

Investigations of the Role of High-Level Cognitive Skills in the Text Production Process

by

Laura K. Allen

A Thesis Presented in Partial Fulfillment
of the Requirements for the Degree
Master of Arts

Approved November 2014 by the
Graduate Supervisory Committee:

Danielle McNamara, Chair
Carol Connor
Arthur Glenberg

ARIZONA STATE UNIVERSITY

December 2014

ABSTRACT

Writing is an intricate cognitive and social process that involves the production of texts for the purpose of conveying meaning to others. The importance of lower level cognitive skills and language knowledge during this text production process has been well documented in the literature. However, the role of higher level skills (e.g., metacognition, strategy use, etc.) has been less strongly emphasized. This thesis proposal examines higher level cognitive skills in the context of persuasive essay writing. Specifically, two published manuscripts are presented, which both examine the role of higher level skills in the context of writing. The first manuscript investigates the role of metacognition in the writing process by examining the accuracy and characteristics of students' self-assessments of their essays. The second manuscript takes an individual differences approach and examines whether the higher level cognitive skills commonly associated with reading comprehension are also related to performance on writing tasks. Taken together, these manuscripts point towards a strong role of higher level skills in the writing process and provide a strong foundation on which to develop future research and educational interventions.

TABLE OF CONTENTS

	Page
LIST OF TABLES.....	IV
CHAPTER	
1 INTRODUCTION.....	1
Current State of Writing Education and Research.....	3
Thesis Project	6
2 STUDY 1	10
Introduction	10
Evaluative Misalignment of Student and Teacher Writing Evaluation Criteria	12
Automated Textual Analyses to Explore Student and Teacher Evaluative Misalignment.....	19
Method.....	24
Results	28
Discussion.....	40
3 STUDY 2	47
Introduction	47
Theoretical Models of Reading Comprehension and Writing.....	49
Connections between Reading Comprehension and Writing Proficiency ...	50
Current Study.....	59
Method.....	60

CHAPTER	Page
Results.....	67
Discussion.....	74
Conclusions	80
4 GENERAL DISCUSSION.....	82
REFERENCES	87

LIST OF TABLES

Table	Page
1. Correlations between Teacher Ratings and Coh-Metrix Variables – Study 1.....	28
2. Correlations between Student Ratings and Coh-Metrix Variables – Study 1.....	32
3. Correlations between Teacher Ratings and LIWC Variables – Study 1.....	35
4. Correlations between Student Ratings and LIWC Variables – Study 1.....	37
5. Descriptive Statistics and Pearson Correlations for Measures Related to Reading and Writing – Study 2.....	68
6. Linear Regression Analysis for Component Processes Predicting Reading Comprehension Scores and Writing Scores – Study 2.....	72
7. Linear Regression Analysis for Measures Predicting Reading Comprehension Scores and Writing Scores – Study 2.....	73

Chapter 1

INTRODUCTION

Writing is an intricate cognitive and social process that involves the production of texts for the purpose of conveying meaning to others (Graham, 2006; Hayes, 1996).

Particularly with respect to academic writing, this process relies on the interplay between both lower level and higher level processes (Flower & Hayes, 1981; Hayes, 1996). Lower level knowledge and processes, such as vocabulary strength and complex sentence development, relate to a writer's general understanding of the language in which they are writing, and are a necessary aspect of the writing process (e.g., Abbott & Berninger, 1993; Fitzgerald & Shanahan, 2000). The successful production of coherent texts, however, does not necessarily emerge once these skills have been developed. Rather, strong and persuasive writing relies on an individual's ability to strategically use language skills to connect ideas and present information in a way that is meaningful to the audience (Donovan & Smolkin, 2006; Graham, 2006; McNamara, 2013; Wong, 1999).

The importance of lower level cognitive skills and language knowledge during the text production process has been well documented in the cognitive literature (Abbott & Berninger, 1993; Fitzgerald & Shanahan, 2000; Graham, 2006; Kellogg, 2001; McCutchen, 1996). Amongst cognitive researchers, however, the role of higher level skills (e.g., metacognition, strategy use, etc.) has been less strongly emphasized. Working memory capacity, for example, has received considerable attention from researchers and is commonly labeled as a central component of the writing process (Berninger & Swanson, 1994; Hayes, 1996; Kellogg, 1996; 2001; 2008; Kellogg, Whiteford, Turner,

Cahill, & Mertens, 2013; McCutchen, 1996). This measured “capacity” of an individual’s working memory is theorized to relate to their writing ability because of the complex and resource-demanding nature of the task. Similarly, the depth of writers’ word knowledge has been identified as key for their production of high-quality texts. In particular, strong vocabulary knowledge has been linked to an increased sophistication in the words and surface level features found within students’ essays (Fitzgerald & Shanahan, 2000; Shanahan & Lomax, 1986).

Despite the stronger emphasis that has been placed on lower level skills within the cognitive literature, evidence from *educational* research suggests that the development and use of higher level skills (e.g., strategies) can significantly reduce the demands of the writing process and enhance writing performance. Saddler and Graham (2007), for example, found that weaker writers had a less sophisticated understanding of the goals for writing ($d = -1.13$), the differences between low- and high-quality writing ($d = -.98$), and efficient writing strategies ($d = -1.10$). Additionally, in a meta-analysis conducted on over 120 published studies of writing interventions, Graham and Perin (2007) found that strategy instruction was the most effective form of writing instruction, demonstrating an average weighted effect size of 0.82. Given these findings, it is clear that the development and use of higher level cognitive skills is important for the successful production of texts. A critical area of research, therefore, lies in the identification of these skills, as well as in the examination of the ways in which they influence the writing process. Findings from such research studies have the potential to inform theoretical models of text production (Bereiter & Scardamalia, 1987; Chenoweth & Hayes, 2001; 2003; Flower & Hayes, 1981; Galbraith, 1999; Hayes, 1996; Kellogg, 1994; 1996;

Torrance & Galbraith, 2006), as well as pedagogical interventions that are intended for struggling writers (Harris, Graham, Mason & Friedlander, 2008; Kieft, Rijlaarsdam, Galbraith, & van den Bergh, 2007; Roscoe & McNamara, 2013).

Current State of Writing Education and Research

Within the classroom, the development of strong writing skills is a critical (and somewhat obvious) goal. Individuals across the world are now expected to reach a high level of writing proficiency to achieve success in both academic settings and in the workplace (Geiser & Studley, 2001; Powell, 2009; Sharp, 2007). Over 90% of mid-career professionals agree that writing is essential for their day-to-day work (Light, 2001), and employees in virtually every area report that writing skills are critical for entry and promotion within their disciplines (Porter, 1997; Sharp, 2007). Further, in a recent survey of 120 American corporations, the National Commission on Writing (2004) found that writing was considered a threshold skill for the decision to hire or promote their salaried employees. Unfortunately, strong writing skills are difficult to develop, as they require individuals to coordinate a number of cognitive skills and knowledge sources through the process of setting goals, solving problems, and strategically managing their memory resources (Flower & Hayes, 1981; Hayes, 1996). Given the difficulty of this process, students frequently underachieve on national assessments of writing proficiency (National Assessment of Educational Progress, 2007; 2009; 2011).

Despite its documented importance, writing is a regularly undervalued domain among both educators and researchers. In the classroom, there is often little focus placed on the explicit instruction of the writing process, nor on the skills and knowledge that are necessary for developing proficiency in this domain. Further, when writing is taught to

students, the primary focus of instruction tends to be placed on surface-level text characteristics, such as the rules for constructing grammatical sentences and avoiding misspellings (Andrews et al., 2006; Crossley, Kyle, Allen, & McNamara, 2014). Instruction on the skills and strategies that are needed to develop strong and coherent arguments within texts, however, is commonly overlooked.

The writing process has been similarly neglected amongst education and cognitive researchers. Compared to other (similar) domains, such as text comprehension, little information is known about the writing process, including the range of individual differences that influence expert and novice writers or the interventions that lead to the greatest improvement in students' writing performance (National Commission on Writing, 2004). More recently, however, the emphasis on writing has been rapidly increasing amongst researchers, educators, and policy makers alike. With the introduction of the Common Core State Standards (CCSS; National Governors Association & Council of Chief School Officers, 2010), students are now expected to receive writing instruction at each grade level and, ultimately, to exit high school with strong writing skills. As these standards are integrated into schools across the country, it is increasingly important that educators and researchers work together to better understand the components of the writing process, and ultimately use this knowledge to improve students' performance on writing tasks.

