WCINFO	$.461^{**}$	$.577^{**}$	$.477^{**}$	$.774^{**}$	-			
6.WCPERS	$.509^{**}$	$.414^{**}$	$.574^{**}$	$.791^{**}$	$.826^{**}$	-		
7.CONDITION	$-.319^{*}$	$-.252^{*}$	-.037	-.171	-.157	-.114	-	
8.GRADE	$.512^{**}$	$.454^{**}$	$.522^{**}$	$.466^{**}$	$.575^{**}$	$.552^{**}$	-.089	-

ANA=Analytic, WC=Word Count, NARR=Narrative, INFO=Informational, PERS = Persuasive
**Correlation is significant at the .01 level (two-tailed)
*Correlation is significant at the .05 level (two-tailed)

As expected, the outcome variables of total word count and analytic writing quality are significantly related to one another given the recognized relationship between writing quantity and writing quality (Graham, Berninger, Abbott, Abbott, & Whitaker, 1997; Rankin, Bruning, & Timme,1994). The condition variable was moderately correlated with the analytic narrative measure, and weakly correlated with the analytic informational measure. In addition, grade was significantly related to all outcome measures.

Effect of Genre

I conducted a series of repeated measures ANOVAs to answer the first research question: what are the effects of genre on 3rd and 5th-grade students' overall writing quality and sentence level skills? The purpose of these tests was to determine if there were statistically significant differences between the means for each genre and for each of the writing measures used (analytic and total word count) in this study.

Assumptions

Repeated measures ANOVA tests follow certain assumptions (Hinkle, Wiersma, and Jurs, 2003). These include: (a) the sample was randomly selected from the population, (b) the dependent variable is normally distributed in the population, (c) the population variances for the test occasions are equal, and (d) the population correlation coefficients between pairs of test occasions scores are equal. In particular, if the last two assumptions are violated the Type I error rate can be seriously affected. Accordingly, I used Mauchy's test of Sphericity (a measure that evaluates the variances of the differences between all possible pairs of groups) for each of the models to ensure that the population variances for the test occasions are equal.

Results

I first assessed the effect of genre as measured by the analytic outcome variable, a measure of writing organization, content, style, and structure. Mauchy's Test of Sphericity indicated that the assumption of sphericity had not been violated, $\chi^2(2) = 3.74$, p = .16. The repeated measures ANOVA determined that the mean analytic scores differed significantly between genres $F(2, 124)=6.94$, $p=0.00$. Post hoc paired sample *t*-tests using the Bonferroni correction revealed that students ability to write using the analytic quality measures in the informational genre over the narrative genre (11.52 vs. 9.95, respectively), which was statistically significant ($p=.00$). However, there were no statistically significant differences ($p=0.30$) between the narrative and persuasive genres (9.95 vs. 10.43, respectively). Finally, there were no statistically significant differences ($p=0.11$) between the informational and persuasive genres (11.52 vs. 10.43, respectively). In effect, based on the analytic measure, students appeared to write best in the informational report genre with less success with overall organization, content, style and conventions in the persuasive and narrative genres.

I next assessed the effect of genre as measured by the total word count variable. Mauchly's Test of Sphericity indicated that the assumption of sphericity had been violated, $\chi^2(2) = 6.05$, p = 0.05. To address this issue, I performed a Greehouse-Geisser procedure, which corrects the degrees of freedom of the *F*-distribution. As such, the *F*-test result is corrected from *F* (2, 10) = 12.53, *p=0.001* to *F* (1.28, 6.38) = 12.53, *p=0.0001*. This correction elicits a more accurate significance value. By increasing the *p*-value I can compensate for the fact that the repeated measures ANOVA test is too liberal when sphericity is violated (Howell, 2002). With the Greehouse-Geisser correction, the mean total word counts differed significantly between genres *F* (1.83, 113.03) = 12.09,

p=0.00. Post hoc paired sample *t*-tests using Bonferroni corrections revealed that students wrote more in the narrative genre over the informational genre (103.22 vs. 82.40, respectively), which was statistically significant (*p*=0.00). However, there were no statistically significant differences (*p*=1.00) between the informational and persuasive genres (82.40 vs. 86.06, respectively). Finally, there were statistically significant differences (*p*=0.02) between the narrative and persuasive genres (103.22 vs. 86.06, respectively). In short, students appeared to write the most in the narrative genre followed by the persuasive and informational genre.

The between subject effects for grade and word count were statistically significant. In order to determine the nature of these effects by each genre I ran additional repeated measures ANOVAS for the writing quality (analytic) and writing quantity (total word count) variables for third and fifth grade separately. At the 3rd grade level, Mauchy's Test of Sphericity was not violated for the analytic outcome variable $\chi^2(2) = 5.98$, p = 0.05. The repeated measures ANOVA determined that the mean analytic scores differed significantly between genres $F(2, 72)=3.71$, *p*=0.02. Post hoc paired sample *t*-tests using the Bonferroni correction revealed that 3rd grade students' ability to write using organization, content, style and conventions was greater in the informational genre over the narrative genre (9.97 vs. 8.73, *p*= 0.04) while there were no statistically significant differences between the narrative and persuasive and persuasive and informational genres. For the word count variable, Mauchy's Test of Sphericity was violated at, $\chi^2(2) = 6.13$, p = 0.04. I therefore used a Greenhouse-Geiser correction yielding statistically significant differences in word count between genres $F(1.72, 72)=7.26$, *p* = 0.00. Post hoc paired sample *t*-tests using the Bonferroni correction show

statistically significant mean differences across all genres ($p < 0.05$) with 3rd grade

students writing more in the narrative genre (M=79.30), followed by the persuasive

(M=62.41), and informational genres (M= 58.65).

At the 5th grade level, Mauchy's Test of Sphericity was not violated for the

analytic outcome variable $\chi^2(2) = 3.73$, p = 0.16. Results of the repeated measures

ANOVA for 5th grade analytic writing quality show no statistically significant differences

between genres. For the word count variable, Mauchy's Test of Sphericity was not

violated at the 5th grade level, $\chi^2(2) = 1.51$, p = 0.47. In addition, there were statistically

significant differences between genres for the word count measure at the 5th grade level *F*

(2, 50) =4.68, *p* = 0.01. Post hoc paired sample *t*-tests using the Bonferroni correction

show statistically significant mean differences (*p*= 0.05) between the narrative

(M=137.27) and informational genres only (M=116.19) favoring the narrative genres.

There were no statistically significant mean word count differences between the

informational and persuasive genres at the 5th grade level.

These findings contradict previous research that suggests that writing fluency

skills are predictive of individual differences in writing quality (Graham, Berninger,

Abbott, Abbott, & Whitaker, 1997; Rankin, Brunning, Timme, & Katkanant, 1993).

Furthermore, in past studies the effect of genre has shown that students appear to write

more and achieve better quality essays in narrative genres relative to expository genres

(e.g., Olinghouse & Wilson, 2012; and Camp, 1993). In this study, children did write

more (*p*=.00) in the narrative genre (M=103.22 words) relative to the informational

report (M=82.40), and persuasive genres (M = 86.06) overall and in both grade levels.

However, in contrast, on the analytic measure 3rd grade students on average seemed to

perform best in the informational report genre with 5th grade students demonstrating no statistically significant differences in their ability to write using organization, style, content, and conventions across the genres.

One possible explanation for students' stronger performance in the expository genres relative to previous studies could be a function of the task environment and prompt structure in the current study. I hypothesized that given the research recommended supports in the expository genres (e.g., specified superior audience, genre specific cues, graphic organizers, and topic and content supports through the provision of a text) the effect of genre or more specifically genre knowledge and topic knowledge may be minimized thereby making the expository genres more accessible to elementary writers. The results of this study with respect to the analytic score results support this hypothesis.

Nevertheless, there was still a difference favoring the narrative genre for the word count measure. A possible explanation for the difference between students' relative performance in each genre by total word count versus analytic measures could be attributed to how raters weighted each of the categories on the analytic measure. For the analytic outcome measure, I used a collapsed analytic variable that combined four 4-point features of writing (organization, content, style and structure) to create a 16-point total analytic writing measure. In Table 9, I share the means for each of these features by genre.

Table 9
Analytic Rubric Means by Category Features

	Organization	Content	Style	Conventions
Narrative	2.39	2.63	2.29	2.63

Informational	2.38	3.20	2.59	2.95
Persuasive	2.21	3.19	2.21	2.69

While the analytic scores for organization seem to be relatively equivalent, both the informational report and persuasive genres appeared to receive much higher ratings for content than the narrative genre. Graham (1990) argued that spelling words and writing letters might interfere with other aspects of composition such as planning, organization, and content generation. In the data above it seems that organization is a relative constant across genres while content and conventions are higher for the informational and persuasive genres. This finding further supports the contention that providing students with content in the form of texts read prior to writing in the expository genres can mitigate the demands of content generation and help aid in areas of conventions such as spelling, thereby yielding essays that may be shorter in length, but higher in quality. In short, although perhaps tempered by a small sample size in the current study, prior results suggesting a better performance on narrative over expository genres only occurred for the length of the students' writing. This suggests that students may not write better in narrative than expository genres.

One concern, however, that is highlighted by the above data across the genres and analytic categories is that there were only two instances when average writing quality was above a 3.0 or effective range with most students performing in the developing range. Of the lowest performing students (total analytic score < 8.0 or an average of < 2.0 across each of the four analytic categories: organization, content, style, and conventions) 52% were general education students, 16% were students with

learning disabilities, 11% were students who are EL, and 20% were dually exceptional

students. Given that the representation of these sub-groups in the overall sample is 65%

(general education), 11% (students with LD), 16% (students who are EL), and 8%

(dually exceptional students), the overrepresentation of students with LD and students

who are dually exceptional in this low performing category is concerning.

<div align="center">**Effect of Condition and Grade Level**</div>

In the following section I share the results of three MANOVA models I used to

identify the effects for the independent variables (condition and grade) in relationship to

the outcome variables. These models seek to answer the remaining research question in

the study that relates to the effect of prompt condition and grade level, on students'

overall writing quality and sentence level performance. In the following section, I will

share the results of the MANOVA assumptions testing I conducted for the data in the

study. I will then share each of the results for each of the dependent variables I modeled.

Assumptions

Tests for the four general assumptions for MANOVA procedures were examined

by pre-screening the data to confirm its robustness to statistical analyses.

Independence. In this study, independence was established through the use of a

stratified random sampling procedure.

Normality Visual examination of a series of histograms for each dependent

measure demonstrated that each was normally distributed.

Given the limited sample size in each group for this study, I conducted a

Mahalanobis' D^2 analysis to check multivariate normality for each of the groups of

dependent variables in the four MANOVA models. This measure is a multidimensional

version of a z-score and it provides information regarding the distance of a case from the multidimensional mean of a distribution, given the covariance of the distribution. For each of the three models' dependent variables there were no significant outliers that would violate the multivariate normality assumption.

      Linearity. Another assumption of the MANOVA procedure is that there are linear relationships among all pairs of dependent variables. This assumption was examined through bivariate correlations. Based on the bivariate correlation statistics, all models have DVs that are significantly correlated ($p<.001$ or $p<.05$).

      Homoscedasticity. The last assumption of MANOVA is that population covariance matrices for the DVs in each group must be equal or in other words the models should have multivariate normality. This assumption is assessed through Box's M statistic using an alpha value of 0.00. The assumption of homoscedasticity was met for all outcome variables in this study: (a) analytic $p=0.10$; and (b) word count $p= 0.00$.

      Procedures for analysis. In the following sections I share the results for each of the three MANOVA models I created to evaluate the effect of prompt condition and grade on each of the dependent variables. I first confirmed multivariate normality for each model. If multivariate significance was found for a variable, I then interpreted the univariate ANOVA results to determine significant group differences for each dependent variable (Mertler & Vanatta, 2005).

Results

      As expected there were statistically significant differences favoring 5th-grade students over 3rd-grade students across all genres and outcome measures: analytic

narrative, $F (1, 59) = 21.34$, $p < 0.00$, (Cohen's $d = 0.27$); analytic informational, $F (1, 59)$ $= 16.86$, $p < 0.00$, (Cohen's $d = 0.22$); analytic persuasive, $F (1, 59) = 23.40$, $p < .0.00$, (Cohen's $d = 0.28$); word count narrative, $F (1, 59) = 15.52$, $p < 0.00$, (Cohen's $d = 0.21$); word count informational, $F (1, 59) = 27.83$, $p < .001$, (Cohen's $d = 0.32$); and word count persuasive, $F (1, 59) = 25.67$, $p < 0.00$, (Cohen's $d = 0.30$). For the analytic informational variable, there was also a statistically significant interaction effect for grade level and condition, $F (2, 58) = 3.93$, $p < 0.03$, (Cohen's $d = 0.12$). Table 11 shows means and standard deviations for writing outcome measures by condition and grade level for the analytic informational outcome variable.

As evidenced in Table 10, 3rd-grade students performed significantly better in the supported condition with respect to the organization, content, style and structure of their essays (ES = 0.90). In contrast, statistical results indicated that students in 5th-grade students performed slightly better in the unsupported condition (ES = 0.12).

Table 10
Means (and Standard Deviations) for Analytic Informational Measure by Condition and Grade Level

CONDITION		GRADE	Means (Standard Deviations)
Analytic Unsupported		3	8.79 (3.29)
		5	13.00 (2.53)
	Supported	3	11.22 (1.93)
		5	12.73 (2.79)

Post-hoc Analysis

To determine if this result was an effect of small sample size, I performed a post-hoc analysis to determine if there were any significant outliers that may be skewing the results. Stem and leaf plots and an Explore analysis in SPSS found three outliers for students in the data set. Students identified as outliers in the dataset included: (a) a third

grade boy in the supported condition identified with a learning disability, (b) a 5th grade boy in the supported condition identified as dually exceptional and (c) a fifth grade boy in the unsupported condition identified as dually exceptional. The initial sample had 33 students in the supported condition with 5 students with LD or who were identified as dually exceptional, and 30 students in the unsupported condition with 7 students with LD or who were identified as dually exceptional. The removal of these three outliers yielded the following total numbers for each subgroup (supported = 31, and unsupported = 29), but decreased the number of non-general education students in the supported condition. To assess if this significantly changed the group means based on the stratifying variable of initial reading and writing ability as measured by the TEWL-3 and student reading levels, I ran an ANOVA test of means for this variable. Because there were no statistically significant differences between groups ($p > .05$), these outliers were removed from the supported and unsupported conditions for all further analyses.

I reran the MANOVA analysis for all outcome variables and genres with these three outliers removed. As expected, grade level continued to be a statistically significant variable on both outcome variables across all genres. However, the data set with the outliers removed yielded a new area of statistical significance across the condition variable. For the analytic narrative variable, condition was statistically significant F (1, 56) = 6.42, $p = 0.01$ (Cohen's $d = 0.10$). For this variable, the effect of condition favored the supported over the unsupported condition for both 3rd (9.82 vs. 7.84, respectively), and 5th-grade (12.50 vs. 11.40) although there was a smaller difference between the 5th grade students across conditions. There was still no statistically significant effect of condition for the persuasive genre or the word count measures.

One possible explanation for why there was no statistically significant effect for the persuasive genre is that it is possible that students in this sample wrote using their general fund of knowledge as opposed to using the texts they read as sources for generating ideas. Support for this idea comes from the fact that the participating school participates in a healthy eating program and curriculum. As a result, the current participants may have had adequate background knowledge on the topic of whether or not Wendy's is a healthy food option, allowing them to bypass the need to read text. Accordingly, future research should consider the inclusion an additional variable that accounts for the use of evidence from the texts read to further isolate the potential effect of prior knowledge across these two genres.

Conclusions

In summary, students on average performed best on an analytic measure (a measure of students' organization, content, style and conventions) in the informational report genre, followed by the persuasive and narrative genres. Past studies have shown that students appear to write more and better quality essays in narrative relative to expository genres (Olinghouse & Wilson, 2012; and Camp, 1993;). In this study, third grade students performed significantly better in the informational genre over the persuasive and narrative genres with fifth grade students showing no difference in performance across genres. This finding supports the work of Huot (1999) and Myhill (2005) that contend that the task environment and the amount of prior knowledge students have on a topic and genre may influence the effect of genre on students' writing performance. With these variables controlled through the use of supports such as a common text, and genre and audience awareness cues, it appears that expository genres

are accessible and manageable for elementary aged students.