The majority of the pedagogical techniques and interventions that have been proposed to writing instructors have been informed by research housed within the cognitive framework. According to this perspective, writing is a problem-solving activity that requires individuals to coordinate their linguistic knowledge, thoughts, and audience

expectations for the purpose of developing coherent texts. A number of models have been developed to describe the specific cognitive components that are involved in this text production process (Hayes, 1996; Flower & Hayes, 1981; Bereiter & Scardamalia, 1987). Hayes and Flower (1980), for instance, developed what was perhaps the most influential framework of writing in which they described the non-linear nature of the writing process and placed an emphasis on the multiple levels of information that are available to be processed during these tasks. A critical component of this model was the acknowledgement that the writing process is interactive – thus, the specific cognitive tasks involved in writing (e.g., planning, translating, reviewing) both influence and are influenced by other aspects of the writing task, such as the environment (e.g., the audience and the assignment) and the information that is in the writer’s long-term memory (e.g., prior world knowledge). This original model was later updated based on more contemporary research that investigated the role of cognitive skills and affective states on writing performance. Importantly, the new version of the model (Hayes, 1996) featured the working memory system as the centralized component of the writing process.

Since the original Hayes and Flower (1980) model, a number of models and frameworks have been proposed to account for the various components that make up the writing process (Bereiter & Scardamalia, 1987; Chenoweth & Hayes, 2001; 2003; Galbraith, 1999; Kellogg, 1996; Torrance & Galbraith, 2006). Despite the significant impact of this work, however, there has been a reduced emphasis on the role of higher level skills – particularly skills related to individuals’ strategic and metacognitive skills. This presents a significant gap in the literature on the writing process. *Instruction* on

higher level skills, such as writing strategies, has been shown to have a profound impact on students' ability to produce high-quality texts (e.g., Graham & Perin, 2007). However, it is unclear how these skills influence students' performance and what role they play in the processes enacted during writing. An important characteristic of these higher level skills is that they are generally *malleable*. Thus, given that one of the principal goals of writing research is to inform writing pedagogy, these skills may be particularly critical areas of study. If researchers can identify and better understand the role of higher level cognitive skills in the writing process, interventions can be developed that will be more effective at improving students' writing performance.

Thesis Project

The proposed thesis project is comprised of two published journal manuscripts. In both manuscripts, higher level cognitive skills are examined within the context of persuasive essay writing. In the first study, we target *metacognition* as an important skill for the writing process. In particular, we examine students' self-assessments of their work and whether these assessments are based on similar or different criteria as those made by their teachers. In the second study, we investigate the potential overlap between the text comprehension and text production processes. Specifically, we examine how individual differences commonly related to reading comprehension ability overlap and contribute to students' writing proficiency. Taken together, these manuscripts point towards a strong role of higher level skills in the writing process and provide a strong foundation on which to develop future research and educational interventions.

Chapter 2. The second chapter of this thesis (i.e., the first published manuscript) comprises a study entitled, "Evaluative Misalignment of 10th-Grade Student and Teacher

Criteria for Essay Quality: An Automated Textual Analysis.” This manuscript was published at the Journal of Writing Research and authored by Laura K. Allen (Varner), Rod D. Roscoe, and Danielle S. McNamara. The abstract for this manuscript is provided below:

Writing is a necessary skill for success in the classroom and the workplace; yet, many students are failing to develop sufficient skills in this area. One potential problem may stem from a misalignment between students’ and teachers’ criteria for quality writing. According to the *evaluative misalignment hypothesis*, students assess their own writing using a different set of criteria from their teachers. In this study, the authors utilize automated textual analyses to examine potential misalignments between students’ and teachers’ evaluation criteria for writing quality. Specifically, the computational tools Coh-Metrix and Linguistic Inquiry and Word Count (LIWC) are used to examine the relationship between linguistic features and student and teacher ratings of students’ prompt-based essays. The study included 126 students who wrote timed, SAT-style essays and assessed their own writing on a scale of 1-6. Teachers also evaluated the essays using the SAT rubric on a scale of 1-6. The results yielded empirical evidence for student-teacher misalignment and advanced our understanding of the nature of students’ misalignments. Specifically, teachers were attuned to the linguistic features of the essays at both surface and deep levels of text, whereas

students' ratings were related to fewer overall textual features and most closely associated with surface-level features.

Chapter 3. Chapter three comprises a study entitled “Reading Comprehension Components and their Relation to Writing.” This manuscript has been accepted to *Topics in Cognitive Psychology* and was authored by Laura K. Allen, Erica L. Snow, Scott A. Crossley, G. Tanner Jackson and Danielle S. McNamara. The abstract is provided below:

Within the educational community, research on student literacy often combines reading and writing measures, as they are presumed to draw on similar skills and background knowledge. However, relatively few studies have investigated the underlying cognitive skills required for both activities, and to what extent the required background knowledge and cognitive processes overlap. The current study investigates how individual differences commonly related to reading comprehension ability overlap and contribute to students' writing proficiency. University students ranging from 19 to 37 years old ($n = 108$) completed assessments to examine their reading comprehension and writing skills, as well as their vocabulary knowledge, lower level cognitive skills (working memory), and higher level cognitive skills (text memory, text inferencing, knowledge access, knowledge integration). Results indicated that reading comprehension was strongly related to both vocabulary knowledge and the higher level cognitive skills. Further, writing ability was moderately associated with a subset of the measured variables, namely vocabulary knowledge and the ability to access prior knowledge. These results

support the hypothesis that reading comprehension and writing share common knowledge sources and higher level cognitive skills, although the writing process is much less reliant on these measured variables than reading comprehension.

The culmination of these two projects suggests that higher level cognitive skills are critical components of the writing process. Although previous educational research suggests that explicit instruction on skills related to metacognition and strategy use has a positive impact on students' writing performance (e.g., Graham & Perin, 2007), these higher level skills have been relatively neglected within the cognitive research on the writing process. The combined results of this thesis are some of the first to investigate the role of higher level skills in the writing process, particularly from multiple perspectives (i.e., fine-grained text analysis, individual differences analysis).

Chapter 2

“Evaluative Misalignment of 10th-Grade Student and Teacher Criteria for Essay Quality: An Automated Textual Analysis”

Published at the Journal of Writing Research

Writing skills are essential for success, both in college and in the workplace (Geiser & Studley, 2001; Powell, 2009; Light, 2001). Unfortunately, national assessments reveal a lack of writing proficiency among high school students. According to the 2007 National Assessment of Educational Progress (NAEP) report, only 33% of United States 8th-graders and 24% of 12th-graders scored at or above the “proficient” level for writing, and only 2% of 8th-graders and 1% of 12th-graders scored at advanced levels for writing. One contributing factor to this problem may be an *evaluative misalignment* between student and teacher writing criteria. Such misalignments have been reported anecdotally in a number of writing studies (Lin, Monroe, & Troia, 2007; Kos & Maslowski, 2001; Hillocks, 1986). According to the evaluative misalignment hypothesis¹ explored here, students do not have an accurate conceptualization of the goals and criteria for quality writing. Their criteria may diverge markedly from the expectations of their instructors. As a result, students produce texts that fail to satisfy the demands of a given genre or assignment and potentially misunderstand teacher evaluations, feedback, and recommendations.

Few studies have sought to explore the phenomenon of *evaluative misalignment* empirically. A key challenge to such work is the highly subjective nature of writing assessment, including both teachers’ assessments of student work and students’ own self-

¹ Not to be confused with alignment in conversational discourse, *evaluative misalignment* refers to a discrepancy between students’ and teachers’ criteria for judging essay quality.

assessments. Ratings of essay quality, for example, are influenced by a myriad of linguistic, syntactic, semantic, and rhetorical features of text (Crossley & McNamara, 2011; McNamara, Crossley, & McCarthy, 2010), but human raters may not always be explicitly aware of the impact of such features on their judgments. In this study, we adopt the novel approach of examining the misalignment of student and teacher ratings of essay quality via *automated textual analysis*. Based upon innovations in computational linguistics and computer science, a number of computational tools now exist that enable researchers to quickly and objectively analyze texts at a fine-grained level for diverse textual features. In this research, we use such tools to analyze the underlying text features associated with student and teacher essay ratings. Subsequently, student and teacher misalignment is revealed by the extent to which their writing quality judgments are attuned to overlapping or divergent sets of textual features.

Our analyses utilize two tools: *Coh-Metrix* (Graesser, McNamara, Louwerse, & Cai, 2004; McNamara & Graesser, 2012) and *Linguistic Inquiry and Word Count* software (LIWC; Pennebaker, Booth, & Francis, 2007). Both of these tools provide measures of textual features that include surface-level components of text and measures related to deeper levels of cohesion and comprehension. Coh-Metrix offers a broad analysis of texts, including indices of cohesion, and text difficulty at the lexical, syntactic, structural, and global levels of text (Graesser & McNamara, 2011). LIWC provides several basic text measures (e.g., number of words and paragraphs), along with word-based analyses of lexical, semantic, and thematic properties of text. As Coh-Metrix and LIWC incorporate both overlapping and unique indices, their combination allows us to obtain converging evidence regarding the misalignment of student and teacher

assessments of writing quality. Additionally, these tools may be able to highlight the more nuanced features of students' and teachers' evaluation criteria that may be more difficult to detect otherwise.

Evaluative Misalignment of Student and Teacher Writing Evaluation Criteria

Students' writing problems cannot be accounted for simply by a lack of instructor expertise or empirical research regarding effective methods for teaching writing. There is ample research on composition instruction, and this work spans a variety of age groups, techniques, first- and second-language proficiencies, and individual differences among students (e.g., Graham & Perin, 2007; Hillocks, 1984; Rogers & Graham, 2008).

Exemplifying this research base, Graham and Perin (2007) conducted a meta-analysis of over 120 published studies on writing interventions among students in grades 4-12.

Interventions were separated into 11 categories: strategy instruction, summarization, peer assistance, setting product goals, word processing, sentence combining, inquiry, prewriting activities, process writing approach, study of models, and grammar instruction. All intervention types, with the exception of grammar instruction, were found to be beneficial (i.e., moderate to large, positive effect sizes). Rogers and Graham (2008) extended this analysis by conducting a separate meta-analysis of studies involving single-subject design writing interventions. They identified nine writing treatments that were proven to be effective in the writing classroom. In sum, research on the instruction of writing is extensive, and has yielded valuable insights and diverse methods for improving the efficacy of writing instruction. Yet, despite the availability of research on the topic, national assessments still reveal that students are struggling to excel in the writing domain.

One caveat for writing instruction is that there is little guarantee that students will internalize or understand the curriculum as intended. Students can ignore key aspects of instruction and fail to develop complete or accurate conceptions of writing goals and criteria. Moreover, when attempting to assess whether their own writing has achieved particular goals, students may judge themselves using limited or faulty criteria. As a result, students may not possess or apply the same evaluation criteria for writing as do their teachers – there may be an evaluative misalignment between student and teacher expectations. This misalignment can serve as a barrier to writing instruction, as students may not internalize the same information that teachers are intending to communicate.

Research on the knowledge of writing has observed that many students, particularly struggling writers, indeed display a lack of knowledge and understanding about key writing goals and processes. Wong (1999) describes this metacognitive knowledge as the “awareness of the purpose and process of writing and self-regulation of writing.” Such knowledge has been linked to writing proficiency in numerous studies (e.g., Donovan & Smolkin, 2006; Graham, 2006; Wong, 1999). A recurring pattern is that skilled writers are more knowledgeable about writing, particularly the higher level principles, such as audience awareness and the development and defense of strong arguments. In contrast, struggling writers tend to think of “better” writing in terms of superficial textual features, such as handwriting, spelling, and punctuation.

For example, Schoonen and de Glopper (1996) investigated the role of writing knowledge in the writing performance of 9th-grade students across three proficiency levels (i.e., low, average, proficient). They instructed older students to write letters to younger students describing the components and features that comprise good writing.

Their analysis of these documents indicated that proficient writers focused more on the organization of ideas, whereas less-proficient writers focused on surface-level features (e.g., spelling and grammar). Similarly, Lin, Monroe, and Troia (2007) investigated the development of writing knowledge in 2nd- and 8th-grade struggling and typical writers. In an interview setting, students were asked questions regarding their perspectives on the purposes, processes, and products involved in writing. As above, they found that younger writers and struggling writers placed a strong emphasis on surface-level features of writing (i.e., handwriting, spelling, and sentence structure). However, older writers and successful writers were able to discuss more global aspects of writing, such as audience awareness and communication of meaning. For instance, when asked to describe the purpose of writing, one elementary-aged, struggling writer responded, “Because they want us to learn...” However, a successful middle school student responded, “When we get our job, we need to know how to write and get our recommendations. To [sic] prepared everything for life, you need to write.” Similarly, when asked about the processes that good writers employ, an elementary student responded, “They put period,” whereas middle school students discussed audience awareness, “They think about who’s reading it...Authors have to write so that others can feel like it’s the author talking and feel connected.” Overall, the results of such research show that as writers develop and expand their knowledge, they move from a shallow and local understanding of writing to a deeper and more global understanding of writing. Nonetheless, across grade levels, a number of students continue to display a misunderstanding of the goals and features of effective writing.

An additional challenge for developing writers resides in the metacognitive processes of self-assessment necessary to judge one's own work. That is, students may display poor writing proficiency because of a failure or inability to accurately assess the quality of their writing. Compared to teachers, who assess students' essays on numerous dimensions and multiple levels of text (e.g., words, organization, meaning, style, and so on), students may apply only a limited set of criteria (e.g., only lexical features) or an incorrect set of criteria (e.g., judging the readability of a text without considering audience needs). More generally, students may struggle with the process of objectively critiquing their own work or with the appreciation of how their outcomes may have fallen short of their intentions.

Research on the metacognitive processes of writing has focused on the benefits and accuracy of students' self-assessment (Andrade & du Boulay, 2003; Andrade, Du, & Wang, 2008; Graham & Perin, 2007; Hillocks, 1986; Ross, Rolheiser, & Hogaboam-Gray, 1999). For instance, Ross and colleagues (1999) investigated the *accuracy* and *benefits* of student self-assessments in writing. They conducted a classroom study that investigated the effects of a training program designed to teach elementary school (4th-6th grades) students how to self-evaluate their narrative writing. Additionally, they explored the direct benefits of this training on students' self-assessment accuracy. As part of the training program, the students helped to define the evaluation criteria for their own writing and were given instructions on how to apply these criteria to their own work. In addition, they were provided with teacher feedback on their self-evaluations. Overall, the treatment condition led to greater gains in the quality of the students' narrative writing than the control condition, particularly for struggling students. Further, the training led to

increased precision of the students' self-assessments; in particular, the treatment group was less likely to overestimate the success of their writing performance.

Similarly, Andrade and colleagues (2008) examined the effects of self-assessment on elementary school students' writing performance in a writing class. In their study, students were first presented with a model essay intended to generate classroom discussion about its strengths and weaknesses. Following the discussion, the students collaborated to generate a list of criteria associated with quality stories and essays. Finally, students received a rubric and were asked to self-assess their own writing based on its criteria. The results revealed that the treatment condition led to higher essay scores over the control condition after controlling for prior English ability. Thus, by providing students with explicit instruction of writing criteria, as well as specific guidance on implementing these criteria, students' performance on writing assignments improved. Overall, these studies reveal the important role of the metacognitive process of self-assessment in writing development. Previous research points to improvements in writing performance as a result of students' self-assessments. Nonetheless, little is known about the characteristics and accuracy of these evaluations. Hence, writing research may benefit from an increased understanding not only of the effects of self-evaluation on writing performance, but also of the features in text that influence students' criterion for quality writing.

Although not explicitly related to writing, prior research has revealed that students are generally inaccurate in their self-assessments of performance. As these findings have spanned numerous domains, it is probably safe to assume that students exhibit these same inaccuracies when assessing their own writing. In a widespread review of the literature,

covering a number of domains in higher education, such as law, medicine, engineering, and psychology, Falchikov and Boud (1989) found that college student and teacher assessments of performance tended to yield only moderate correlations around $r = .39$, and students' predictions of their anticipated grades exceeded teachers' assigned grades about 68% of the time. The accuracy of self-assessments was slightly higher in well-defined domains, such as engineering, and among students taking advanced courses (i.e., higher-performing students). More recently, Tousignant and DesMarchais (2002) evaluated the accuracy of medical students' self-assessments in a problem-based learning program. Self-assessments for three tests were obtained both prior to an examination and immediately following the examination. Results indicated that students were inaccurate in predicting their performance prior to the tests, with only weak correlations between students' predictions and their actual test scores (r ranged from 0.04 to 0.24 for the three tests). Immediately after completing the exam, students' self-assessment accuracy improved slightly, but the correlation with actual scores remained low ($r = 0.26$ to 0.33). In sum, available research suggests that many students apply limited or faulty criteria in the assessment of their performance, and this misalignment likely extends to self-assessments of writing quality.