In addition, while students still wrote more in the narrative genre relative to the informational and persuasive genres, in this study, the relationship between quantity and quality was not as explicit. While the word count and analytic outcome measures were highly correlated across genres, further analysis of the means for the sub-categories of the analytic measure show that content, style and conventions were higher in the informational genre relative to the narrative and persuasive genres while the organization category was relatively stable across genres. This finding also supports the hypothesis that the inclusion of topic content supports through a commonly read text in the informational and persuasive genres may mitigate the demands on aspects of the writing process such as spelling and content generation. As such, while students may be able to write more in the narrative genre, this may not necessarily equate with quality when controls for topic and genre cues are provided in the informational and persuasive writing prompts.

Another significant finding in this study was the effect of condition and grade level on students' performance on analytic writing quality and total word count. As expected, grade level was a statistically significant factor in student performance across all outcome measures and genres. In contrast, the effect of condition was more variable. For the analytic informational variable, there was an interaction effect for condition and grade suggesting that for 3rd graders, there was a significant effect (ES=0.90) for the read aloud accommodation on the informational report genre favoring the supported condition. In contrast, at the 5th grade level students performed slightly better in the unsupported condition based on the means (ES=0.12). One possible reason for the discrepancy in the

effect of condition by grade level is the level of reading proficiency between 3rd and 5th grade students. Presumably, most 3rd grade students are still learning how to read. As such, the read aloud accommodation may help younger or weaker students overcome the additional challenge of the read and respond writing context. For 5th graders who are typically more proficient readers, the read aloud accommodation may potentially hinder a student's ability to read and apply what they have learned from the texts to their writing. Further research into this effect is needed to confirm this hypothesis.

In a post-hoc MANOVA analysis of the outcome measures with outliers removed from the data set, condition was also found to have a statistically significant effect for the narrative genre favoring the supported condition for both the 3rd and 5th grade although the difference in means was greater for 3rd than 5th (1.98 vs. 1.10). This finding further supports the hypothesis that the benefit of the read aloud accommodation may decrease as children become more proficient readers and writers.

While there was no statistically significant effect of condition for the persuasive genre, mean differences between conditions across 3rd (supported = 9.88, unsupported = 9.11) and 5th (supported = 12.00, unsupported = 12.6) grade levels still demonstrate the same trend as the narrative and informational genres with the benefit of the read aloud accommodation decreasing as students get older and become more proficient readers.

CHAPTER 5: DISCUSSION

This study provides insight into two important questions: What effect does genre have on early and intermediate elementary students' writing performance? And, what effect does condition and grade have on students' writing performance? Findings from this study are particularly relevant given the new Common Core writing assessment context that will require students to write in response to texts read thus revealing important considerations for both policy and practice. In this chapter, I first discuss the main concerns that this study addresses with respect to testing policy. I also discuss potential implications for instructional practice. I will then discuss limitations of the current study, and will conclude with a discussion of how the present findings relate to future areas for research in the field of writing assessment and instruction.

Implications for Policy

Olinghouse and Santangelo (2010) suggest that there are four primary purposes for assessing students' writing. These include assessing: (a) to identify children who are at-risk for school failure, (b) to inform instructional planning and modification, (c) to monitor students' progress, and (d) to identify students for eligibility for special education services. Since the reauthorization of the Elementary and Secondary Education Act (ESEA) and the implementation of No Child Left Behind (NCLB) in 2001, a fifth possible purpose of assessment has been to determine the allocation of federal funding to schools.

Given the high-stakes nature of standardized testing, it holds that an effort should be made to ensure that such assessments are fair and valid. The primary purpose of this study was to examine the effect of prompt condition, and genre on the writing and

sentence level quality of 3rd and 5th grade students. In an effort to mirror the testing conditions proposed in the new Common Core context, I assessed students in three genres: narrative, persuasive, and informational report (www.corestandards.org) using a read and respond to text writing prompt condition that will be used by both the PARCC and SBAC consortiums.

The PARCC and SBAC consortiums have published accommodations manuals that limit the use of the read aloud accommodation on their writing assessments without a research-base to determine if there are potential writing construct validity threats given the new read-and-respond writing assessment context (Laitusis et al., 2012). The results of this dissertation study indicate that developing or younger readers benefit from a read aloud accommodation before writing, suggesting that the new writing assessments may pose a threat for construct validity at the earlier grade levels. It holds that if developing or younger readers and writers perform more poorly in the read and respond writing assessment context without the read aloud accommodation, than these types of assessments may in fact be just another assessment of students' reading abilities. Given the aforementioned purposes for writing assessment, the results of the current study support new concerns that writing assessments without support for reading text may be unfairly biased against students who are struggling readers. Another result from this study to take into consideration is the inverse relationship that the read aloud accommodation has on more proficient or older children. In this study, the benefits of the read aloud accommodation appeared to cease for older children and in some cases appeared to even hinder a child's ability to apply what they had read effectively to their writing.

There are a number of potential policy implications that can be drawn from this study. First, it is clear that there is an effect of the read aloud accommodation on writers in the new assessment context. It is also clear that this effect may be different based on grade level and learner characteristics. Given this possibility, assessment consortiums such as PARCC and SBAC should consider how to appropriately accommodate all learners, but particularly developing readers and writers in the new writing assessment context. Starting in the 2015 school year, all 3rd through 12th-grade students will be formally assessed in the three genres explored in this study. In many states, this will be the first time students in the 3rd grade will be asked to write in a standardized testing context. For younger and developing readers, specialized testing formats or supports may be needed until they reach a level of reading proficiency that allows them to integrate what they read into their writing effectively to prevent any construct validity threats.

Another area that warrants consideration based on the results of this study is the potential effect of task environment on students' performance across the genres. Research on writing at the elementary level has shown that prior knowledge and task environment can have a significant effect on the writing performance of students in general education, as well as those with LD and students who are EL (Crosson et al., 2012; Olinghouse & Santangelo, 2010, Donovan, 2001; and McCutchen, 1998). Huot (1990) and Myhill (2005) suggest that a number of prompt condition factors can promote students' writing performance across genres. These include a specified superior audience, the inclusion of genre specific cues, and the inclusion of background knowledge content support in the topic of the prompt particularly for expository genres.

In an effort to isolate the effect of genre and prompt condition, I implemented all

of the recommended practices proposed by current research in all of the writing assessments in this study regardless of prompt condition. These included a specified audience for each prompt (narrative = Mini-Pages, informational report = Time for Kids, and persuasive = Kids Post), supportive content texts for the informational report and persuasive genres, a focused picture prompt in the narrative genre, and genre specific checklists for each genre. Finally, all students were provided the opportunity to use a generic graphic organizer across all writing assessments, which has proven to have a supportive effect in the writing process (Marzano, Pickering & Pollock, 2001).

Previous work suggests that children's development in writing may vary by genre with most students in the elementary grades performing better in narrative over expository genres (e.g., Olinghouse & Wilson, 2012 and Scott & Windsor, 2000). In these previous studies, researchers predominantly used open-ended writing prompts across the genres where students were asked to write exclusively from their personal background knowledge and experience. For example, in Olinghouse and Wilson's study students were asked to write in the following ways: (a) the narrative task prompted students to write a story in response to a picture of astronauts on the moon uncovering something on the moon's surface, (b) the informational task asked students to write a report about outer space, and (c) the persuasive task asked students to write a letter about whether or not President Obama should build places to live in outer space. When viewing open-ended prompts of this nature, it is more difficult to isolate the potential effects of background knowledge of the topic from students' ability to write effectively within each genre. This study contradicts earlier findings suggesting that when students are provided the task environment supports listed above such as common content from texts read,

expository genres such as persuasive and informational report writing may become more accessible. Additionally, it may further support the findings by researchers like Crosson et al. (2012) that suggest that background knowledge can have a significant impact on students' performance across the genres. For example, there is evidence from this study that students may have used their background knowledge on healthy eating from their school program in lieu of the provided text to write their response thereby negating the need for a read aloud accommodation. Accordingly, an effort should be made to control for this effect.

In addition, the above findings suggest that test-makers, researchers, and educators should make a concerted effort to include research-based writing supports such as a specified audience, genre specific cues, graphic organizers, and background knowledge supports when assessing students' writing. While the use of a read and respond testing prompt may help to accommodate potential background knowledge deficits particularly with respect to informational report and persuasive writing topics, this study shows that a students' reading ability may influence their ability to write. In order to isolate the construct of writing from that of reading, other vehicles for providing background knowledge on the topic in testing contexts should be considered. These could include read-aloud presentations or videos alongside texts such as those used in Scott and Windsor's (2000) study. In particular, the PARCC and SBAC consortiums should consider applying these strategies to the writing assessments they create to ensure that the targeted writing skills of organization, content, conventions, and use of evidence are isolated from other biasing factors such as background knowledge and reading ability.

Implications for Research

A potential challenge facing researchers in the new writing assessment context is the interpretation of the Common Core writing standards and its subsequent assessment. In particular, how researchers assess the role of genre in a research context presents particular difficulties for the field. . Genre theory represents a multi-faceted and complex collection of ideas that are marked by regional, field-specific, and individual differences and similarities (Hyland, 2009). Swales (2009) suggests that there are multiple theories of genre including (a) a balance between constraint and choice, (b) the role of culture in the realization of genre exemplars, (c) a sense that genre evolves in response to various constraints and demands, and (d) a nuanced approach to teaching and gaining and understanding of genres. In the Common Core writing assessment context, it appears that there are also numerous interpretations of genre. For example, a PARCC published document that shares sample writing forms (http://www.parcconline.org/sites/parcc/files/PARCCSampleofWritingForms.pdf) lists 40 sub-genres of writing as potential vehicles for assessing the Common Core standards in grades 3-8. These range from prompting children to write satires, spoofs, testimonials, apologies, endings, biographies, fables, explanations, and more. Given the wide range of possible genre prompts, that are being proposed by testing consortiums, researchers will likely have difficulty isolating and standardizing an approach to writing assessment research that can be repeated and also aligned with the demands of the school-based writing assessment context.

Furthermore to date, most writing assessment research has relied heavily on holistic measures similar to those used by the National Assessment of Educational

Progress (see http://nces.ed.gov/nationsreportcard/). In contrast, the new PARCC and

SBAC rubrics are analytic measures that include a new focus on the role that reading

comprehension has on writing. Accordingly, an effort should be made among researchers

to begin to standardize measures and assessment protocols to better align with the new

writing assessment sample items being published by these two national assessment

consortiums.

Implications for Practice

In this study the average rubric scores for the analytic measures placed most

students at an inadequate or developing writing level. Students appeared to perform the

most poorly in organization, style and conventions and relatively better in the general

inclusion of content. While the participants in this study may not be wholly representative

of the elementary school population at large based on the fact that they attended a

bilingual school, the results from this study imply that many elementary students may

have good ideas to share, but lack the facility to express these ideas clearly and accurately

through the written word. Research suggests that the quality of students' writing is often

impeded if they have difficulties with writing conventions such as spelling and

transcription (McCutchen, 2000). Given the link between these two variables, additional

emphasis on the mechanics of writing may be beneficial.

The newly adopted Common Core writing standards place a great deal of

emphasis on structure and content, and place less emphasis on writing mechanics. There

are standards that specify language use in both speaking and writing, which include the

use of various verb forms, sentence structures, and standard conventions. There is a

concern that the separation of these skills from the core writing standards may lead to

their omission in daily instruction. As teachers work to design writing units of study, they should make a concerted effort to systematically identify the language standards that appropriately match the overall goals of the writing units they plan around the Common Core standards to ensure that sentence level skills are consistently and repeatedly taught throughout the language arts curriculum.

A commonly held belief in the world of reading instruction is that in "In K–3 children are learning to read, and in 4–12 children are reading to learn" (Chall, Jacobs, & Baldwin, 1990; Chall and Jacobs, 2003). Unfortunately, in the new Common Core Context with students being asked to read and respond to texts in writing starting in 3rd grade, schools and educators can no longer afford to wait until 4th grade to explicitly make the connection for students that reading is a skill that needs to be applied to a broader context. Likewise, Houck and Ross (2012) argue that learning to read should continue well past the early grades as children are taught more sophisticated strategies for comprehension. In effect, learning to read and reading to learn should happen simultaneously across the grade levels.

A similar myth exists in the world of writing that suggests that in the early grades; narrative writing is more accessible to children than expository genres (Calkins, 1986, Olinghouse & Wilson, 2012; and Camp, 1993). This study disproves this assumption. Data now shows that when students are provided the appropriate task environment supports, elementary aged students can write effectively in expository genres. Accordingly, an effort should be made to give students sufficient opportunities to write across multiple genres. Teachers should shift writing instruction away from daily journal entries, to authentic and purposeful reasons to write across the genres on a regular basis.

Furthermore, based on the results of this study, best practices for writing assessment such as specification of audience, genre cues, and topic control through the use of shared texts would likely be beneficial to students during general writing instruction throughout the school day.

Given the poor performance of students across genres in this study and in previous studies (e.g., Hebert, Graham & Harris, 2010), it is clear that students in the younger grades need more practice writing from an earlier age. In particular, students will likely need significantly more practice writing from texts they read. One possible effect of the learning to read and reading to learn myth may be that up until recently, elementary educators in the early grades have placed undue emphasis on learning to read without providing students sufficient opportunities to apply what they have read across multiple contexts such as writing. This coupled with the emphasis on narrative genre structures in both reading and writing in the early grades may provide students with insufficient exposure to informational and persuasive text structures.

Bridges (2012) argues, "Approximately half the texts an elementary school student should encounter should be nonfiction increasing to 70% by the time students are in high school." (p.9). She also highlights the work of Yopp and Yopp (2006) and Jeong, Gaffney and Choi (2010) that suggest that in preschool through grade 3 children seldom encounter informational texts with 2nd graders experiencing 1 minute per day, and 3rd and 4th graders averaging 16 minutes per day of exposure to these types of texts. Given the demands of the new Common Core context and writing assessment formats, it is imperative that teachers increase the amount of time that students get to work with non-narrative genres.

One way to improve students' level of engagement with complex texts across the genres is through close reading. Close reading is an instructional method for reading instruction that is commonly seen at the secondary and post-secondary level, but less so in the primary grades (Fisher & Frey, 2012). The purpose of close reading is to give students the opportunity to read complex and higher Lexile leveled texts through supporting text-dependent questions that highlight the metacognitive skills necessary to make inferences about the text, determine the author's purpose, and identify when something is confusing. Fisher, Frey, and Lapp (2012) contend that as students gain experience with reading texts that are quantitatively and qualitatively more complex, the materials they are able to read independently will increase as well. This is an important goal given that the Common Core standards have also included the adoption of higher Lexile band requirements from grade level to grade level (www.corestandards.org).

According to Fisher and Frey (2012), close reading includes the following key features: (a) short passages, (b) complex texts, (c) limited front-loading, (d) text-dependent questions, (e) repeated readings, and (e) annotations. The role of annotating in the close reading process is of particular interest in the writing assessment context. Text annotation is a common practice at the secondary level as students learn to make notes about the texts they read to support their analysis. Fisher and Frey note that this strategy is also "useful in analytic writing about text, as students consult their annotations to formulate arguments, analyze information, and make connections within and outside of the text." (p. 186).

In their meta-analyses of the relationship between reading and writing, Graham and Perin (2007) and Graham and Hebert (2010) both show that writing has a strong and

positive impact on reading comprehension and development. Unfortunately, up until recently, writing instruction and practice has not been a core part of the elementary curriculum. This is likely driven by the fact that prior to the adoption of the Common Core and the implementation of the PARCC and SBAC writing assessments, the assessment of writing has not been an expectation at the state and district level as it is not required for the purposes of meeting No Child Left Behind requirements (www2.ed.gov/nclb/accountability). Graham (2010) noted that in the primary grades, students spend only 20 to 30 minutes a day writing with little time spent writing in expository texts. Graham also notes that 40% of teachers make few or no adaptations and spend very little time teaching the writing process. At the 4th-6th grade levels, students on average receive only 15 minutes a day of writing instruction with only 25 minutes a day for actual writing practice.