This misalignment may lead to serious consequences for student writers. First, students may produce texts that fail to meet particular writing goals (e.g., persuasive essays that lack evidence) or that achieve those goals inappropriately (e.g., evidence that is overly subjective and speculative rather than objective and factual). Second, students may find it difficult to understand or apply the feedback received from teachers. For instance, a teacher may urge the student to "pay closer attention to appropriate word

choice,” with the intention that the student should employ more diverse, precise, and descriptive wording. However, the student may interpret feedback about word choice to mean they should “use bigger words to impress the teacher.” In this case, the revised essay would be more likely to contain more multisyllabic words, perhaps used inappropriately, and continue to display problems of word usage. Similarly, misalignment may contribute to students’ difficulties in assessing their own writing. Students may make inaccurate or overly positive judgments of their own work, because they are misapplying the criteria needed to make those judgments. Thus, not only may student and teacher misalignment directly contribute to students’ poor writing, it may hinder the very communication and metacognitive processes needed for students to learn and improve.

We propose that a better understanding of students’ writing proficiency calls for further research on how students’ beliefs and perceptions of “good” writing diverge from teachers’ conceptions. It is possible that teachers’ perceptions of essays are colored by subtle influences that are underlying a more specific rubric. Thus, even if provided with a scoring rubric, students’ perceptions of the rubric components may still diverge markedly from those of their teachers. In this study, we address two principal questions concerning student and teacher evaluative misalignment. First, how are teachers’ ratings of essay quality related to the linguistic features of student essays, such as syntax, cohesion, or emotional word use? Second, how do students’ self-assessments of their own writing diverge from teachers’ ratings? To address these questions, we analyzed student essays using two automated text analysis tools: Coh-Metrix and LIWC. These tools have the power to provide numerous measures of the nuanced text features in students essay.

Thus, the use of these tools may help to highlight some of the less obvious or explicit aspects of students' and teachers' evaluation criteria.

Automated Textual Analyses to Explore Student and Teacher Evaluative

Misalignment

In this study, we conducted two automated analyses of student essays in relation to student and teacher quality ratings. These analyses use Coh-Metrix and LIWC, both of which have been widely used in previous studies on text and discourse.

Coh-Metrix. Coh-Metrix is a computational tool that analyzes text on a variety of textual dimensions relating to cohesion, and text difficulty (Graesser et al., 2004; McNamara & Graesser, 2012). A sampling of key indices provided by Coh-Metrix is discussed below.

Basic Text Measures. Coh-Metrix assesses fundamental properties of text, such as the total number of words, words per sentence, incidence of parts of speech, number of paragraphs, and so on. Many of these measures have been shown to be important predictors of expert ratings of essay quality, and capable of discriminating essays based on the grade levels of the writers (Crossley & McNamara, 2011; Crossley, Weston, McLain-Sullivan, & McNamara, 2011).

Lexical Indices. Coh-Metrix also assesses text using many word-level measures. Many of these indices are calculated through the WordNet computational lexical database (Fellbaum, 1998), which is organized into lexical networks based upon connections between related concepts. Example Coh-Metrix indices provided by WordNet include *polysemy* (the number of senses attributed to a word) and *hypernymy* (the specificity of a word). Other Coh-Metrix lexical indices are obtained using the MRC Psycholinguistic

Database (Wilson, 1988). The MRC database is comprised of over 150,000 words that have been rated along 26 possible linguistic and psycholinguistic dimensions. For example, *age of acquisition* refers to the fact that some words appear in a child's language before others. Additionally, *word familiarity* provides a measure of how familiar printed words seem to a typical person.

Cohesion Indices. Coh-Metrix provides over 50 measures of textual cohesion. For example, one way to establish cohesion is through the use of connective phrases that make conceptual relations explicit (e.g., the phrase *on the other hand* can be used to signal the presentation of counterevidence or opposing viewpoints). Use of connectives can be assessed based on valence (i.e., whether the connective phrases are positive or negative) and based on functional categories (Halliday & Hasan, 1976; Louwrese, 2001). Such categories include clarifying connectives (e.g., *in other words*), additive connectives (e.g., *moreover*), temporal connectives (e.g., *subsequently*), and causal connectives (e.g., *consequently*).

Another Coh-Metrix cohesion index is lexical overlap, which includes *content word overlap*, *argument overlap*, *stem overlap*, and *noun overlap* (see McNamara, Louwrese, McCarthy, & Graesser, 2010, for more detail). Content word overlap measures the proportion of content words shared between two sentences. Argument, stem, and noun overlap are binary measures of the frequency that two sentences share nouns and pronouns, stems, and identical nouns, respectively.

Latent Semantic Analysis (LSA). Coh-Metrix uses LSA to analyze text cohesion at the semantic level (McNamara, Cai, & Louwrese, 2007). LSA uses a statistical method to reduce a large matrix of word co-occurrence into approximately 100-500 dimensions

and is typically used to compute the similarity between sentences or between a sentence and an entire passage (Landauer, McNamara, Dennis, & Kintsch, 2007). Coh-Metrix provides multiple LSA measures, including *sentence-to-paragraph*, *sentence-to-text*, *paragraph-to-paragraph*, and *paragraph-to-text* to measure the semantic co-referentiality of texts. These measures reflect the *semantic* similarities, rather than *surface* similarities, that occur at the sentence, paragraph, and overall text levels.

Validity of Coh-Metrix. Coh-Metrix has been shown to be an informative and reliable text analysis tool in a number of prior studies. One line of studies has focused on examining the linguistic features of high-quality essays (Crossley & McNamara, 2010; Crossley & McNamara, 2011; Crossley & McNamara, 2012; McNamara, et al., 2010). Crossley and McNamara (2011), for instance, used Coh-Metrix to investigate the role of cohesion and coherence in expert evaluations of essay quality. They found that coherence as assessed by expert raters was an important characteristic of high-quality essays, but that coherence was marked by an *absence* of cohesive cues in the text, rather than a presence. Similarly, McNamara et al., (2010) used Coh-Metrix to determine the linguistic differences in essays rated *high* and *low* by experts. The indices most predictive of essay quality were *syntactic complexity* (number of words before the main verb in a sentence), *lexical diversity*, and *word frequency*. Their results indicated that expert judgments of essay quality were sensitive to linguistic features associated with text difficulty and a refined use of language. Indeed, these results have emerged across a number of corpora for both first and second language writers (for a review of the studies, see Crossley & McNamara, 2011).

In addition, Coh-Metrix has been also been used to assess student paraphrases (Rus, Lintean, Graesser, & McNamara, 2009), assess paragraph quality in student essays (Roscoe, Crossley, Weston, & McNamara, 2011), and to detect grade level of student writers (Crossley, Weston, et al., 2011). Thus, Coh-Metrix has been established as a useful tool that is capable of detecting subtle differences within student writing.

Linguistic Inquiry and Word Count (LIWC). LIWC is a text analysis tool that uses categorical word dictionaries to provide information corresponding to thematic and rhetorical language use (Pennebaker et al., 2007). The categorical dictionaries are hierarchical and each consists of a list of words that are related to a particular concept or theme. LIWC contains approximately 4,500 words and word stems across a number of dictionaries; select indices are described in detail below.

Basic Text Measures. Like Coh-Metrix, LIWC provides measures of basic textual information. For a given text, LIWC calculates word count, incidence of punctuation, frequency count of words containing more than six letters, incidence for some parts of speech, incidence of verb tenses, and so on. These text measures provide information about surface-level text features.

Psychological Processes. The psychological process categories of LIWC relate to social, affective, cognitive, perceptual, and biological processes of humans. Each category is further divided into sub-dictionaries that reflect specific characteristics of these processes. For instance, the *perceptual processes* category contains words related to sensory and perceptual concepts, which are divided into three sub-dictionaries: *see*, *hear*, and *feel*. Each sub-dictionary contains numerous words related to that specific category. For instance, the sub-dictionary *see* contains words, such as ‘gaze’ and ‘bright,’ whereas

the *hear* sub-dictionary contains words, such as ‘ring’ and ‘whisper.’ The *cognitive processes* category contains numerous sub-dictionaries, such as *insight* (e.g., ‘aware’ and ‘notice’), *certainty* (e.g., ‘absolute’ and ‘factual’), and *exclusion words* (e.g., ‘if’ and ‘just’). The more frequent occurrence of words within a category is assumed to reveal information about the semantic content of an essay. For instance, a high incidence of *social* words can reflect a text that relates to broader social or cultural concerns, but a high incidence of *cognitive* words signals a more opinionated and analytical text.

Personal Concerns. LIWC also provides measures of personal issues that reflect the theme of a text. Some examples of these sub-dictionaries are *work* (e.g., ‘boss’ and ‘publish’), *leisure* (e.g., ‘bath’ and ‘read’), and *money* (e.g., ‘account’ and ‘portfolio’). These measures provide a better understanding of the specific themes and topics that are being discussed in a given text.

Validity of LIWC. LIWC has been employed in numerous studies to measure the emotional, cognitive, structural, and process components present in a text. Many of the studies have focused on discriminating emotional states and personality features of the speakers within a given text. For example, Hancock, Landrigan, and Silver (2007) investigated the ways that individuals express emotion during text-based communication. They found that individuals relied on four strategies to express levels of happiness: negations, negative affect terms, punctuation, and verbosity. All indices except verbosity were calculated using the LIWC software. Gill, French, Gergle, and Oberlander (2008) similarly analyzed the emotional language use of authors in blog entries. They collected blog posts of 50 and 200 words that had been previously coded by expert and naïve raters. The results showed that the “angry” authors used more affective language and

negative affect words, whereas “happy” authors used more positive affect words. Moreover, they found that the LIWC results were consistent with human ratings.

In addition to text analyses, LIWC has been used in various ways to understand natural speech. Pennebaker, Mayne, and Francis (1997) found that LIWC indices successfully detected improvements in physical and mental health following traumatic events, and Hancock, Curry, Goorha, and Woodworth (2008) used LIWC to analyze the differences between deceptive and truthful conversations in an instant messaging environment. The results of these studies show that LIWC has the potential to detect changes in individuals’ language use, as well as distinguish between groups based on language use. In sum, LIWC has been established as a useful tool that provides information about themes, content, and genre within a text.

METHOD

In this study, we investigate the degree to which linguistic features of text, as measured by Coh-Metrix and LIWC, are predictive of student and teacher ratings of essay quality. Through the use of automated text analysis tools, we aim to identify some of the more nuanced features of students’ essays that potentially affect student and teacher ratings of essay quality. Throughout our analyses, alignment is defined as the extent to which student ratings are predicted by, or related to, the same features as teacher ratings. Evaluative misalignment, on the other hand, is defined as the extent to which student ratings are predicted by *fewer* or *different* linguistic indices than teacher ratings. In this manner, our analysis potentially reveals both the complexity and the actual content of student and teacher misalignments.

Participants

Participants were 126 students enrolled in tenth-grade English courses (approximately 15-16 years of age) at a high school in the Washington, DC area. The school enrolled over 2,400 students, with a student population comprised of 49.0% female students, with 22.3% Asian, 4.2% Black, 9.0% Hispanic, and 59.9% White students. Only 7.0% of the students were described as limited English proficiency, and 10.9% qualified for free or reduced-price meals. As students typically begin to prepare for college entrance exams (including writing assessments) around grade 10, this age group provides a representative sample of students who may be strongly affected by evaluative misalignment.

Essay Corpus

Each participating student wrote a timed (25 minute), prompt-based, argumentative essay. The essays were written in response to an SAT-style prompt selected by the teachers:

A sense of happiness and fulfillment, not personal gain, is the best motivation and reward for one's achievements. Expecting a reward of wealth or recognition for achieving a goal can lead to disappointment and frustration. If we want to be happy in what we do in life, we should not seek achievement for the sake of winning wealth and fame. The personal satisfaction of a job well done is its own reward.

Are people motivated to achieve by personal satisfaction rather than by money or fame? Plan and write an essay in which you develop your point

of view on this issue. Support your position with reasoning and examples taken from your reading, studies, experience, or observations.

Essay Evaluations

Teacher Evaluations. Several weeks after students completed their essays, teachers from several classrooms exchanged essays such that no teacher graded his or her own students' work. Teachers rated student essays using the scoring rubric published by the SAT and College Board (Camara, 2003), which resulted in a single, holistic score on a 1-to-6 scale. A "1" is the lowest rating that an essay can receive and a "6" is the highest. Teacher scores had an average rating of 3.67 ($SD = 1.01$) and were normally distributed.

Student Evaluations. Approximately one week after writing their essays, students were asked to predict the score of their essays. As with teachers, students provided a rating on a scale of 1-to-6, with "1" being the lowest score and "6" being the highest. As this was an in-class activity, there was insufficient time to train students on the complete SAT rubric. Thus, students were given a simplified version of the rubric to assess their own essays. The survey provided both qualitative and quantitative choices for student ratings. For example, the highest rating students could choose stated, "My essay was '*Great*' and will get a 6 out of 6 (highest score)" whereas the lowest rating students could choose stated, "My essay was '*Poor*' and will get a 1 out of 6 (lowest score)."

Students' self-assessments were normally distributed and had an average score of 4.04 with a standard deviation of 0.82. They were only moderately and positively correlated with teacher scores ($r = .26, p < .01$). Relative to teachers, students tended to slightly overestimate their scores; $t(125) = 3.86, p < .001$, which represents a small to

moderate effect size ($d = .40$). Overall, the pattern of means, and the low correlation between student and teacher scores, suggest a potential misalignment between the students' and teachers' expectations for the essay quality. We used the automated text analysis tools Coh-Metrix and LIWC to further explore the *characteristics* of this misalignment.

Statistical Analyses

To examine the misalignment between student and teacher ratings of essay quality, we conducted correlation and regression analyses using essay ratings and textual features. Specifically, we examined how and whether specific linguistic text features were correlated with, or predictive of, student and teacher ratings. Analyses were conducted for each set of scores (i.e., teacher and student ratings) and each computational tool (i.e., Coh-Metrix and LIWC) separately. First, correlations were calculated between text indices provided by the automated tool and the essay scores. The pattern of correlations was examined for indices related to scores at the $p < .05$ level, and the variables with the strongest relations to the scores were included in the regression model. To address multicollinearity, when variables correlated with each other above $r = .70$, the variable with the lowest relation to the student and teacher scores was removed. To avoid overfitting the model, we chose a ratio of 15 essays to 1 predictor, which allowed 8 indices to be entered, given that there were 126 essays included in the analyses.

RESULTS

Coh-Metrix Analyses

Teacher Ratings. Correlations were calculated between the Coh-Metrix indices and the teacher scores. As shown in Table 1, 12 variables were significantly correlated with the teachers' scores.

Table 1.

Correlations between Teacher Ratings and Coh-Metrix Variables

Coh-Metrix Variable	Correlation with Teacher Ratings
Total number of words	.337**
Familiarity of content words	-.294**
Total number of paragraphs	.240**
Incidence of locational entities	.239**
LSA paragraph to paragraph	-.226*
Age of acquisition of content words	.224*
Lexical diversity (VOCD)	.218*
Polysemy of words	-.207*
Noun incidence	.205*
Content word overlap	-.194*
Total number of sentences	.192*
Frequency of content words	-.192*

* $p < .05$; ** $p < .01$

Corroborating past research with expert raters (Crossley & McNamara, 2011; McNamara et al., 2010), these correlations indicate that teachers' ratings were influenced by aspects of essay elaboration, essay organization, skillful use of language, and a lack of cohesion. For example, the *total number of words*, *total number of sentences*, and *total*

number of paragraphs indices are indicators of the overall length and structure of essays. In particular, the *total number of words* and *total number of sentences* provide measures of the length of the essay, and longer essays can be indicative of the elaboration of ideas and examples. Similarly, the total *number of paragraphs* broadly measures the organization of an essay. Essays with more paragraphs may possess clearer demarcations between separate ideas, especially when compared to the commonplace “one-paragraph essays” written by novice writers. Overall, teachers rated longer and organized essays more highly, which is to be expected. This finding indicates that teacher ratings were related to the elaboration and organization of ideas in students’ essays.

The teachers’ essay ratings were also correlated with lexical features of the essays. The *familiarity and frequency of content words*, *age of acquisition* of content words, and *polysemy of words* indicate more sophisticated language use. *Word familiarity*, *word frequency*, and *age of acquisition of content words* indicate that a given essay is composed of more uncommon and sophisticated words. *Word polysemy* is indicative of the degree of ambiguous language utilized in an essay. Thus, teachers’ ratings were related to students’ use of specific words, which indicate more precise descriptions of ideas and concepts. Albeit somewhat weak, the correlations with these variables suggest, not surprisingly, that teachers are sensitive to vocabulary use, particularly students’ use of less familiar and abstract words. Teachers’ scores were also influenced in part by more concrete language, as measured by their relationship to *incidence of locational entities* and *noun incidence*. This suggests that teachers prefer persuasive essays that contain more nouns, particularly those that refer to a specific

location (e.g., house, store, Georgia). These measures are indicative of more concrete language, as they refer to specific places and objects.