Literacy blocks in elementary schools continue to be largely dedicated to reading work (Graham, 2010). One way to better integrate the goals of writing instruction and development into the literacy block is to more explicitly link the close reading instructional sequence to writing outcome measures. In this context, writing outcome measures should go beyond the use of annotations or short journal responses to text dependent questions. In the current close reading lesson sequence, students are asked to read a text multiple times in an effort to answer a series of text dependent questions that reveal aspects of the text such as author's purpose, figurative language, text structures, and perspectives (Fisher & Frey, 2012). As close reading instructional approaches are used with greater frequency, an effort should be made to include an additional component to the close reading sequence that requires a culminating writing task that goes beyond

journal responses to assignments that require students to respond to reading in one of the three targeted genres of the Common Core (e.g., a narrative, argumentative/persuasive, or informational essay). This would make the connection between reading and writing in the classroom more explicit for both students and teachers.

Finally, both the PARCC and SmarterBalanced consortiums have reported that the new assessments will include some form of computer adaptive testing technology. Accordingly, students will need to gain familiarity and expertise with a number of technology based skills particularly with respect to word processing. MacArthur (1996) notes that word processing can support writing in several ways. These include: (a) the ability to produce neat and legible text for students who may have challenges with fine-motor processing skills, (b) the ability to use editing tools to improve spelling, (c) the ability to use features such as cut and paste for the purposes of revising, and (d) the ability to publish work in a uniform format. Nevertheless, all of the benefits of word processing cannot be maximized without giving students opportunities for practice with such technologies.

Graham and MacArthur (1993) found that typing skills and familiarity with different software and hardware features of technology were necessary in order for children to effectively utilize word processing features in their writing. They noted that this required regular access to technology. Unfortunately, Graham (2010) noted that over 20 years later, students' opportunities to work with technology in the elementary grades is still extremely limited with many elementary-aged children receiving little to no time to work on computers. Alternatively, there is a concern that children may have access to certain technologies (e.g., iPads), but not necessarily technology that promotes writing

development and skills specifically. Furthermore, in a study of the differential effects of dictation, handwriting, and word processing for 5th and 6th-grade students with LD, MacArthur and Graham (1987) found that students with LD struggled with word processing skills and there was little to no difference between their handwritten and typed essays. For example, in their study students with LD wrote on average 4.6 words per minute, which was less than half of what they could produce by hand. In addition, the quality of these essays were significantly lower than those produced through dictation. In contrast, for general education students, research has shown that there is little difference in quality measures between dictated and handwritten (Hidi & Hildyard, 1983), and handwritten and typed essays (Daiute, 1986).

Moving forward there is a concern that with the introduction of technology based assessments, additional variance that is outside of the realm of writing ability may be introduced into writing assessment. It holds that if children's abilities to effectively use technology to demonstrate writing proficiency hinges upon their ability to access technology for regular practice, than students in school systems with less access to technology both at home and at school will be additionally disadvantaged. Furthermore, research to date shows that students with LD in particular may have greater difficulty mastering technology based writing applications than their general education peers (MacArthur & Graham, 1987). In a review by De La Paz (1999), she argues that one way to provide student with LD greater access to technology based writing requirements is through the use of adaptive speech-recognition systems (SR). She noted "while our understanding of the impact of dictation and SR systems on the composing of persons with LD remains unclear, the latter may allow individuals to transcribe at rates closer to

the speed of speech, a development that may result in improved writing for some." (p. 180). De La Paz's work coupled with the findings from MacArthur and Graham (1987) regarding students with LD's ability to compose higher quality texts under dictation conditions bring to the forefront another important consideration for writing assessment. It holds that if the new writing assessment measures will require some form of computer-based technology, than the role of word processing and assistive technologies for writing warrants further investigation particularly with respect to students who may be disadvantaged socially, economically, linguistically, or academically.

Limitations

A significant limitation to this study was the sample size. While there were 63 participants in this study, my goals of assessing the effects of interest required the reduction of this total population into two conditions shrinking the sample size across these groups to 33 and 30 respectively. As a result, there were significant limits to my ability to disaggregate student performance across student subgroups such as students with LD, students who are EL, and students who are dually exceptional. Additionally, due to limits in sample size and its resulting effect on power, I was unable to run statistical analyses for all outcome variables that I collected data for. In addition to the analytic and total word count variables, I had also collected data using a holistic and correct-word sequence variable. For the purposes of this study, as I was limited in the number of variables I could enter into the models to achieve statistically significant results (Faul, et. al., 2009). Accordingly, of the two measures of writing quality (holistic and analytic) I prioritized the analytic outcome variable that was best aligned with the PARCC and SBAC assessment measures. I then selected the sentence level variable that

was most frequently used in research to date at this developmental level (e.g., Beers & Nagy, 2010; Olinghouse & Graham, 2009, Hudson, Lane & Mercer, 2005; Scott & Windsor, 2000, etc.), in this case the total word count measure.

Additionally, the population at my selected school setting is very unique. The combination of its urban setting, immersion program and the diversity of students may present threats to external validity factors such as generalizability of individuals. Unfortunately, there was also very limited information available on students' baseline skills across grade levels and learner sub-groups. While I was able to obtain some information on students EL levels, students with LD were only designated by the presence or absence of an IEP. In addition, due to the fact that 3rd grade was the first standardized testing year, there was limited baseline data on the 3rd grade students to compare. Another limitation in this study was that I was not able to isolate the potential effects of instruction due to limited access to teachers for follow-up interviews regarding their reading and writing instruction practices in the classroom. In the persuasive writing task, students' background knowledge on healthy eating habits, types of food, etc., through their school-based program may have skewed the results for the persuasive genre assessment that asked students to argue whether Wendy's was a healthy choice. Given the possibility of instructional effects on student performance, additional information on students' backgrounds and learning experiences in the classroom would have been helpful to contextualize the above findings.

Finally, while the use of multiple assessment measures and multiple assessments helped to control for mono-method bias and mono-operation bias, it may have lead to threats to internal validity in the areas of fatigue, and history.

Future Research

There are a number of potential areas for future research. First, future studies should attempt to include more schools in order to include a larger sample size particularly with respect to subgroups of interest such as students with LD and students who are EL. In addition, researchers should consider duplicating the above study with varying settings to see if the findings above might be unique to urban and/or immersion classroom populations.

Another area to further explore is the effect of the read aloud accommodation on developing readers and students with different learner characteristics. This study presents preliminary findings to suggest that the read aloud accommodation is supportive of 3rd graders, but less so of 5th graders. Furthermore, the effect of the read aloud accommodation on struggling learners appears to be extremely variable suggesting that individual learning characteristics beyond broad categories such as students with LD, students who are EL and students who are dually exceptional may influence the efficacy of this accommodation.

While the read aloud accommodation appears to consistently support younger and developing readers overall, its effect for students with LD, students who are EL, and dually exceptional students is not yet known and should not be taken as wholly beneficial to these subgroups in the writing assessment context. Descriptive statistics for these subgroups in this study show that there was a large amount of variability in students' performance across conditions and genres. One possible explanation for this phenomenon comes from the field of neuroscience. Recent research suggests that some language-based learning disabilities that lead to challenges with literacy skill development can be

attributed to auditory processing issues (Heim, Keil, Choundhury, Friedman & Benasich, 2013 and Lehongre, Ramus, Villiermet, Schwartz, & Giraud, 2011). Given that the read aloud accommodation relies on providing children auditory stimuli, this accommodation may not be as supportive to struggling readers as originally hypothesized. Unfortunately, given the small sample size for each of the subgroups in this study, and the likelihood for selection bias and the influence of individual differences on the group comparisons requires additional research to generate conclusive findings.

Accordingly, another possible research avenue could be to further investigate the procedure for providing the read aloud accommodation to see if there is a differential effect for students depending on the population. For example, all students may receive more benefits from the read aloud accommodation if they are first asked to read texts on their own. This could be true for not only writing assessments, but also for reading and math assessments that warrant the read aloud accommodation as well. Alternatively, other reading accommodations other than the read aloud accommodation could be more effective such as providing students the opportunity to read smaller segments of text over an extended period of time.

In this study, depending on the measure of interest, students on average appeared to write better or equivalently well in the expository genres (persuasive and informational report) as they did in the narrative genres. Potential causes for this shift in students' performance from stronger narrative writing than expository writing as seen in previous studies (e.g., Olinghouse & Wilson, 2012; and Camp, 1993) could be related to various features included in the task environment or prompting condition. These included supports such as a specified audience, genre specific checklists, graphic organizers, and

the inclusion of texts to read to support background knowledge on the topic. Nevertheless, the design of the current study did not allow for the isolation and exploration of how and why these factors may have influenced students' stronger performance in the expository genres over the narrative genre or to factor in other influences such as increased attention by teachers to these genres in response to policy changes such as the CCSS-ELA. Future studies should consider testing the effects of each of these supports individually and conjointly within and across genres to identify the best possible combinations of supports for developing writers.

The new writing assessment formats also warrant further exploration. If the next generation writing assessments will all follow the read and respond type of structure, additional research into the effect of different types of texts and task environments are needed to ensure that these types of writing assessments can reliably assess the construct of writing as opposed to reading. In particular, further analysis of writing samples to determine the relationship between students' use of evidence from texts and their resulting writing quality is warranted. While in this study I did not use an analytic rubric that included a category on the use of evidence from text read, this is a category on both the PARCC and SBAC rubrics. Accordingly, adding this variable should be an area for future research as well.

In addition to exploring the effects of different types of accommodations and supports for struggling learners, specific attention should be directed to exploring how dually exceptional students who are EL and have LD write. Very little is known on how this growing population of students write and given their poor performance relative to all other peers in this study, this population deserves additional support and attention.

Moving forward, qualitative analysis of the writing of the participants in this study could yield interesting insights into similarities and differences, as well as areas of strengths and needs in the writing of this sub-group of students.

Finally, given the new computer based technology assessment formats that will be introduced with the SBAC and PARCC assessments, research into the effect of word processing and assistive technology programs on writing performance warrants further review. In particular, an effort should be made to explore the potential biasing effects that may be included in writing assessments that require proficiency in technology for students who may have limited access to such resources such as those students who are socially, economically, linguistically, and academically disadvantaged.

Conclusions

The current study is significant for several reasons. First, the findings in this study offer a more refined view on the factors that affect student performance within and across genres such as prior knowledge on the topic and genre, as well as the potentially unique effect of the read aloud condition on students in different grade levels. By better isolating these factors, researchers will hopefully be able to gain a better understanding of whether or not students' writing performance across different genres is a function of development, differences in levels of cognitive demands across tasks, prior knowledge, or learner characteristics. This may in turn allow educators to appropriately identify benchmarks and standards for learning. Second, findings from this study may afford the opportunity to gain a more holistic understanding of how to better prepare students to write in in the SBAC, PARCC, and Common Core writing context. Finally, findings from the current study could be used to further develop standardized writing assessments and formative

classroom assessments that minimize known biases to improve measurement practices

and ultimately, instruction and curriculum.

APPENDIX A

Content Review Matrix

Topic	Authors (Date)	Research Questions or Purpose	Design	Sample/ Participants	Ind. Variables	Dep. Variables	Results
Prior Know. Effects	Degroff (1987)	Knowledge of topic on expository writing	Correlational	4th-grade students 20 high knowledge (17 boys and 3 girls) 20 low knowledge (15 girls and 5 boys)	Low and high prior knowledge of baseball as measured by Voss et al. (1980) questionnaire	Voss et al. (1980) baseball grammar Mean t-unit length Students' self-rated holistic scores.	Prior knowledge was found to be related to goal-related in-formation in high-knowledge writers' first and second drafts and to comments from conferences with high-knowledge respondents. Also, prior knowledge was related to non-goal-related information in low-knowledge writers'. High knowledge writers also had more syntactically complex pieces of writing and longer pieces of writing.
Prior Know. Effects	Englert, Raphael, Fear, Anderson (1988)	Students' metacognitive knowledge about how to write informational texts	Quasi-experimental	30 (LD, low-achieving, high achieving) groups 4th and 5th grade students	Metacognitive knowledge Learner characteristics Genre	Writing interviews (high to low knowledge scores) Primary trait score Holistic score	When performance levels among the three ability groups were compared, the results suggested that learning disabled students were less aware than high-achieving students of modeled writing strategies, steps in the writing process, strategies for presenting expository ideas, and procedures for selecting and integrating information from multiple sources. Discrepancies between learning disabled and low-achieving writers also emerged in the metacognitive interview in terms of ability to (a) control and regulate the writing process, (b) use organizational strategies or text structures to generate or group ideas, and (c) monitor the quality of texts. When performance on the composition and metacognitive measures was related, the results revealed that the strongest relationship existed between writing performance and the following metacognitive variables: students' awareness of modeled writing strategies, students' knowledge of processes related to monitoring the completeness of text, and students' categorizing abilities. These findings suggest that writing instruction should focus on both the development of students' metacognitive knowledge of the expository writing process and the organizational strategies for generating,

							organizing, and monitoring expository prose.
Prior Know. Effects	Lee, Penfield, Buxton (2011)	Relationship between form and content knowledge on expository science writing of Students who are EL	Quasi-experimental	3 years of a 5-year study of 3rd graders

Year 1: 683
Year 2: 661
Year 3: 676 | ESOL classification (ESOL 1-4; ESOL 5; Exited ESOL or never in ESOL) | Two scoring rubrics: "form" (conventions, organization, and style/voice), "content" (specific knowledge and understanding of science). Both rubrics used a 5-level system based on a continuum of "minimal" (score of 1) to "comprehensive" (score of 4), with a score of 0 used to indicate irrelevant or no response. | The results indicated significant relationships between writing form and content at both pretest and posttest, with a stronger relationship at posttest. The effect of English proficiency on the magnitude of the relationship was significant only at posttest, for which the relationship was stronger for non-ELL students. The results suggest that through our intervention over the course of the school year, students with greater English proficiency learned science content and developed English literacy simultaneously, whereas students with lower English proficiency did not show this simultaneous growth to the same degree. |
| Prior Know. Effects | Olinghouse & Graham (2009) | Discourse knowledge effects on narrative writing | Correlational | Grade 4 students (18 boys, 14 girls) possessed more discourse knowledge than Grade 2 students (18 boys, 14 girls). | Discourse knowledge (substantive, production, motivation, story elements, and irrelevant) | Story quality, length, and vocabulary diversity beyond the 7 control variables--4 writing (handwriting fluency, spelling, attitude toward writing, advanced planning) and 3 nonwriting (grade, gender, basic reading skills) variables. | Five aspects of this discourse knowledge (substantive, production, motivation, story elements, and irrelevant) together made a unique and significant contribution to the prediction of story quality, length, and vocabulary diversity beyond the 7 control variables. In addition, older students possessed greater knowledge about the role of substantive processes, motivation, and abilities in writing. Findings support the theoretical propositions that discourse knowledge is an important element in early writing development and that such knowledge is an integral part of the knowledge-telling approach to writing. |
| Task Effects | Brodney, Reeves, Kazelskis | Selected Prewriting Treatments: | Quasi-experimental | 96 5th-grade students | Condition (pre-writing groups) | Used multivariate analysis of covariance to examine the effects of the four prewriting | A significant ($p < .001$) multivariate F-ratio indicated that type of prewriting treatment significantly affected scores on expository |