Finally, variables such as *lexical diversity*, *LSA paragraph-to-paragraph*, and *content word overlap* reveal a benefit of using diverse language and developing low overt text cohesion. The correlation analyses indicate that teachers associate low-cohesion essays with higher overall essay quality (as found in previous research on expert ratings: Crossley & McNamara, 2010; Crossley & McNamara, 2011; Crossley & McNamara, 2012; McNamara, et al., 2010). It is possible that high-cohesion essays are too reliant on repetitive vocabulary and examples to connect ideas. On the other hand, low-cohesion essays may rely on deeper (i.e., not surface-level) arguments structures to develop text coherence. This finding is line with prior research on text comprehension, which suggests that cohesive devices can support or hinder the development of coherent text representations depending on readers' level of prior knowledge (McNamara & Magliano, 2009; O'Reilly & McNamara, 2007).

Overall, these results confirm, and allow us to document, common intuitions about teachers' criteria. The teachers' ratings are related to numerous aspects of student essays, ranging from lower- to higher-level features. Namely, teachers seem most sensitive to the elaboration and organization of ideas, sophisticated vocabulary and language use, and a reduced cohesion.

A regression analysis was conducted to assess how and whether the above variables predicted teachers' essay ratings. All variables were tested for multicollinearity ($r > .70$) and two variables (*frequency of content words* and *total number of sentences*) were eliminated due to a strong relationship to other variables. The analysis yielded a

significant model, $F(8, 125) = 6.89, p < .001; R^2 = .32$. The significant predictors in the model were *total number of words* ($B = .27, p < .01$) and *LSA paragraph-to-paragraph* ($B = -.29, p < .001$). Two additional variables in the model were statistically significant, if tested one-sided: *number of paragraphs* ($B = .15, p = .097$) and *word polysemy* ($B = -.16, p = .07$). These results suggest that the linguistic features most predictive of the teacher ratings in this sample were related to essay elaboration (i.e., length of the essays), followed by less abstract wording and reduced cohesion. In general, the Coh-Metrix analysis reveals that teacher quality ratings are associated with numerous essay components, including lexical, syntactic, and cohesive features.

Student Ratings. As shown in Table 2, seven variables were significantly correlated with the students' scores. Importantly, these correlations reveal that students' ratings were partially associated with *different* features of the essay than were teachers' ratings, indicating some degree of misalignment in the criteria. The correlations indicate that students were most sensitive to the level of personalization in their essays along with strong vocabulary and language use. In particular, the *incidence of 2nd person pronouns* is indicative of the level of personal or familiar language in students' essays. Essays that contain a high incidence of second person pronouns often rely too heavily on personal stories and anecdotes as examples and evidence statements. Accordingly, students seemed to be aware that this overly familiar language potentially reduces the quality of their essays. In addition, students rated their essays more positively when they had a higher *mean number of syllables per word* and *mean age of acquisition of words*. These variables represent the length of the words used (i.e., the number of syllables), as well as the sophistication of the words (i.e., the age at which the vocabulary words are typically

acquired). As academic and professional writing typically contain more sophisticated vocabulary, students may have perceived their essays to be of higher quality if they incorporated longer and less common words.

Table 2.

Correlations between Student Ratings and Coh-Metrix Variables

Coh-Metrix Variable	Correlation with Student Ratings
Second person pronoun incidence score	-.249**
Noun incidence	.226*
Age of acquisition of content words	.226*
Lexical Density	.221*
LSA verb overlap	.180*
Incidence of locational entities	.180*
Average syllables per word	.178*

* $p < .05$; ** $p < .01$

Student essay ratings were also positively related to the incidence of concrete language. In particular, *noun incidence*, *lexical density* (proportion of function words in the text), and *incidence of locational entities* represent more concrete language, which provides more examples and facts. Essays with a higher incidence of nouns and a higher proportion of lexical items typically contain more concrete and grounded language, as they are less reliant on function words and verbs. Further, the *incidence of location entities* provides a count of the nouns that refer to a specific location (e.g., Arizona or house). The positive relation between students' ratings and these location nouns indicates a preference for specific (i.e., not abstract) facts and examples. Overall, when assigning

quality ratings, students seemed to be attuned to the level of specific, concrete language use.

Finally, students seemed to rate their essays more highly if they were more cohesive, as suggested by the *LSA verb overlap* measure. Specifically, this measure indicates that students assigned higher quality ratings when their essays were more semantically connected. This result is contrary to the teachers' ratings, and suggests that students are unaware of the level of cohesion appropriate for high essay quality. Because students are often taught to develop clearly connected ideas, their evaluation criteria may require explicit cohesive devices in the essays. Thus, they may provide higher overall ratings to their essays when they contain these overt cohesion features. Overall, students' essay ratings were most highly associated with lexical features and concrete language use. This is somewhat in contrast to teacher ratings, which were associated with a wider variety of indices, including organization and elaboration of ideas.

A regression analysis was conducted to assess how and whether the correlated variables predicted students' essay ratings. All variables were assessed for multicollinearity ($r > .70$) and one variable (*average syllables per word*) was eliminated due to a strong relationship with other variables. Students' predicted scores were regressed onto the six remaining variables in a linear regression, yielding a significant model, $F(6, 125) = 3.47, p < .01; R^2 = .15$. No single variable was a significant predictor in the model, although three variables approached significance: *age of acquisition of content words* ($B = .18, p = .05$), *LSA verb overlap* ($B = .15, p = .095$), and *incidence of locational entities* ($B = .16, p = .08$). The results of the regression suggest that students' ratings of essay quality are less systematically related to the linguistic features of essays

than teachers. Thus, they are utilizing an incomplete or different set of criteria when providing self-assessments of essays. While some linguistic variables are associated with students' overall ratings, they are not strongly related to or predictive of student ratings. Thus, students may focus on other aspects of their essays when assigning ratings, such as the theme or content, or even how they felt emotionally while they wrote it.

Summary of Coh-Metrix Analysis. An analysis of textual features related to students' and teachers' ratings of essays revealed that there was, indeed, misalignment in the evaluation criteria. Overall, teacher ratings were more strongly related to the Coh-Metrix variables with an $R^2 = .32$, compared to the student ratings, which reported an $R^2 = .15$. In addition, teachers' ratings were significantly correlated with a larger number of indices than were students' ratings. This is unsurprising, given that teachers necessarily have a broader understanding of how multiple text features interact to produce quality essays. For example, students seemed attuned to word length (e.g., *number of syllables*) as an indicator of lexical sophistication, whereas teachers attended to whether the words were less common and more precise. Indeed teachers have a more thorough understanding of the different features related to essay quality at both superficial and deep levels. One explanation for this low relationship between student ratings and linguistic variables is that students are paying attention to different levels of the content of their essays (e.g., themes or genres). Thus, our subsequent analyses also evaluated teacher and student misalignment using LIWC, which places a stronger emphasis on the textual features that related to thematic or genre content. Because LIWC is a similar, yet more thematic and idea-based tool, this second analysis serves as a triangulation, providing converging evidence for evaluative misalignment.

Table 3.

Correlations between Teacher Ratings and LIWC Variables

LIWC Variable	Correlation with Teacher Ratings
Word count	.333**
Cognitive mechanisms	-.282**
Tentative words	-.278**
Future tense words	-.275**
Present tense words	-.272**
Verbs	-.261**
Certainty words	.243**
Third person plural pronouns	.239**
Human words	-.233**
Exclusion words	-.223*
Insight words	-.217*
Words containing more than 6 letters	.213*
Feeling words	-.209*
Auxiliary verbs	-.203*
Past tense words	.202*
Perception Words	-.181*

* $p < .05$; ** $p < .01$

LIWC Analyses

Teacher Ratings. As shown in Table 3, 17 LIWC indices were significantly correlated with the teachers' ratings. The results in Table 3 indicate that teachers' ratings were most strongly related to essay elaboration, vocabulary strength, and the skilled use of language in student essays. Not surprisingly, teachers seemed to be most attuned to the length or

elaboration of the student essays, as evidenced by the correlation with *word count*. Essays composed of more words often contain more detailed elaborations of arguments and examples. Similarly, the teachers rated essays more highly when they included longer vocabulary words (*words containing more than 6 letters*). Because strong vocabulary is typically associated with longer words, this correlation indicates that teachers score higher essays with more sophisticated word choices.

Other correlated variables suggested that teachers were sensitive to the use of objective and fact-based language in student essays. For example, hypothetical language (*exclusion words*), hedging language (*tentative words*), emotional language (*feeling words*), and other subjective words (*insight words*, *perception words*, and *cognitive mechanisms*) had a negative association with teacher ratings. The *exclusion words* and *tentative words* measure uncertain language, as they represent ungrounded and hesitant word choices. For instance, the *tentative words* category includes words such as “might,” “possibly,” and “could,” which establish weaker arguments and examples. On the other hand, objective and confident language (*certainty words* and *third person plural pronouns*) was associated with higher ratings by teachers. Essays with objective and confident word choices may develop stronger and more sophisticated arguments. Overall, as one would expect, teachers are attentive to the strength and objectivity of the language that students used when developing their arguments.

A regression analysis was conducted to examine the extent to which LIWC variables predicted teacher ratings. The indices were checked for multicollinearity ($r > .70$) and two variables (*perception words* and *verbs*) were eliminated due to a high relationship to other variables. The regression yielded a significant model, $F(8, 125) =$

5.81, $p < .01$; $R^2 = .28$, with two significant predictors: *word count* ($B = .26$, $p < .01$) and *cognitive mechanisms* ($B = -.21$, $p < .05$); and one variable was statistically significant, if tested one-sided: *future tense words* ($B = -.16$, $p = .08$). The results of this analysis suggest that teachers were most concerned with essay elaboration and a more objective use of language; longer essays were most likely to receive a high rating by teachers. Additionally, objective language, as indicated by a lack of subjective words (e.g., *think*, *should*, and *maybe*) was a factor in teachers' assignment of high ratings to student essays.

Student Ratings.

Table 4.

Correlations between Student Ratings and LIWC Variables

LIWC Variable	Correlation with Student Ratings
Third person plural pronouns	-.270**
Tentative words	-.261**
Second person pronouns	-.251**
Third person singular pronouns	.226*
Certainty words	.200*
Sadness words	-.196*
Function words	-.193*
Words containing more than 6 letters	.187*
Present tense verbs	-.185*

* $p < .05$; ** $p < .01$

Nine variables were significantly correlated with student scores (see Table 4). The results indicate that student ratings were more highly correlated with linguistic variables

related to three major factors: objectivity of language, level of confidence expressed in the essays, and vocabulary strength.

For example, students' essay ratings were also related to the objectivity of their language. Objective language was measured by a positive correlation with *third person singular pronouns*, as well as a negative correlation with *second person pronouns*. Second person pronouns indicate a higher incidence of personal and familiar language, in contrast to third person pronouns, which are representative of more objective language. Thus, students were somewhat aligned with teachers in the focus on strong vocabulary and objective language use when rating essays.

Similarly, students' ratings were dependent on the level of confidence expressed in their essays. When arguments were developed with confidence words (*certainty words*) and with a low incidence of hedges (*tentative words*), students were more likely to rate their essays higher. Essays with more confident language typically develop stronger arguments and examples. Thus, students' perceptions of their writing quality are somewhat aligned with teachers' ratings regarding the use of confident language in the essays.

Finally, vocabulary strength (*words containing more than 6 letters*) was related to high student ratings. Thus, students may feel more confident in quality of their writing if they utilize more complex vocabulary. Overall, the LIWC analysis revealed that student ratings were highly associated with the strength of vocabulary, language use, and use of confident language in the essays.

A regression analysis was conducted to assess which LIWC variables, if any, predicted student essay ratings. The variables were assessed for multicollinearity, but no

two variables were correlated above the .70 threshold. The regression yielded a significant model, $F(8, 125) = 4.02, p < .001; R^2 = .22$, with two significant predictors: *third person plural pronouns* ($B = -.21, p < .01$) and *second person pronouns* ($B = -.20, p < .05$), and one predictor that approached significance: *certainty words* ($B = .15, p = .08$). The positive relation to *third person plural pronouns* and the negative relation to *second person pronouns* suggest that students were sensitive to the level of personalization in the essays when providing quality ratings. Additionally, the predictor, *certainty words*, implies that the level of confidence expressed in the essays influenced student self-assessments of writing quality.

Summary of LIWC Analysis. The linguistic features captured by the LIWC measures provide further information about the characteristics of the misalignment between student and teacher evaluation criteria. Overall, the LIWC indices were able to capture approximately one-fourth of the variance in both teacher ($R^2 = .28$) and student ($R^2 = .22$) ratings of the essays. In addition, LIWC analyses revealed a partial alignment between student and teacher evaluation criteria. Similar to the Coh-Metrix analysis, the LIWC analysis suggested that students' and teachers' ratings relied somewhat on sophisticated vocabulary and objective and confident language use. When students expressed their ideas confidently (*certainty words*) and avoided personalized language (*third person plural pronouns; second person pronouns*), both students and teachers assigned higher quality ratings. The LIWC indices, however, also demonstrated areas of misalignment between the students and teachers. Specifically, the teachers were sensitive to deeper issues and strong language use than were the students. For instance, in addition to the relationship to objective and confident language use, teacher ratings were also associated

with a lack of hypothetical, emotional, and perceptual language. This suggests that teachers were better able to assess texts based on a larger number of textual features than the students. The results of the LIWC analyses suggest that students and teachers were, at least partially, misaligned in their criteria for quality essays. As one would expect, teachers have a more expansive conceptualization of the different features that interact to produce quality essays.

DISCUSSION

The results of this study indicate that teachers do indeed assess student essays on a variety of linguistic measures at surface- and deep-levels of text. In contrast, students' ratings are associated with a smaller subset of variables, namely surface-level features. Thus, our results are in line with the hypothesis that there is an evaluative misalignment between the criteria of students and teachers. Despite the importance of students understanding the feedback that they receive on their writing, the nature of potential misalignments between students' and teachers' writing evaluation criteria has not been examined in the composition literature. In this study, we explored one assumption of the misalignment hypothesis using a textual analysis of students' and teachers' assessments of SAT-style essays. Specifically, we took the novel approach of investigating how misalignments manifest themselves in terms of linguistic textual properties.

One substantial contribution of this study is the analysis of the linguistic features that most accurately predict teacher ratings of essay quality. Although researchers have investigated the textual features related to *expert* ratings, no study to our knowledge has explored the features that characterize *teacher* ratings. A second contribution is the comparison of teacher ratings to students' self-assessments of their own writing.

Specifically, our analyses reveal the areas in which students' and teachers' evaluation criteria are disparate. With this analysis, we were able to establish linguistic features that characterize students' and teachers' evaluations of essays and confirm the presence of student-teacher misalignment in essay evaluation.

Teacher Ratings

The results of this study provide an extensive analysis of the textual features that are most predictive of teacher ratings of essay quality. Although the results of the teacher analyses are not surprising, they have now been assessed empirically. Through the use of two related, yet different, automated text analysis tools, Coh-Metrix and LIWC, we measured a number of surface- and deeper-level linguistic features of student essays with which teachers' ratings were highly associated. Coh-Metrix provided a broad analysis of the student essays, including indices of cohesion, and text difficulty at the lexical, syntactic, structural, and global levels of text (Graesser & McNamara, 2011). In addition, LIWC offered word-based analyses of the lexical, semantic, and thematic properties of student essays. With these linguistic measures of student essays, we are able to account for a significant amount of variance in teachers' essay assessments.

Although prior work has utilized text analysis tools to investigate expert raters' scores of essay quality, it is unclear whether and how these ratings correspond to classroom teachers' assessments. As classroom teachers and expert raters differ in their goals, training, and context of their scoring, their evaluation criteria also differ. Our analyses revealed that teacher ratings were, indeed, similar to expert ratings of essay quality. Consistent with prior research on expert ratings of essay quality, the teachers' ratings were related to skillful language, text organization, and text elaboration. In

addition, a notable similarity between experts and teachers was the negative influence of cohesive devices. Similar to experts, the teachers in the current study did not associate cohesive essays with higher quality ratings (Crossley & McNamara, 2010; Crossley & McNamara, 2011; Crossley & McNamara, 2012; McNamara, et al., 2010). In fact, the teacher ratings were *negatively* correlated with indices of cohesion. Overall, the analyses suggest that classroom teachers, like expert raters, are able to assess student essays on myriad different features, ranging from surface-level lexical features to deep-level properties of text cohesion.