	(1999)	Effects on Expository Compositions Written by Fifth-Grade Students			• Reading-prewriting • Reading only • Prewriting only • comparison	treatments on the holistic measure (HM); the four analytic measures of ideas (ID), style (ST), organization (OR), and mechanics (MC); and total words per T-unit (TU). Raw reading achievement served as co-variate	compositions. Reading paired with prewriting before composing was found to be the most effective prewriting instructional strategy.
Task Effects	Hudson, Lane, and Mercer (2005),	(1) What are the effects of writing prompts on the compositional fluency of second-grade students? and (2) Are these effects different for students who vary on spelling achievement and handwriting fluency?	Correlational	195 2nd grade students Four schools in one north central Florida district free or reduced-price lunch 49.4 to 86.9 (M ¼ 69.4).	Condition Occasion	compositional fluency, including handwriting fluency and spelling achievement	Priming condition and writing occasion had significant effects. Interactions were found between priming condition, occasion, and each of the covariates (handwriting fluency and spelling achievement). Analysis of the sample based on the covariates revealed differential effects for the slowest and fastest writers and poorest and best spellers.
Task Effects	Midgette and Haria & MacArthur (2007)	The purpose of the study was to investigate the effects of revising goals focused on content and audience awareness on the persuasive writing of fifth- and eighth-grade students.	Quasi-experimental	181 5th and 8th graders one elementary school and two middle schools in an urban/suburban school district in the mid-Atlantic region of the United States. Free reduced lunch 29% elementary No SWD or ESL	Condition Grade Gender	Final drafts of essays were scored for elements of persuasive discourse relevant to content and audience and for overall persuasiveness.	Students in the audience goal group were more likely than both other groups to consider opposing positions and rebut them. Students in both the content and audience goal groups wrote essays that were more persuasive than essays by students in the general goal group. The results also indicate that eighth grade students wrote more persuasively than fifth-grade students and that girls wrote more persuasively than boys.
Genre Effects	Beers and Nagy	How does syntactic	Correlational	83 students in grades 3 and 5	Grade level 3, 5 and 5, 7	2 measures of syntactic	For clauses per T-unit, significant differences were found between persuasive essays, which had

144

	(2010)	complexity develop across genres from mid-elementary years to early adolescence?		96 students in grades 5 and 7 Same cohort of students on 2 occasions 2 years apart as part of a larger longitudinal study	Genres of narrative, descriptive, compare/contrast and persuasive	complexity: clauses per T-unit and words per clause on four genres of text: narrative, descriptive, compare/contrast, and persuasive	more subordinate clauses, than the 3 other genres. For words per clause, significant differences were found btw descriptive texts, which had more words per clause than persuasive essays, which did not differ from the compare/contrast texts. For text length grade x genre effects were significant for both cohorts. Findings suggest that although students could produce each kind of genre, their ability to do so may have been compromised by their limited knowledge of the syntactic structure necessary to achieve text-level genres.
Genre Effects	Prater and Padia (1983)	What are the effects of three modes of discourse on student writing performance at two elementary grade levels? Will grade/ topic, sex/topic, or grade /sex/ topic interaction occur?	Correlational	70 fourth grade students 70 sixth grade students from six elementary schools in California. Schools were selected to represent a mixture of urban/ suburban settings and a diversity in SES.	Sex, grade, discourse topic	Each essay was scored using a four-point holistic scale. Essays were given two independent readings and scores varying by more than two points were arbitrated by a table leader.	Expressive writing tasks generated essays at both grade levels for both sexes that were judged to be higher in quality than either Explanatory or Persuasive writing tasks. Persuasive writing tasks were found to be the most difficult type of writing for all subgroups with the exception of fourth grade boys whose lowest performance was on the Explanatory writing task. Predictable sex and grade differences were found across all modes of discourse.
Genre Effects	Scott and Windsor 2000	Does GLPM distinguish school-age children with and without learning disabilities based on (a) language group; (b) discourse genre (expository vs. narrative); (c) modality (spoken or	Correlational	60 students 20 each in matched triplets (LLD, Chronological Age, and Language-Age Peers) upper-elementary to middle-school age students in grades 3-7 average or above average intelligence on Test of Nonverbal	LLD, LA, and CA matched for gender and socioeconomic status Spoken and written narrative and expository responses to 2 videos (one for narrative and one for expository)	10 General Language Performance Measures (e.g., productivity measures, total T-units, total words, etc.)	Group differences were seen for 5 of the 10 GPLM Productivity differences between LLD and CA was notable. All students, regardless of language status, were affected in similar ways by genre. Main effects for genre favored narrative over expository contexts.

		written); and (d) interactions among group, genre, and modality?		Intelligence			

APPENDIX B

Validity Criterion and Definitions

Category	Criterion	Definition
Internal Validity	Selection	Systematic differences in participants may account for outcome differences.
	Testing	Repeated exposure to an assessment measure may impact subsequent scores, making it difficult to ascertain effect of an independent variable.
	Instrumentation	Changes in measurement protocols (e.g., instruments, rubrics, assessments) may account for outcome differences.
Construct Validity	Inadequate explication of constructs	Construct defined too broadly or narrowly which may lead to incorrect inferences about the relationship between the variables examined.
	Construct confounding	Presence of other possible constructs that may mask the effects of the measured construct.
	Mono-operation bias	Single method of measurement is used to measure a construct.
Statistical Conclusion Validity	Violated assumptions of statistical tests	Selecting an inappropriate statistical assumption may lead to over- or underestimating of an intervention's effect size.
	Unreliability of measures	Measurement error weakens the relationship between two variables and strengthens or weakens the relationships among three or more variables.
External Validity	Generalize to individuals	An effect found with certain individuals might not hold if other individuals had been studied.
	Generalize to other assessment outcomes	An effect found on one kind of assessment may not hold if other assessments had been used.

APPENDIX C

Methods Review Matrix

Focus	Study Authors	Internal Validity			Construct Validity			Statistical Conclusion Validity		External Validity	
		Selection Bias	Testing	Instru-mentation	Mono-Operation Bias	Expli-cate con-structs	Confound-ing of con-structs	Violation of Statisti-cal Assump-tions	Unreliability of Measurement	Generalize to individuals	Generalizability of Assessment Outcomes
Prior Know. Effects	DeGroff (1987)	NO	n/a	NO	YES	YES	YES	YES	NO	NO	NO
Prior Know. Effects	Englert, Raphael, Fear, Anderson (1988)	YES	YES	YES	NO	YES	YES	YES	NO	NO	NO
Prior Know. Effects	Lee, Penfield, Buxton (2011)	YES	NO	YES	NO	YES	YES	YES	YES	YES	NO
Prior Know. Effects	Olinghouse & Graham (2009)	NO	NO	YES	YES	YES	YES	NO	YES	NO	NO
Task Effects	Brodney, Reeves, Kazelskis (1999)	NO	n/a	NO	YES	YES	YES	YES	YES	NO	NO
Task Effects	Hudson, Lane, and Mercer (2005)	YES	NO	NO	NO	NO	NO	NO	YES	YES	NO
Task Effects	Midgette, Haria & MacArthur (2007)	YES	n/a	NO	NO	NO	NO	YES	NO	NO	NO
Genre Effects	Beers and Nagy (2010)	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO
Genre Effects	Prater and Padia (1983)	YES	YES	NO	NO	NO	NO	NO	NO	YES	NO
Genre Effects	Scott & Windsor 2000	YES	YES	YES	NO	YES	YES	YES	YES	NO	NO

APPENDIX D: Narrative, Informational and Persuasive Prompts

Story Writing Contest

Take a look at the picture. What do you think has happened? Can you create an interesting and exciting story about this picture for the "Washington Post Mini-Pages"? You can use the space below to plan or the graphic organizer on page 4.
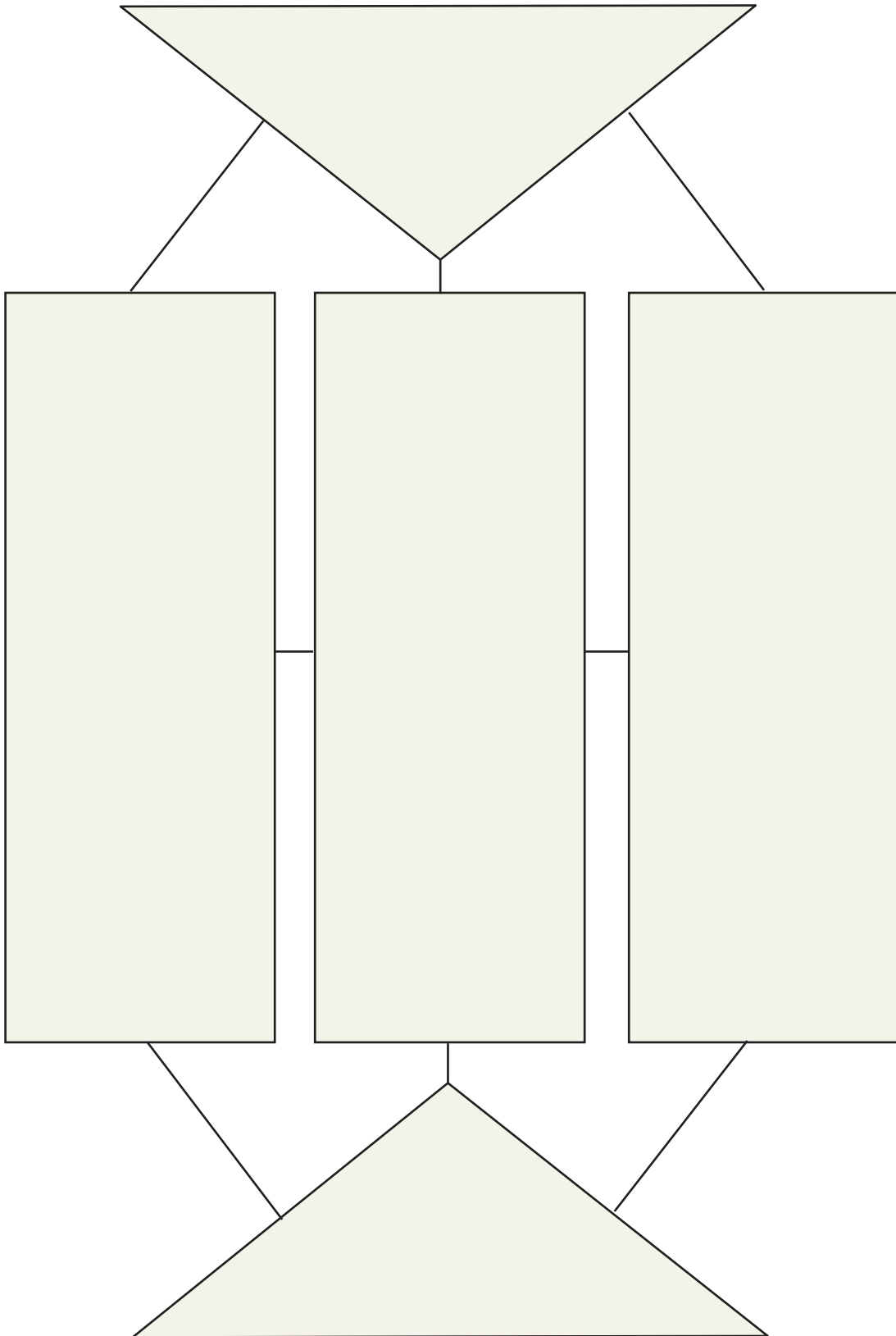
Remember . . .

- ☐ A good story has a beginning that includes the setting, characters, and an introduction to the problem. Take a moment to think about the characters. What are their names? What might his problem be?

- ☐ A good story also has a middle that has interesting details and action. Take a moment to think about what the most exciting part of your story will be. What will the characters do and feel?

- ☐ Finally, a good story has an end with a great solution. How will your character solve his problem?

Plan your story for the "Washington Post Mini-Page" here or on the graphic organizer:

Write your story for the "Washington Post Mini-Page" here:

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____

Edit your paper to be sure that:

- ☐ You used good grammar;
- ☐ You used capital letters and punctuation marks correctly;
- ☐ You spelled words correctly; and
- ☐ You let your readers know where you started new paragraphs.
- ☐ You checked your paper to make sure that it is the way that you want readers to read it.

--3rd-Grade Student Copy—

Panda cubs are born with their eyes closed. Panda cubs are about the size of a stick of butter at birth. When they are born they are hairless and helpless.
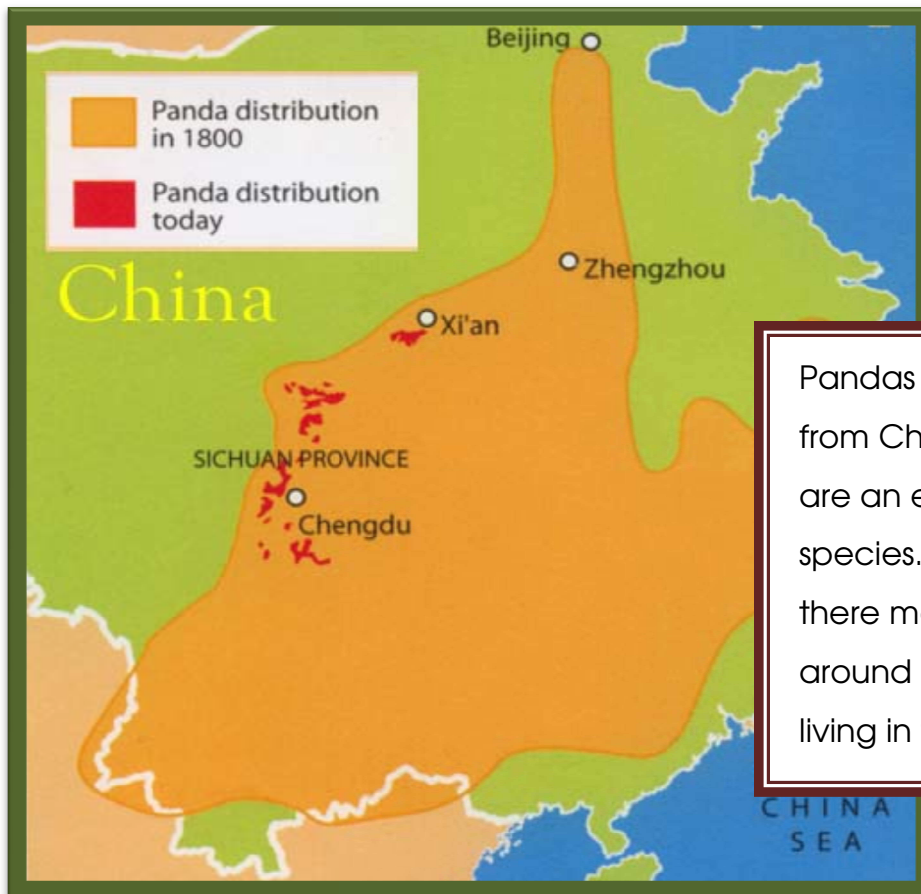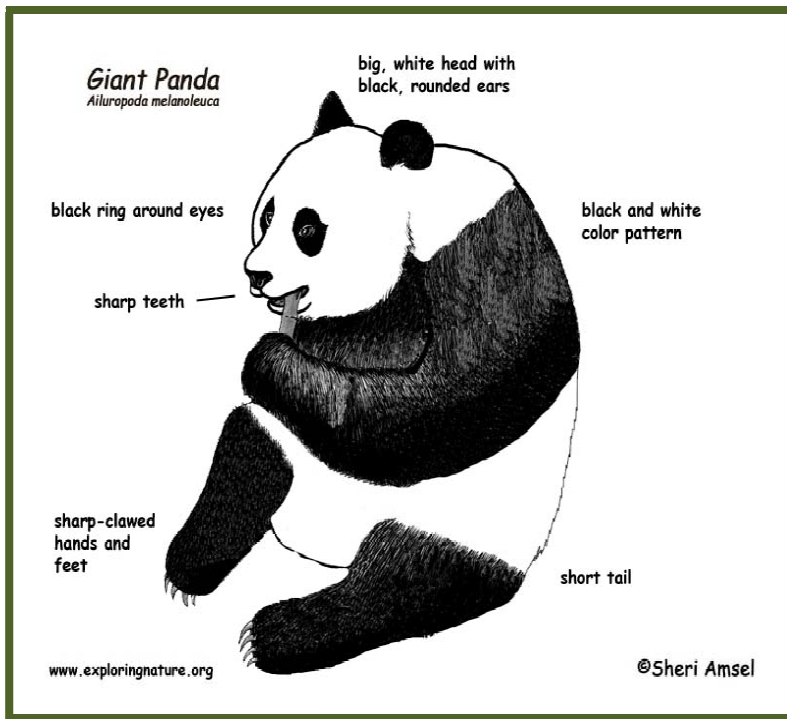
Pandas are originally from China. Pandas are an endangered species. Currently, there may be only around 2000 left living in the wild.

Pandas live in the forest and eat bamboo, insects, bulbs and fruit. Giant pandas eat as much as 22 pounds of bamboo a day.

155

Giant Panda
Ailuropoda melanoleuca

big, white head with black, rounded ears

black ring around eyes

black and white color pattern

sharp teeth

sharp-clawed hands and feet

short tail

www.exploringnature.org

©Sheri Amsel

Pandas have black and white fur and are about 5 feet long. Male pandas can weigh up to 330 pounds.

Pandas use their teeth to peel off the tough outer layers of the bamboo stalk to reveal the soft inner tissue. They also eat the leaves.

The panda mother takes great care not to harm the baby panda. For several days after birth, the mother does not leave the den, not even to eat or drink!

156

Remember that as informational report writers, you'll want to include your information about pandas in a particular order. You can use the space below to plan or the graphic organizer on page 5:

0   You should start by introducing your topic in a sentence. Take a moment to think. What is your informational report going to be about?

0   You should include main ideas and supporting details. What were some of the main ideas in the passage? What were the supporting details that went with them? Take some time to think about how to organize these ideas.

0   You should also include words that help the reader follow your thinking, such as *for example* and *also*.

0   Finally, you should provide a conclusion.

Plan your informational report about pandas for "Time for Kids" here or on the graphic organizer:

EFFECT OF GENRE AND PROMPT
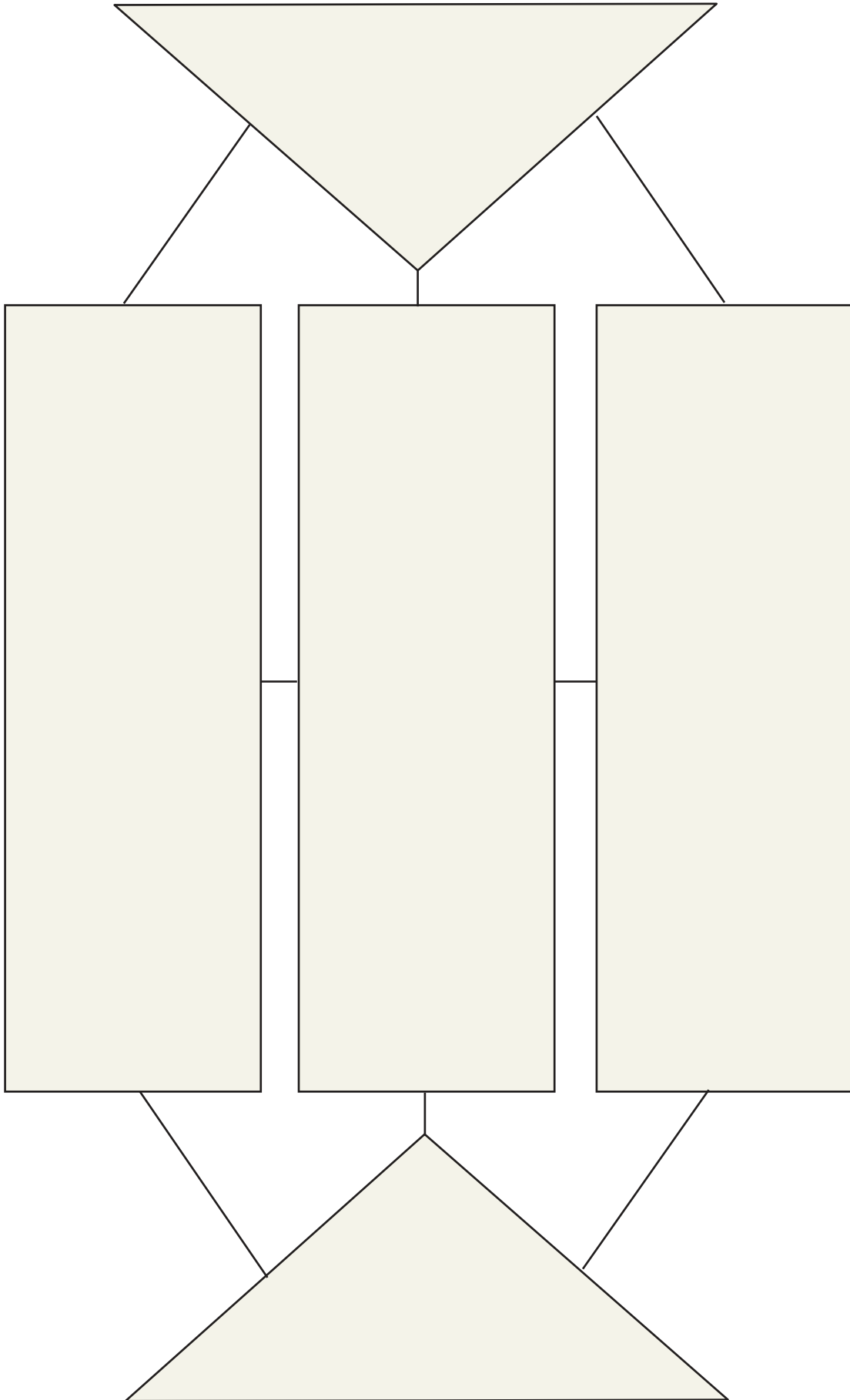
Write your informational report about pandas for "Time For Kids" here:

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____

Edit your paper to be sure that:

- ☐ You used good grammar;
- ☐ You used capital letters and punctuation marks correctly;
- ☐ You spelled words correctly; and
- ☐ You let your readers know where you started new paragraphs.
- ☐ You checked your paper to make sure that it is the way that you want readers to read it.

-5th Grade Student Copy-

Nutrition Information

Wendy's food is fast and inexpensive. Wendy's offers healthy options like baked potatoes and salads as well as "smart snacks" in smaller sizes. For example, instead of a regular sized burger, you can buy a junior burger, which has significantly fewer calories. Wendy's food is also made fresh to order with fresh ingredients. Unlike many other fast food restaurants, Wendy's does not pre-make their menu items with frozen ingredients.

Approximately 1/3 of Americans are overweight. One reason for this is the amount of fast food that Americans consume on a regular basis. Wendy's food is high in fat and calories. Even their junior bacon cheeseburger has 400 calories and 24 grams of fat! That's over half the total amount of fat a child should eat in one day. Children should not eat foods at Wendy's because such foods may lead to heart attacks and other serious health problems.

Side Selections	Nutrition	
	Total	
Garden Side Salad	25	0
Caesar Salad	60	3.5
Apple Slices	40	0
Plain Baked Potato	270	0

Side Selections	Nutrition	
	Total Calories	Fat (g)
Small Natural-Cut Fries*	320	16
Medium Natural-Cut	420	21
Fries*	530	25

Sandwiches Made when you order it.	Nutrition	
	Total Calories	
Grilled Chicken Go Wrap	260	10
Spicy Chicken Go	340	16

Sandwiches Made when you order it.	Nutrition	
	Total Calories Fat (g)	
Jr. Bacon Cheeseburger	400	24
Jr. Cheeseburger	350	19

162

You decide: Is Wendy's a good choice for kids?

Remember that as persuasive writers, you'll want to be sure to include information in a particular order.   You can use the space below to plan or the graphic organizer on page 4:

0   You should start by stating your opinion. Are you for Wendy's or against it?

0   You should include reasons and evidence to support your opinion. What are some of your reasons for or against Wendy's? What is the evidence that goes with those reasons? Take some time to think about how to organize these ideas.

0   You should also include words that help the reader follow your thinking, such as *for example* and *because*.

0   Finally, you should provide a conclusion that restates your opinion.

Plan your persuasive essay about Wendy's for the "Kid's Post" here or on the graphic organizer:
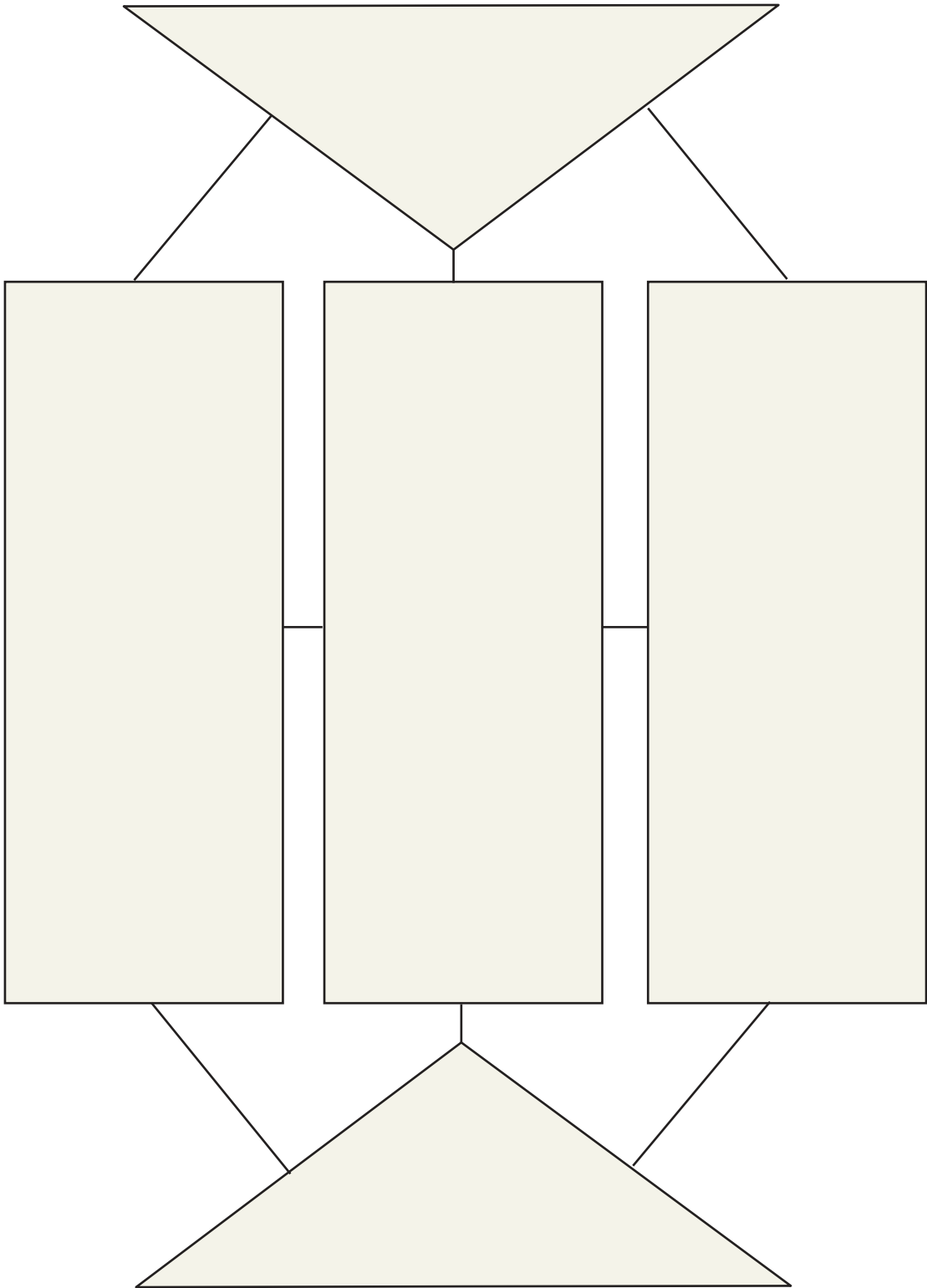
Write your persuasive essay for the "Kid's Post" on the lines below.
Is Wendy's a good choice for kids? Why or why not?

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____

Edit your paper to be sure that:

- ☐ You used good grammar;
- ☐ You used capital letters and punctuation marks correctly;
- ☐ You spelled words correctly; and
- ☐ You let your readers know where you started new paragraphs.
- ☐ You checked your paper to make sure that it is the way that you want readers to read it.

APPENDIX E: Prompt Directions

Informational Report **SUPPORTED** Condition Prompt

---Administration Directions---

1. Before starting the test, tell students: **"Please write your name, teacher, grade, and date on the front of this packet."**

2. Teacher says, **"When you are ready, turn to page 2. Imagine that you are a reporter for "Time for Kids" Magazine. You've just been assigned to write an informational report article about pandas for their latest edition on endangered animals. Before you being, I am going to read to you from a passage about pandas for five minutes. You may highlight and mark key information as I read. Then you will write an informational report about what you have learned."** Read all of the content in the passage including the diagram (read counter-clockwise) and the map (use the class color print out as needed). **"In the remaining time, you may reread the passage."**

3. After the 5 minutes are up the teacher says, "**Before you begin writing, you will have 5 minutes to plan your report in the space provided.** (Set timer for 5 minutes and read the following in approximately 1 minute intervals). **Turn to page 4. Remember that as informational report writers, you'll want to include your information about pandas**

in a particular order. You can use the space below to plan or the graphic organizer on page 5:

You should start by introducing your topic in a sentence. Take a moment to think. What is your informational report going to be about? Pause to allow students time to record their thoughts.

You should include main ideas and supporting details. What were some of the main ideas in the passage? What were the supporting details that went with them? Take some time to think about how to organize these ideas. Pause to allow students time to record their thoughts.

You should also include words that help the reader follow your thinking, such as *for example* **and** *also***.** Pause to allow students time to record their thoughts.

Finally, you should provide a conclusion. Pause to allow students time to record their thoughts.

3. Teacher says, "**You may now have 15 minutes to write your informational report for the Time for Kids article. I will warn you when you have 3 minutes left. You may begin now.**" At the 3 minute warning state "**You have 3 minutes left. Please begin to finish your writing. Remember to edit your paper to be sure that**

☐ **You used good grammar;**

- [ ] **You used capital letters and punctuation marks correctly;**

- [ ] **You spelled words correctly; and**

- [ ] **You let your readers know where you started new paragraphs.**

- [ ] **You checked your paper to make sure that it is the way that you want readers to read it.**

At the end of the 15 minutes say, "**Your 15 minutes are up, please put your pencils down and turn to the last page in your packet. I will now ask you to answer questions to a brief questionnaire. I will read the directions and questions to you."**

Read the questionnaire items on the last page. Then collect all writing packets.

Informational Report **UNSUPPORTED** Writing Prompt

---Administration Directions---

1. Before starting the test, tell students: **"Please write your name, teacher, grade, and date on the front of this packet."**

2. Teacher says, **"When you are ready, turn to page 2. Imagine that you are a reporter for "Time for Kids" Magazine. You've just been assigned to write an informational report article about pandas for their latest edition on endangered animals. First, you are going to read a passage about pandas for five minutes. You may highlight and mark key information as you read. Then you will write an informational report about what you have learned. You may begin now."**

3**.** After the 5 minutes are up the teacher says, "**Before you begin writing, you will have 5 minutes to plan your report in the space provided. Turn to page 4. You may begin now."**

4. After 5 minutes, teacher says, "**You will now have 15 minutes to write your report for the Time for Kids article. You may begin now."** When there are 3 minutes left say, "**You have 3 more minutes to write your report. Be sure to reread what you have written to make sure it makes sense. "** After the 15 minutes are over say, "**Your 15 minutes are up, please put your pencils down and turn to the last page in your packet. I will now ask you to answer questions to a brief questionnaire. I will read**

the directions and questions to you." Read the questionnaire items on the last page.

Then collect all writing packets.

Narrative **Supported** Writing Prompt

(25 minute administration time)

----Administration Instructions----

TO BEGIN ADMINISTRATION:

1. Before starting the test, tell students: **"Please write your name, teacher, grade, and date on the front of this packet."**

2. Begin testing by saying: **"The Washington Post Mini-Pages is having a story writing contest for a picture they have posted. In a few moments I am going to show you the picture they want you to write your story about. Open your booklet to page 2."** Show the students the Picture on the overhead or whiteboard. Say: **"I want you to write a story about this picture." Take the next five minutes to carefully look at the picture. Be sure to examine all of the details in the picture. Now, before you start, take some time to plan your story.** (Set timer for 5 minutes and read the following in approximately 1 minute intervals)**.**

Take a look at the picture. What do you think has happened? Can you create an interesting and exciting story about this picture? You can use the space below to plan or the graphic organizer on page 4.

Remember . . .

☐ **A good story has a beginning that includes the setting, characters, and an introduction to the problem. Take a moment to think about the characters. What are their names? What might his problem be?** Pause to give students time to record their thoughts.

☐ **A good story also has a middle that has interesting details and action. Take a moment to think about what the most exciting part of your story will be?**

What will the characters do and feel? Pause to give students time to record their thoughts.

☐ **Finally, a good story has an end with a great solution. How will your character solve his problem?** Pause to give students time to record their thoughts.

3. After 5 minutes have elapsed, say: **"You will have 15 minutes to write your story for the "Washington Post Mini-Page" contest. Use your imagination to make your story as interesting as you can. Also, use paragraphs, good spelling, and the right punctuation to make your story the best it can be. Remember to write neatly."** Pause, then say: **"Begin writing now."**

4. When 12 minutes have lapsed, say: "**You have 3 minutes left. Please begin to finish your writing. Remember to edit your paper to be sure that**

☐ **You used good grammar;**

☐ **You used capital letters and punctuation marks correctly;**

☐ **You spelled words correctly; and**

☐ **You let your readers know where you started new paragraphs.**

☐ **You checked your paper to make sure that it is the way that you want readers to read it.**

At the end of the 15 minutes say, "**Your 15 minutes are up, please put your pencils down and turn to the last page in your packet. I will now ask you to answer questions to a brief questionnaire. I will read the directions and questions to you."**

Read the questionnaire items on the last page. Then collect all writing packets.

Narrative **Unsupported** Writing Prompt

(25 minute administration time)

----Administration Instructions----

TO BEGIN ADMINISTRATION:

1. Before starting the test, tell students: **"Please write your name, teacher, grade, and date on the front of this packet."**

2. Begin testing by saying: **"The Washington Post Mini-Pages is having a story writing contest for a picture they have posted. In a few moments I am going to show you the picture they want you to write your story about. Open your booklet to page 2."** Show the students the Picture on the overhead or whiteboard. Say: **"I want you to write a story about this picture. Take the next five minutes to carefully look at the picture. Be sure to examine all of the details in the picture. Now, before you start, take some time to plan your story. Take 5 minutes to plan your story.**

3. After 5 minutes have elapsed, say: **"You will have 15 minutes to write your story for the "Washington Post Mini-Page" contest. Use your imagination to make your story as interesting as you can. Also, use paragraphs, good spelling, and the right punctuation to make your story the best it can be. Remember to write neatly."**
Pause, then say: **"Begin writing now."**

4. When 12 minutes have lapsed, say: **"You have 3 minutes to finish writing your story. Reread your writing to make sure it makes sense."** At the end of 15 minutes, say: **"Your 15 minutes are up, please put your pencils down and turn to the last page in your packet. I will now ask you to answer questions to a brief questionnaire. I will read the directions and questions to you."** Read the questionnaire items on the last page. Then collect all writing packets.

Persuasive Prompt: **Unsupported** Condition

(25 minute administration time)

(adapted from Teacher's College Reading and Writing Project Elementary Persuasive

Performance Assessment)

---Administration Directions---

1. Before starting the test, tell students: **"Please write your name, teacher, grade, and date on the front of this packet."**

2. Teacher says, "**When you are ready, turn to page 2. Imagine that you are a writer for "The Washington KidsPost." Your boss has asked you to write a persuasive essay about whether or not kids should eat at Wendy's. Before you write your essay, I'd like you to examine the following nutrition information from a Wendy's menu. You may have 5 minutes to read this information. Feel free to highlight important information and take notes as you read. You may begin now."** After the 5 minutes are up, please move on to administration item number 2.

3. Teacher says, "**Writers, you've done some good research now by reading and studying the nutrition information from a Wendy's menu. Next I'd like you to write your persuasive essay about whether or not kids should eat at Wendy's. Turn to page 3. Your job is to persuade the readers of the KidsPost about why people should or shouldn't go to Wendy's. Before you begin writing, you will have 5 minutes to**

plan your persuasive essay using the planning space in your booklet. You may begin now.

4. Teacher says, "**You may now have 15 minutes to write your persuasive essay for the KidsPost. I will warn you when you have 3 minutes left. You may begin now.**" At the 3 minute warning state "**You have 3 minutes left. Please begin to finish your writing. **" At the end of the 15 minutes say, "**Your 15 minutes are up, please put your pencils down and turn to the last page in your packet. I will now ask you to answer questions to a brief questionnaire. I will read the directions and questions to you.**"

Read the questionnaire items on the last page. Then collect all writing packets.

Persuasive Prompt: Supported Condition
(25 minute administration time)

---Administration Directions---

1. Before starting the test, tell students: **"Please write your name, teacher, grade, and date on the front of this packet."**

2. Teacher says, "**When you are ready, turn to page 2. Imagine that you are a writer for "The Washington KidsPost." Your boss has asked you to write a persuasive essay about whether or not kids should eat at Wendy's. Before you write your essay, I'm going to read a passage with nutrition information from a Wendy's menu to you while you read along. Feel free to highlight important information and take notes as I read."** Teachers should read the speech bubbles and read the tables with the nutritional information alternating from the table on the left to the table on the right for each food category. **"In the remaining time, you may reread the passage."** After 5 minutes, please move on to administration item number 3.

3. Teacher says, **Writers, you've done some good research now by reading and studying the nutrition information from a Wendy's menu. Next I'd like you to write your persuasive essay about whether or not kids should eat at Wendy's. Turn to page 3. Your job is to persuade the readers of the KidsPost about why people should or shouldn't go to Wendy's. Before you begin writing, you will have 5 minutes to plan your persuasive essay using the planning space in your booklet. Turn to page 3.**

(Set timer for 5 minutes and read the following in approximately 1 minute intervals)**. You decide: Is Wendy's a good choice?**

Remember that as persuasive writers, you'll want to be sure to include information in a particular order.   You can use the space below to plan or the graphic organizer on page 4:

You should start by stating your opinion. Are you for Wendy's or against it? Pause and give students time to record their thoughts.

☐ You should include reasons and evidence to support your opinion. What are some of your reasons for or against Wendy's? What is the evidence that goes with those reasons? Take some time to think about how to organize these ideas. Pause and give students time to record their thoughts.

You should also include words that help the reader follow your thinking, such as *for example* and *because*. Pause and give students time to record their thoughts.

Finally, you should provide a conclusion that restates your opinion. Pause and give students time to record their thoughts.

3. Teacher says, "**You may now have 15 minutes to write your persuasive essay for the KidsPost. I will warn you when you have 3 minutes left. You may begin now."** At

the 3 minute warning state "**You have 3 minutes left. Please begin to finish your writing. Remember to edit your paper to be sure that**

- ☐ **You used good grammar;**

- ☐ **You used capital letters and punctuation marks correctly;**

- ☐ **You spelled words correctly; and**

- ☐ **You let your readers know where you started new paragraphs.**

- ☐ **You checked your paper to make sure that it is the way that you want readers to read it.**

At the end of the 15 minutes say, "**Your 15 minutes are up, please put your pencils down and turn to the last page in your packet. I will now ask you to answer questions to a brief questionnaire. I will read the directions and questions to you."**

Read the questionnaire items on the last page. Then collect all writing packets.

APPENDIX F

Analytic Rubric

Student Name: _____

CATEGORY	4-Effective/Comprehensive	3-Effective	2-Developing	1-Underdeveloped
Organization and structure	Organization is a logical progression of ideas/events and is unified and complete.	There is a logical progression of ideas/events and is reasonably complete, although minor lapses may be present.	One or more major lapses in the logical progression of ideas/events is evident.	Ideas/events are presented in a random fashion.
Content focus	Maintains focus on topic/subject throughout response.	May exhibit minor lapses in focus on topic/subject.	May lose or may exhibit major lapses in focus on topic/subject.	May fail to establish focus on topic/subject.
STYLE	Consists of specific, developed details. Exhibits skillful use of vocabulary that is precise and purposeful. Demonstrates skillful sentence fluency (varies length, good flow rhythm, and varied structure).	Consists of some specific details. Exhibits reasonable use of vocabulary that is precise and purposeful. Demonstrates reasonable sentence fluency.	Consists of general and/or undeveloped details, which may be presented in a list-like fashion. Exhibits minimal use of vocabulary that is precise and purposeful. Demonstrates minimal sentence fluency.	Elaboration is sparse; almost no details. Lacks use of vocabulary that is precise and purposeful. Sentence fluency is lacking.
Conventions	Exhibits STRONG CONTROL of grammatical conventions appropriate to the writing task: sentence formation; standard usage including agreement, tense, and case; and mechanics including use of capitalization, punctuation, and spelling.	Exhibits REASONABLE CONTROL of grammatical conventions appropriate to the writing task: sentence formation; standard usage including agreement, tense, and case; and mechanics including use of capitalization, punctuation, and spelling.	Exhibits MINIMAL CONTROL of grammatical conventions appropriate to the writing task: sentence formation; standard usage including agreement, tense, and case; and mechanics including use of capitalization, punctuation, and spelling.ing agreement, tense, and case; and mechanics including use of capitalization, punctuation, and spelling.	LACKS CONTROL of grammatical conventions appropriate to the writing task: sentence formation; standard usage including agreement, tense, and case; and mechanics including use of capitalization, punctuation, and spelling.

APPENDIX G: Scoring Directions

PURPOSE: Determine quality of students' writing AND score reliably

1. You are to apply the scoring rubric (attached). Use essays from our training session as benchmark papers. Papers in the data set can become benchmark papers if you both decide they exemplify specific features of the rubric.

2. Each person will eventually read all papers and score all papers independently. As you decide on a score, write the number on the paper. Then provide a reason for this rating using 1-2 descriptors from the rubric. Write this directly on the student's paper.

3. Only read 5 papers at a time. As you read, write notes on the essay to help you determine a specific value on the rubric, so that you can use it accurately.

4. After you each finish reading the same 5 essays, share/report scores for each paper, one at a time. The first purpose of the sharing is so you both learn use the rubric in the same way. In other words, you both need to have agreement on what a "4" means – or what a "1" means. Second, you need to use these values consistently.

5. When you disagree (i.e., you each assign different scores to the same paper)

    A. First, talk about what led each of you to the original score. Perhaps one person missed something in the essay or another person credited the child's writing too much. Regardless, decide on the "true" score, and write a second, final score on the paper (use an arrow, → 4). Note: one person may keep her original score, but this helps us see each final score by the same designation.

    B. The second purpose of the discussion is to remind each other what each value of the rubric means, and whether each of you is evaluating the papers consistently. Use benchmark papers, recently scored papers, and recent decisions to remind each other of decisions in applying the rubric.

6. If you *have* made an "error" in scoring a paper, do not share scores for any more papers in a given set of five papers before checking the remaining papers for the same mistake. Change any score that reflects the same problem before continuing to report your values. Doing this will improve your reliability.

7. Start each new session with a review of benchmark papers and the most recently scored papers. Do not score too many papers in a given session, and take breaks to keep your focus.

APPENDIX H

Sentence Level Measure: Correct Word Sequence

Counting correct word sequences is one quantitative method of measuring and monitoring students' use of conventions. Correct word sequences (CWS) are two adjacent, correctly spelled words that are grammatically acceptable within the context of the phrase (Videen, Deno, & Marston, 1982). Capitalization and punctuation also can be considered within the sequence. To calculate the proportion of CWS:

1. Place a caret (^) over every correct sequence between the two words that form the sequence.

2. Place a large dot between every incorrect sequence. Place dots before and after misspelled words.
   Example: o my ^ dog o chasd o the ^ ball^.

3. The first sequence is not comprised of two words but marks how the sentence was begun. (Sentence beginning to first word "my" is marked as an incorrect sequence because the M is not capitalized.) The last sequence is the last word to period, question mark, or other appropriate ending punctuation.

4. To control for length of composition divide the number of CWS by the total number of sequences (total word count + 1), which gives the proportion of CWS.

REFERENCES

Afflerbach, P. (1986). The influence of prior knowledge on expert readers' importance assignment processes. In J.A. Niles & Lalik, R.V. (Eds.), *Solving problems in literacy: Learners, teachers, and researchers* (pp. 30-40). Rochester: National Reading Conference.

Afflerbach, P.P. (1990). The influence of prior knowledge on expert readers' main idea construction strategies. *Reading Research Quarterly, 25* (1), 31-46.

Alexander, P.A., Gillingham, M.G., & Kulikowich, J.M. (1990). *The role of interest in learning from text*. Paper presented at the annual meeting of the American Educational Research Association, Boston.

Alvermann, D.E., Smith, L.C., & Readence, J.E. (1985). Prior knowledge activation and the comprehension of compatible and incompatible text. *Reading Research Quarterly, 20*, 420-436.

Atwell, N, (1987). *In the middle.* Portsmouth NH: Boyton/Cook.

Barritt, L., Stock, P.L., & Claark, F. (1986). Researching practice: Evaluating assessment essays. *College Composition and Communication, 37,* 315-327.

Bazerman, C. (2009). Genre and cognitive development: Beyond writing to learn. In C. Bazerman, A. Bonini, & D. Figueiredo (Eds.), *Genre in a Changing World* (pp. 279-294). Colorado: The WAC Clearinghouse.

Beers, S.F., & Nagy, W.E. (2010) Writing development in four genres from grades three to seven: Syntactic complexity and genre differentiation. *Reading and Writing, 24*(2), 183-202.

Bereiter, C. (1980). Development in writing. In L.W. Gregg & E.R. Steinberg (Eds.), *Cognitive processes in writing*, Hillsdale, NJ: Erlbaum, 73-93.

Bereiter, C., and Scardamalia, M. (1987) *The psychology of written expression.* Hillsdale, NJ: Erlbaum.

Berninger, V.W. and Swanson, H.L., (1994) Modifying Hayes and Flower's model of skilled writing to explain beginning and developing writing. In Carlson, J.S. (series ed.) and Butterfield , E.C., (vol. ed.), *Advances in Cognition and Educational Practice, Vol. 2: Children's Writing: Toward a Process Theory of the Development of Skilled Writing,* JAI Press, Greenwich, CT. pp. 57-81.

Bakhtin, M. (1982). Forms of time and of the chronotope in the novel. In M. Holquist & V. Liapunov (Eds.) *The dialogic imagination: Four essays,* Austin: University of Texas Press.

Barritt L., Stock., P. L., & Clark, F. (1986).Researching practice: Evaluating assessment essays. *College Composition and Communication 3*(7), 315-327.

Bridges, L (2014). Beyond the surface: Data-driven whole-book comprehension instruction, assessment, and evaluation. USA: Scholastic Inc.

Brisk, M. E. (2012) Young bilingual writers' control of grammatical person in different genres. *The Elementary School Journal, 112* (3), 445-468.

Britton, B., & Tesser, A. (1982). Effects of prior knowledge on use of cognitive capacity in three complex cognitive tasks. *Journal of Verbal Learning and Verbal Behavior*, *21*, 421-436.

Brodney, B., Reeves, C., & Kazelskis, R. (1999). Selected prewriting treatments:Effects on expository compositions written by fifth-grade students. *The Journal of*

Experimental Education, 68(1), 5-20.

Brosell, G. (1983). Rhetorical specification in essay topics. *College English, 45*, 165-173.

Brossell, G., & Ash, B. H. (1984). An experiment with the wording of essay topics. *College Composition and Communication 3*(5) 423-425.

Bymes, J.P., & Guthrie, J.T. (1992). Prior knowledge and textbook search. *Contemporary Educational Psychology*, *17*, 8-29.

Calkins, L. M. (1986). *The art of teaching writing*. Portsmouth, NH: Heinemann.

Camp, R. (1993). Changing the model for the direct assessment of writing. In M. Williamson & B. Huot (Eds.), *Validating holistic scoring for writing assessment*, 45–78. Cresskill, NJ: Hampton Press.

Ceci, S.J., Caves, R.D., & Howe, M.J.A. (1981). Children's long-term memory for information that is incongruous with their prior knowledge. *British Journal of Psychology, 72*, 443-450.

Chall, J., Jacobs, V., & Baldwin, L. 1990. *The reading crisis: Why poor children fall behind*. Cambridge, MA: Harvard University Press.

Chall, J., & Jacobs, V. (2003, Spring). The classic study on poor children's fourth-grade slump. *American Educator*, *27*(1). Retrieved from http://www.aft.org/newspubs/periodicals/ae/spring2003/hirschsbclassic.cfm

Chenoweth, N.A. & Hayes, J.R., (2001). Fluency in writing: Generating text in L1 and L2. *Written Communication, 18,* 80-98.

Chiang, C.S., & Dunkel, P. (1992). The effect of speech modification, prior knowledge, and listening proficiency on EFL lecture learning. *Tesol Quarterly, 26* (2), 345-374.

Chiesi, H.L, Spilich, G.J., & Voss, J.F. (1979). Acquisition of domain related information in relation to high and low domain knowledge. *Journal of Verbal Learning and Verbal Behavior, 18*, 257-273.

Clifton, C., Jr., & Slowiaczek,M .L. (1981). Integrating new information with old knowledge. *Memory & Cognition, 9* (2), 142-148.

Colombi, M. C., & Roca, A. (2003). Insights from research and practice in Spanish as a heritage language. In A. Roca & M. C. Colombi (Eds.), *Mi lengua: Spanish as a heritage language in theUnited States; research and practice* (pp. 1–24). Washington, DC: Georgetown University Press.

Colombi, M. C.,&Schleppegrell, M. J. (2002). Theory and practice in the development of advanced literacy. In M. J. Schleppegrell &M. C. Colombi (Eds.), *Developing advanced literacy in first and second languages: Meaning with power* (pp. 1–19).

Cox, B.E., Shanahan, T. & Tinzmann, M.B. (1991). Children's knowledge of organization, cohesion and voice in written exposition. *Research in the Teaching of English, 25,* 179-218.

Crosson, A. C., Matsumara, L.C., Correnti, R., & Guerrero, A. (2012). The quality of writing tasks and students' use of academic language in Spanish. *The Elementary School Journal. 112*(3), 469-496.

Crowhurst, M. (1980). Syntactic complexity and teachers' ratings of narrations and arguments. *Research in the Teaching of English, 14*, 223-231.

Cummins, J. (1979). Linguistic interdependence and the educational development of bilingual children. *Review of Educational Research*, **49**, 222–251.

Daiute, C. (1984). Performance limits on writers. In Beach, R., and Bridwell, L.S. (Eds.), *New Directions in Composition Research,* The Guilford Press, New York, pp. 205-224.

Daiute, C.A. (1986). Physical and cognitive factors in revising: Insights from studies with computers. *Research in the Teaching of English, 20,* 141-158.

Darling-Hammond, L. (2006). No Child Left Behind and high school reform*. Harvard Educational Review, 76*(4), 642–667.

DeGroff, L.J. (1987). The influence of prior knowledge on writing, conferencing, and revising. *The Elementary School Journal, 88* (2), 105-118.

De La Paz, S., Swanson, P., & Graham, S. (1998). The contribution of executive control to the revising by students with writing and learning difficulties. *Journal of Educational Psychology, 90,* 448–460.

De La Paz, S. (1999). Composing via dictation and speech recognition systems: Compensatory technology for students with learning disabilities *Learning Disability Quarterly*, *22*(3), 173-182.

Dellerman, P., Coirier, P., & Marchand, E. (1996). Planning and expertise in argumentative composition. In G. Rijlaarsdam, H. can den Bergh, & M. Couzijn (Eds.), *Theories, models, and methodology in writing research* (pp. 182-195). Amsterdam: Amsterdam University Press.

Deno, S. L. (1985). Curriculum-based measurement: *Exceptional Children*, 52(3), 219-232.

Deno, S.L. (2003). Intervention assessment for effective. *Assessment for effective intervention, 28*(3), 3-12.

Dochy, F., Segers, M., & Buehl, M. (1999). The relation between assessment practices and outcomes of studies: The case of research on prior knowledge. *Review of Educational Research, 69* (2), 145-186.

Donovan, C. A. (2001). Children's development and control of written story and informational genres: Insights from one elementary school. *Research in the Teaching of English. 35*(3), 394-447.

Duke, N. K. (2000). 3.6 minutes per day: The scarcity of informational texts in the first grade. *Reading Research Quarterly, 35*(2), 202-224.

Engelhard, G., Gordon, B., & Gabrielson, S. (1992). The influences of modes of discourse, experiential demand, and gender on the quality of student writing. *Research on the Teaching of English. 26,* 315-336.

Englert, C.S., Raphael, T.E., Fear, K.L., & Anderson, L.M. (1988). Students' metacognitive knowledge about how to write informational texts. *Learning Disability Quarterly, 11*(4), 18-46.

Espin, C.A., De La Paz, S., Scierka, B.J., & Roelofs, L. (2005) The relationship between curriculum-based measures in written expression and quality and completeness of expository writing for middle school students. *The Journal of Special Education. 38*(4), 208-217.

Espin, C. A., Weissenburger, J. W., & Benson, B.J. (2004). Assessing the writing performance of students in special education. *Exceptionality, 12*(1), 55-66.

Faul, F., Erdfelder, E., Buchner, A., & Lang, A.G. (2009). Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods*, *41*, 1149-1160.

Fayol, M., Largy, P., and Lemaire, P. (1994). Cognitive overload and orthographic errors: When cognitive overload enhances subject-verb agreement errors in a study in French written language. *Quarterly Journal of Experimental Psychology, 47,* 437-464.

Fisher, D., Frey, N., & Lapp, D. (2012). *Text complexity: Raising rigor in reading.* Newark, DE: International Reading Association.

Fisher, D., & Frey, N. (2012). Close reading in elementary schools.*The Reading Teacher*, *66*(3), 179-188.

Freedman, S. W. (1984). The registers of student and professional expository writing: Influences on teachers' responses. In R. Beach & S. Bridwell (Eds.), *New Directions in Composition Research*(pp. 334-347). New York: Guilford Press.

Gabrielson, S., Gordon, B., & Engelhard, G. (1995) The effects of task choice on the quality of writing obtained in a statewide assessment. *Applied measurement in education, 8*(4), 273-290.

Garer, R., & Gillingham, M.G. (1992). Topic knowledge, cognitive interest, and text recall: A microanalysis. *Journal of Experimental Education, 59*, 310-319.

Gebhard, M. & Harman, R. (2011). Reconsidering genre theory in K-12 schools: A response to school reforms in the United States. *Journal of Second Language Writing, 20*, 45-55.

Goldman, S., Et Rueda, R. (1988). Developing writing skills in bilingual exceptional children. Exceptional Children, 54 (6), 543-551.

Graham, S. (1990). The role of production factors in learning disabled students' compositions. *Journal of Educational Psychology, 82,* 781-791.

Graham, S. (1997). Executive control in the revising of students with learning and writing difficulties. *Journal of Educational Psychology, 82,*781–791.

Graham, S. (2006). Writing. In P. Alexander & P. Winne (Eds.), *Handbook of Educational Psychology*. Mahwah, NJ: Erlbaum.

Graham, S. (2010). Writing matters: Overview of writing next and writing to read. Symposium conducted at the Center for Teaching and Learning Conference at the University of Oregon.

Graham, S., Berninger, V., Abbott, R., Abbott, S., & Whitaker, D. (1997). The role of mechanics in composing of elementary school students: A new methodological approach. *Journal of Educational Psychology, 89*, 170-182.

Graham, S. & Harris, K. (2000). The role of self-regulation and transcription skills in writing and writing development. *Educational Psychologist, 35*, 3-12.

Graham, S., Harris, K. R., & Fink, B. (2000). Is handwriting causally related to learning to write: Treatment of handwriting problems in beginning writers. *Journal of Educational Psychology.*

Graham, S., Harris, K, & Hebert M. (2011). Informing writing: The benefits of formative assessment. New York: Carnegie Corporation of New York.

Graham, S., Harris, K. R., MacArthur, C., & Schwartz, S. (1991). Writing and writing instruction with students with learning disabilities: A review of a program of research. *Learning Disability Quarterly, 14,* 89-114.

Graham, S., Harris, K. R., & Mason, L. (2005). Improving the writing performance, knowledge, and motivation of struggling young writers: The effects of self-regulated strategy development. *Contemporary EducationalPsychology, 30,* 207–

241.

Graham, S., & MacArthur, C. (1993). Integrating strategy instruction and word

processing into a process approach to writing instruction. *School Psychology*

Review. 22 (4) 671-82.

Graham, S., Schwartz, S., & MacArthur, C. (1993). Learning disabled and normally

achieving students' knowledge of the writing and the composing process, attitude

toward writing, and self-efficacy. *Journal of Learning Disabilities, 26,* 237-249.

Graham, S., & Weintraub, N. (1996). Review of handwriting research: Progress and

prospects from 1980 to 1994. *Educational Psychology Review, 8,* 7-82.

Graves, A. W., Valles, E. C., & Rueda, R. (2000). Variations in interactive writing

instruction: A study in four bilingual special education settings. Learning

Disabilities Research and Practice 15(1), 1-9.

Graves. D. H. (1989). Writing: Teachers and Children at Work. In D. S. Strickland and L.

M. Morrow (Eds.), *Emerging Literacy: Young Children Learn to Read and Write.*

Newark. Del.: International Reading Association.

Greenberg, K. (1981). *The effects of variations in essay questions on the writing*

performance of CUNY freshman. New York: The City University of New York

Instructional Resource Center.

Halliday, M. A. K. (1975). *Learning how to mean: Explorations in the development of*

language. London: Edward Arnold.

Hamilton, L. (2003). Assessment as a policy tool. In R. Floden (Ed.), *Review of*

research in education, 27, 25–68. Washington, DC: American Educational

Research Association.

Hasselhor, M., & K6rkel, J. (1986). Metacognitive versus traditional reading instructions: The mediating role of domain-specific knowledge on children's text-processing. *Human Learning, 5,* 75-90.

 Hayes, J. R. (1996). A new framework for understanding cognition and affect in writing. In M.C. Levey & S. Ransdell (Eds.), *The science of writing* (pp. 1-27). Mahwah, NJ: Erlbaum.

Hayes, J., & Flower, L. (1980). Identifying the organization of writing processes. In L. Gregg, & Steinberg, E. (Ed.), *Cognitive processes in writing* (pp. 3-30). NJ: Erlbaum: Hillsdale.

Hebert, M., Graham, S., & Harris, K. (2011, February). *Assessing struggling writers: A comparison of writing performance and behavior across four genres.* Paper session presented at the Writing Research Across Borders II Conference, George Mason University, Fairfax, VA.

Heim, S., Keil, A., Choudhury, N., Thomas Friedman, J. & Benasich, A. (2013). Early gamma oscillations during rapid auditory processing in children with a language-learning impairment: Changes in neural mass activity after training. Neuropsychologia, 51, 990-1001.

Heit, E. (1994). Models of the effects of prior knowledge on category learning. Journal of Experimental Psychology, 20 (6), 1264-1282.

Hidi, S. & Hildyard, A. (1983). The comparison of oral and written productions of two discourse types. *Discourse Processes, 6,* 91-105.

Hildreth, G. (1936). Developmental sequences in name writing. *Child Development, 7,* 291-302.

Hinkle, D. E., Wiersma, W., & Jurs, S. G. (2003). *Applied statistics for the behavioral sciences.* Belmont, CA: Wadsworth Cengage Learning.

Hoetker, J., & Brosell, G., (1989). The effects of systematic variations in essay topics on the writing performance of college freshman. *College Composition and Communication, 40*, 414-421.

Houck, C. & Billingsley, B. (1989). Written expression of students with and without learning disabilities: Differences across the grades. *Journal of Learning Disabilities. 22,* 561-572.

Houck, B. D. & Ross, K. (2012) Dismantling the Mylth of Learning to Read and Reading to Learn. *Reading: The core skill, 7* (11) .Retrieved from http://www.ascd.org/ascd-express/vol7/711-houck.aspx.

Howell, D. (2002). Statistical methods for psychology. Duxbury: Pacific Grove, CA

Hudson, R. F., Lane, H. B., & Mercer, C.D., (2005). Writing prompts: The role of various priming conditions on the compositional fluency of developing writers. *Reading and Writing, 18,* 473-495.

Hult, C. A. (1987). Assessment topics: The importance of the rhetorical frame. *Writing Program Administration 1* (0) 19-28.

Hunt, K. (1965). Grammatical structures written at three grade levels. *NCTE Research report No. 3.* Champaign, IL, USA: NCTE

Huot, B. (1990). Literature of direct writing assessment: Major concerns and prevailing trends. *Review of Educational Research, 60*(2), 237-263.

Hyland, K. (2009). Genre-based pedagogies: A social response to process. *Journal of Second Language Writing, 12*, 17-29.

Jeong, J. S., Gaffney, J. S., & Choi, J. O. (2010). Availability and use of informational text in second, third and fourth-grade classrooms. *Research in the Teaching of English. 44*(4): 435–456.

Jewell, J. & Malecki, C. K. (2005). The utility of CBM written language indices: An investigation of production-dependent, production- independent, and accurate-production scores. *School Psychology Review, 34,* 27-44.

Johnston, P., & Pearson, P.D. (1982). *Prior knowledge, connectivity, and the assessment of reading comprehension.* Technical report, no. 245. Cambridge: Center for the study of reading.

Kamberlis, G. (1999). Genre development and learning: Children writing stories, science reports, and poems. *Research in the Teaching of English. 33*(4), 403-460.

Kelley, K. (2007). Confidence intervals for standardized effect sizes: Theory, application, and implementation. *Journal of Statistical Software*, *20*(8), 1–24.

Keppel, G. & Wickens, T.D. (2004). *Design and analysis: A researcher's handbook* (3rd ed.). Upper Saddle River, NJ: Pearson.

King, M. L., & Rentel, V. (1979). Toward a theory of early writing development. *Research in the Teaching of English, 13*, (3), 243-253

Kline, R. B. (2009). *Becoming a behavioral science researcher: A guide to producing research that matters.* New York: Guilford Press.

Knudson, R. E. (1991) Effects of instructional strategies, grade, and sex on students' persuasive writing. *The Journal of Experimental Education. 59*(2), 141-152.

Laitusis, Buzick, Stone, Hansen and Hakkinen (2012). Smarter Balanced Assessment Consortium: Literature review of testing accommodations and accessibility tools

for students with disabilities. New Jersey: Educational Testing Service.

Lambiotte, J.G., & Dansereau, D.F. (1992). Effects of knowledge maps and prior

knowledge on recall of science lecture content. *Journal of Experimental*

Education, 60 (3), 189-201.

Lavoie, D.R. (1989*). Effects of prior knowledge and Piagetian cognitive devel-opment on*

the process skill of prediction in the learning cycle. Technical 178 Assessment

and Outcomes of Studies report. Detroit: Wayne State University.

Lee, O., & Buxton, C.A., (2011). Relationship between "form" and "content" in science

writing among English language learners. *Teachers College Record 113*(7), 1401-

1434.

Leu, D., Keech, C., Murphy, S., & Kinzer, C. (1982). Effects of variation in a writing test

prompt upon holistic score and other factors. In J. Gray (Ed.). *Properties of*

writing tasks: A study of alternate procedures for writing assessment, a final

report (pp. 15-23). Berkeley: University of California, Berkeley, College of

Education, Bay Area Writing Project. (ERIC Reproduction Service No. ED 230

576).

Lipson, M.Y. (1982). Learning new information from text: The role of prior knowledge

and reading ability. *Journal of Reading Behavior, 14* (3), 243-261.

Lloyd-Jones R. (1977). Primary trait scoring. In C. R. Cooper & L. Ode (Eds*.),*

Evaluating writing: Describing measuring, judging (pp. 3-32). Urbana: NCTE.

Loban, W. (1976) *Language development: Kindergarten through grade twelve.* (National

Council of Teachers of English Research Report No. 18). Urbana, Illinois:

National Council of Teachers of English.

Lehongre, K., Ramus, F., Villiermet, N., Schwartz, D., & Giraud, A. (2011) Altered

Low-Gamma Sampling in Auditory Cortex Accounts for the Three Main Facets

of Dyslexia. Neuron*, 72*, 1080–1090.

Lomax, R.G. (2007). *Statistical concepts: A second course.* (3rd ed.). Mahwah, NJ:

Lawrence Erlbaum Associates.

Luria, A. (1982). *Language and Cognition.* New York: John Wiley and Sons.

MacArthur, C.A., (1996). Using technology to enhance the writing processes of students

with learning disabilities. *Journal of Learning Disabilities, 29*(4), 344-354.

MacArthur, C. A. (2009). Reflections on research on writing and technology for

struggling writers. *Learning Disabilities Research & Practice, 24*(2), 93-103.

MacArthur, C. A., & Graham, S. (1987). Learning disabled students' composing under

three methods of text production: Handwriting, word processing, and dictation.

Journal of Special Education, 21, 22–42.

MacArthur, C. A., Graham, S., Haynes, J. A., & De La Paz, S. (1996). Spelling checkers

and students with learning disabilities: Performance comparisons and impact on

spelling. *Journal of Special Education, 30,* 35–57.

MacArthur, C. A., Graham, S.,&Schwartz, S. S. (1991).Knowledge of revision and

revising behavior among learning disabled students. *Learning Disability*

Quarterly, 14, 61–73.

MacArthur, C. A., Graham, S., & Schwartz, S. S. (1993). Integrating word processing

and strategy instruction into a process approach to writing. *School Psychology*

Review, 22, 671–681.

MacArthur, C. A., Graham, S., Schwartz, S. S., & Shafer,W. (1995). Evaluation of a

writing instruction model that integrated a process approach, strategy instruction, and word processing. *Learning Disability Quarterly,18,* 278–291.

Marshall, N. (1985). *The effects of prior knowledge and instruction on recall for well-written expository text*. In J.A. Niles & R.V. Lalik (Eds.),Issues in literacy: A research perspective (pp.89-96). Chicago, IL: National Reading Conference, Inc.

Marzano, R. J., Pickering, D. J., & Pollock, J. E. (2001). *Classroom Instruction that works: Research-based strategies for increasing student achievement*. Virginia: Association for Supervision and Curriculum Development. 72-83.

Mathews, S.R. (1982). *The impact of prior knowledge on accessibility and availability of information from prose.* Pensacola, U.S.A.: The University of West Florida, Educational Research and Development Center

McAndrew, D. A. (1982). The effects of an assigned rhetorical context on the syntax and holistic quality of the writing of first year college students. *Dissertation Abstracts International, 43,* 2911 A.

McCutchen, D. (2006). Cognitive factors in the development of children's writing. In C. MacArthur, S. Graham, & J. Fitzgerald, (Eds.), *Handbook on Writing Research..*(pp. 115-130). New York: The Guilford Press.

McCutchen, D. (2000). Knowledge, processing, and working memory: Implications for a theory of writing. *Educational Psychologist, 35*(1), 13-23.

McCutchen, D. (1996). A capacity theory of writing: working memory in composition. *Educational Psychology Review, 8*(3). 299-325.

McCutchen, D. (1988). "Functional Automaticity" in children's writing: A problem with metacognitive control. *Written Communication, 5*, 306-324.

McMaster, K. L.. & Campbell. H. (2008). New and existing curriculum based writing measures: Technical features within and across grades. *School Psychology Review, 37,* 550-566.

McMaster, K. L., & Espin, C. A. (2007). Technical features of curriculum-based measurement in writing: A literature review. *Journal of Special Education*, *41*, 68–84.

Mertler, C. A. & Vannatta, R. A. (2005). *Advanced and multivariate statistical methods: Practical application and interpretation.* California: Pryczak Publishing.

Midgette, E., Haria, P., & MacArthur, C. (2007). The effects of content and audience awareness goals for revision on the persuasive essays of fifth and eighth-grade students. *Reading & Writing*, *21,* 131-151.

Moffett, J. (1968). *Teaching the universe of discourse.* Boston: Houghton-Mifflin.

Monte-Sano, C. & De La Paz, S. (In Press). Using writing tasks to elicit adolescents' historical reasoning. *Journal of Literacy Research.*

Mosenthal, P. (1984). The effect of classroom ide-ology on children's production of narrative text. *American Educational R esearch J ournal, 21,* 679-689.

Mosenthal, P., Conley, M. W., Colella, A., & Davidson-Mosenthal, R. (1985). Prior knowledge and teacher ideology as determinants of text in classroom lessons. *Elementary School Journal, 85*, 1-13.

Murphy, S. & Camp, R. (1996). Toward systemic coherence: A discussion of conflicting perspectives on portfolio assessment. In R. Calfee & P. Perfumo (Eds.), *Writing portfolios in the classroom: Policy and practice, promise and peril*. Lawrence Erlbaum Associates.

Murphy, S. (2008). Some consequences of writing assessment. In A. Havnes & L. McDowell (Eds.) *Balancing dilemmas in assessment and learning in contemporary education*. Routledge.

Myhill, D. (2005). Testing times: The impact of prior knowledge on written genres produced in examination settings. *Assessment in Education,12*(3), 289-300.

Neuman, W.L. (1989). Which students learn the most, and why? A replication and extension of the Szafran pretest study. *Teaching Sociology, 17 (*1), 19-27.

Nelson, N. (2011, February). *Writing the news: Development of a standardized test of written language development.* Paper session presented at the Writing Research Across Borders II Conference, George Mason University, Fairfax, VA.

Nold, E. W. (1978, December). *The basics of research: The evaluation of writing*. Paper presented at the meeting of the Modern Language Association of America, New York. (ERIC Document Reproduction Service No. 166 713)

Nystrand, M. (2006). The social and historical context for writing research. In C. MacArthur, S. Graham, & J. Fitzgerald, (Eds.), *Handbook on Writing Research.* (pp.11-27). New York: The Guilford Press.

Olinghouse, N. & Graham, S. (2009). The relationship between the discourse knowledge and the writing performance of elementary-grade students. *Journal of Educational Psychology 101*(1), 37-50.

Olinghouse, N. & Santangelo, T., (2010). Assessing the writing of struggling writers. *Focus on Exceptional Children, 43*(4), 1-27.

Olinghouse, N. & Wilson, J. (2012). The relationship between vocabulary and writing quality in three genres. *Reading and Writing Research Quarterly, 26*, 45-65.

Ortmeier-Hooper, C. & Enright, K.A. (2011). Mapping new territory: Toward an understanding of adolescent L2 writers and writing in US contexts. *Journal of Second Language Writing, 20*, 167-181.

Partnership for Assessment of Readiness for College and Careers, (2013) .PARCC draft accommodations manual. Retrieved from http://www.parcconline.org/parcc-draft-accommodations-manual.

Parker, R., Tindal, G., & Hasbrouck, J. (1991a). Countable indices of writing quality: Their suitability for screening-eligibility decisions. *Exceptionality, 2,* 1-17.

Parker, R., Tindal, G., & Hasbrouck, J. (1991b). Progress monitoring with objective measures of writing performance for students with mild disabilities. *Exceptional Children, 58,* 61-73.

Persky, H. R., Daane, M. C., & Jin,Y. (2003). *The nation's report card:Writing 2002*. (NCES 2003–529). U.S. Department of Education. Institute of Education Sciences. National Center for Education Statistics.Washington, DC: Government Printing Office.

Piaget, J. (1955). *The language and thought of the child.* Cleveland, Ohio: Meridian Books.

Powell-Smith, K. A., & Shinn, M. R. (2004). *Administration and scoring of written expression curriculum-based measurement (WECBM) for use in general outcome measurement.* Retrieved from AimsWeb website: http://www.aimsweb.com/measures-2/writtenexpression-Cbm

Prater, D. & Padia, W. (1983). Effects of modes of discourse on writing performance in grades four and six. *Research in the Teaching of English, 17*(2), 127-134. '

Puma, V. D. (1986). *The effects of degree of audience intimacy on linguistic features and quality in the audience specified essays of first year college students*. Unpublished doctoral dissertation Indiana University of Pennsylvania.

Quellmalz, E.S., Capell, F. J., & Chou, C. C.(1982), Effects of discourse and response mode on the measurement of writing. *Journal of Educational Measurement, 19*(4) 241-258.

Rankin, J., Bruning, R., & Timme, V. (1994). The development of beliefs about spelling and their relationship to spelling performance. Applied Cognitive Psychology, 8, 213-232.

Redd-Boyd, T. M., & Slater, W. H. (1989). The effects of audience specification on undergraduates attitudes, strategies and writing. *Research in the Teaching of English, 23,* 77-108.

Reed, W. M., Burton, J. K., & Kelly, P. P. (1985). The effects of writing ability and mode of discourse on cognitive capacity engagement. *Research in the Teaching of English,19*, 283-297.

Reynolds, D. W. (2005). Linguistic correlates of second language literacy development: Evidence from middle-grade learner essays. *Journal of Second Language Writing, 14*, 19–45.

Rosemblum, S., Weiss, P.L., & Parush, S. (2003). Product and process evaluation of handwriting difficulties. *Educational Psychology Review, 15*(1) 41-81.

Rosen, H. (1969). *An investigation of the effects of differential writing assignments on the performance in English composition of a selected group of 15-to-16 year old pupils*. Unpublished doctoral dissertation University of London.

Roth, M. W. (2009). Realizing Vygotsky's program concerning language and thought: t

    tracking knowing (ideas, conceptions, beliefs) in real time, *Language and*

    Education, 23(4), 295-311.

Ruiz, N., Vargas, E., & Beltran, A. (2002). Becoming a reader and a writer in a bilingual

    special education classroom. *Language Arts, 79* (4) 297-309

Sapir, E. (1921). *Language an introduction to the study of speech.* New York: Harcourt

    Brace Jovanovich.

Sanbonmatsu, D.M., Kardes,F .R., & Herr,P .M. (1992). The role of prior knowledge and

    missing information in multiattribute evaluation. *Organizational Behavior and*

    Human Decision Processes, 51, 1-16.

Scarcella, R. (2002). Some key factors affecting English learners' development of

    advanced literacy. In M. J. Schleppegrell & M. C. Colombi (Eds.), *Developing*

    advanced literacy in first and second languages: Meaning with power (pp. 209–

    226). Mahwah, NJ: Erlbaum.

Scarcella, R. (2003). *Academic English: A conceptual framework* (University of

    California Linguistic Minority Research Institute Technical Report No. 2003-1).

    Retrieved from http://lmri.ucsb.edu/publications/03_scarcella.pdf

Schiefele, U. (1990). *The influence of topic interest, prior knowledge, and cognitive*

    capabilities on text comprehension. In J.M. Pieters, K. Breuer,& P.R.J. Simons

    (Eds.), Learning Environments. Berlin: Springer-Verlag

Schleppegrell, M. J. (2004). *The Language of schooling: A functional linguistics*

    perspective. Mahwah, NJ: Erlbaum.

Scott, C. M., & Windsor, J. (2000). General language performance measures in spoken

and written narrative and expository discource of school-age children with language learning disabilities. *Journal of Speech, Language, and Hearing Research, 43*, 324-339.

Shadish, W., Cook, T. & Campbell, D. (2002). *Experimental and quasi-experimental designs for generalized causal inference*. Boston: Houghton Mifflin.

Slavin, R., Madden, N., & Karweit, N. (1989). Effective programs for students at risk: Conclusions for practice and policy. In R. Slavin, N. Karweit & N. Madden (Eds.), *Effective programs for students at risk* (pp. 21-54). Boston: Allyn & Bacon.

Smith, W. L., Hull, G.A., Land, R. E. , Moore, M.T., Ball, C., Dunham, D.E., Hickey, L., S., & Ruzich, C. W. (1985). Some effects of varying the structure of a topic on college students' writing. *Written communication. 2, 73-89.*

Spandel, V., & Stiggins, R. J. (1980). *Direct measures of writing skil: Issues and applications*. Portland, OR: Northwest Regional Educational Development Laboratory.

Spiro, R. J. & Taylor, B. M. (1980). *On investigating children's transition from narrative to expository discrouse: The multidimensional nature of psychological text classification* (Report No. 195), Center for the Study of Reading. Urbana, IL: University of Illinois. (ERIC document Respoduction Service No. ED 199 666).

Stevenson, M., Schoonen, R., Kees de Glopper (2006). Revising in two languages: A multi-dimensional comparison of online writing revisions in L1 and FL. Journal of Second Language Writing 15, 201–233.

Stewart, M. F., & Grobe, C.H. (1979). Syntactic maturity, mechanics of writing, and

teacher's quality ratings. *Research in the Teaching of English, 13*(3), 207-215.

Strickland, D. S., & Alvermann, D. E. (Eds.). (2004). Bridging the literacy achievement gap grades 4–12. New York: Teachers College Press.

Stock, P. L., & Robinson, J. L. (1987). Taking on testing: Teachers as testers researchers. *English Education, 19*, 93-121.

Swales, J. (2009). Worlds of genre—Metaphors of genre. In C. Bazerman, A. Bonini, & D. Figueiredo (Eds.), *Genre in a Changing World* (pp. 3-16). Colorado: The WAC Clearinghouse.

Tabacnick, & Fidell (1996). *Using multivariate statistics* (3rd ed.). New York: Harper Collins.Teachers of English to Speakers of Other Languages. (2009). *ESL standards for pre-K–12 students*. Retrieved from http://www.tesol.org/s_tesol/seccss.asp?CID_95&DID_1565

Tindal, G., & Parker, R. (1989). Assessment of written expression for students in compensatory and special education programs. *Journal of Special Education, 23,* 169-183.

Troia, G. A. (1999). Phonological awareness intervention research: A critical review of the experimental methodology. *Reading Research Quarterly, 34*(1), 28-52.

Troia, G. (2006). Writing instruction for students with learning disabilities. In C. MacArthur, S. Graham, & J. Fitzgerald, (Eds.), *Handbook on Writing Research.*(pp. 324-336). New York: The Guilford Press.

Veal, L. R., & Hudson, S . A.(1983). *Direct and indirect measures of large-scale evaluation of writing*. Research in the Teaching of English, 17, 285-296.

Vygotsky, L. S. (1962). *Thought and Language.* Massachusetts: Massachusetts Institute

of Technology.

Vygotksy, L. (1987). *Problems of general psychology: the collected works of L.S.*

Vygotsky, volume 1. (N. Minick, trans.). New York, NY: Plenum.

Weissenburger, J.W., & Espin, C.A. (2005). Curriculum-based measures of writing

across grade levels. *Journal of School Psychology, 43,* 153–169.

Wells, G. (1986). *The meaning makers: Children learning language and using language*

to learn. Portsmouth, NH: Heinemann.

Willoughby, T., Waller, T.G., Wood, E., & MacKinnon, G.E. (1993). The effect of prior

knowledge on an immediate and delayed associative learning task following

elaborative interrogation. *Contemporary Educational Psychology, 18* (1), 36-46.

Willoughby, T., Wood, E., & Khan, M. (1994). Isolating variables that impact on or

detract from the effectiveness of elaboration strategies. *Journal of Educational*

Psychology, 86 (2), 279-289.

Yopp, R., & Yopp, H. (2006). Informational texts as read-alouds at school and home.

Journal of Literacy Research. March 38(1), 37–51.