Student-Teacher Evaluative Misalignment

Beyond our analysis of teacher ratings, we further investigated the degree to which students' and teachers' ratings of essay quality were *misaligned*. Although prior research has investigated the *benefits* and, less commonly, the *accuracy* of students' self-assessments (Andrade & du Boulay, 2003; Andrade, et al., 2008; Graham & Perin, 2007; Hillocks, 1986; Ross et al., 1999), no work has explored the linguistic features that predict students' self-assessments. In our study, we investigated the linguistic features associated with students' self-assessments in order to determine the degree of misalignment between students' and teachers' essay ratings.

The correlation and regression analyses confirmed that there was, indeed, misalignment of student and teacher expectations for writing quality. In line with prior research on students' performance self-assessments (Falchikov & Boud, 1989; Dunning et al., 2003; Kruger & Dunning, 1999), the students participating in this study overestimated the quality of their writing. Teachers gave essays an average score of 3.67, whereas students had an average self-assessment of 4.04. In addition, the scores were

weakly correlated ($r = .26$) indicating that while the score means are not vastly different, the essays rated as low quality by the teachers may have been given higher ratings by students, and vice versa.

A potential limitation of this study lies in the differences between the types of rubrics and training used by the teachers and students. Indeed, this factor deserves further attention in future studies, as student-teacher misalignment may be remediated through more specific student rubrics or more extensive evaluation training. Nonetheless, the differences in the linguistic features of the essays that are associated with the scores are less likely to be driven solely by the rubrics. First, our analyses of the linguistic features related to student and teacher essay ratings indicated that student ratings were related to fewer measured variables than the teacher ratings. While teachers' assessments were influenced by multiple features of the essays, such as sophistication of vocabulary, text organization, and objective and factual language use, students seemed to focus only on a subset of these features. For instance, the LIWC analysis revealed that students and teachers were both sensitive to objective language use, as revealed by the association between essay ratings and pronoun usage. However, teachers' ratings were associated with numerous additional aspects of language use, such as a lack of hypotheticals, hedging words, and emotionally charged language. These findings suggest that, while students can understand the individual aspects of writing quality highlighted in classroom instruction, they likely find it challenging to understand how the effects of some features depend on others. As a result, they focus on fewer or more simplistic aspects of writing quality when assessing essay quality.

Second, in addition to focusing on fewer essay features, students rated essays based on a *different* collection of features than the teachers. For instance, student ratings were positively affected by high semantic overlap (*LSA verb overlap*) within the essays. Thus, the more essays exhibited semantic cohesion among sentences, the more highly students rated their essays. As previously discussed, however, teachers' ratings were *negatively* associated with measures of essay cohesion (*LSA paragraph-to-paragraph* and *content word overlap*). Our analysis, therefore, reveals distinct misalignment between the evaluation criteria of the students and teachers. Specifically, student ratings were based on both fewer and different textual features than teachers' ratings.

We assume that students have yet to develop a complete criterion for evaluating their own essays, and, as a result, are missing the more nuanced and dynamic features that contribute to quality essays. However, students' inaccurate self-assessments are the consequence of numerous factors, including students' knowledge of writing, students' metacognitive monitoring, classroom environment, and pedagogical methods. Of course, this study only addressed part of the evaluative misalignment hypothesis. Here, our goal was to establish the presence of a misalignment and to explore its nature in terms of the linguistic features of the essays that influence the students' and teachers' assessments of essay quality. As such, it is important to note that this type of analysis is complementary, not alternative, to analyses of teachers' explicit criteria for writing. In the future, researchers should investigate the misalignments revealed through students' and teachers' explicit reports of writing quality.

In addition, future studies should explore the sources and causes of student-teacher misalignments. One potential source of the student-teacher misalignment could

be the teachers' ability to compare student essays with those from their classmates. That is, because teachers are able to view multiple essays along numerous levels of quality, they are better able to make sophisticated quality judgments than students. Future studies should investigate this question by presenting students with their own essays in conjunction with several other essays varying in quality. Analyses could then investigate whether students' self-ratings are more aligned with teachers' ratings, and whether this peer-review process results in improved performance on subsequent writing tasks. Along these lines, prior research on peer review and writing has revealed that students benefit from peer review exercises (Nelson & Schunn, 2000; Cho & Schunn, 2007).

Future research should also investigate this evaluative misalignment from a developmental perspective. Because students' aptitude for writing is a result of multiple developing factors, such as interest, motivation, and conceptual competence (Lipstein & Renninger, 2006), it is important to investigate how these variables interact with students' development of accurate evaluation criteria. In the future, the method outlined in this paper should be used to provide important insight into students' development of criteria. In particular, these textual analyses can be applied to longitudinal data of student and teacher essay assessments to determine how misalignment changes over time.

Regardless of their source or cause, misalignments in expectations pose a threat to students' successful writing development. Our results here indicate that students lack stringent evaluation criteria for their essays that consider the myriad of linguistic, rhetorical, and semantic characteristics associated with essay quality. Despite teachers' best approaches and practices in the classroom, students can still misinterpret or misunderstand the writing process, as well as the associated lessons and strategies.

Students without systematic criteria to evaluate their writing will have difficulties with certain phases of the writing process (e.g., revision) as well as improving performance on future writing assignments.

Chapter 3
“Reading Comprehension Components and their Relation to Writing”

Accepted at Topics in Cognitive Psychology

In today’s society, the ability to read and write is crucial for communication with the surrounding environment. With the increasing volume of electronic messaging, online news stories, and other printed sources of information, strong literacy skills are more important than ever. Unfortunately, despite this apparent need, assessments consistently report ubiquitous student underachievement in the literacy domain. This problem has been particularly poignant in the United States, where according to the National Assessment of Educational Progress (2011) only 27% of eighth graders and twelfth graders score at or above proficiency levels in writing, and only 3% of these students score at advanced levels (NAEP, 2011). Comparable findings have been reported for reading, with only 32% of eighth grade and 38% of twelfth grade students scoring at or above reading proficiency (NAEP, 2009). Similar problems persist in many countries across the world with international assessments of literacy suggesting that students, on average, only have intermediate levels of literacy proficiency (e.g., Baer & McGrath, 2007).

One way that educators and researchers have attempted to alleviate these proficiency gaps is by identifying the skills and background knowledge required for the successful comprehension and production of text. Indeed, a good deal of research has focused on the shared *knowledge* required to complete both reading and writing tasks (e.g., knowledge about vocabulary or prior world knowledge; for reviews, see Galda, 1983; Shanahan, 1988; Stotsky, 1983). Along these lines, some researchers hypothesize that the identification of these common factors and the integration of reading and writing

instruction is essential to informing classroom practices designed to enhance literacy skills (Corden, 2007; Couzijn, 1999; Parodi, 2007).

Extensive research has identified a number of cognitive processes involved in reading comprehension (Daneman & Carpenter, 1980; Gernsbacher, 1997; Graesser, Singer, & Trabasso, 1994; Kintsch, 1998; McNamara, Jacobina, & Allen, in press; Myers & O'Brien, 1998; Zwaan, Langston, & Graesser, 1995). These processes range from lower level skills, such as working memory, to higher level cognitive skills, such as generating inferences based on background knowledge. Although these constructs have been thoroughly examined within the domain of reading comprehension (Daneman & Carpenter, 1980; Gernsbacher, 1997; Graesser et al., 1994; Kintsch, 1998; Myers & O'Brien, 1998; Zwaan et al., 1995), relatively little is known about how these higher level skills overlap with writing proficiency. The current work draws from previous research on reading comprehension to examine how lower- and higher level cognitive skills (in addition to vocabulary knowledge) relate to comprehension and, in turn, how these skills map onto students' writing ability. Specifically, we aim to confirm relations between lower and higher level cognitive skills and reading comprehension ability, and further to examine how these same processes potentially relate to writing proficiency. The overall motivation for this study is driven by the hypothesis that reading comprehension and writing share components (e.g., background knowledge, cognitive skills, etc.) that can be leveraged to provide more effective interventions for struggling students.

Theoretical Models of Reading Comprehension and Writing

A number of discourse models have been proposed to account for the cognitive processes that allow a reader to collect information from a text and develop a comprehensive understanding of that text at various levels (Gernsbacher, 1997; Graesser et al., 1994; Kintsch, 1998; Myers & O'Brien, 1998; Zwaan et al., 1995). While these models diverge in more specific components, the majority of contemporary comprehension models highlight the *constructive* and *active* nature of the reading comprehension process. For instance, Kintsch, in his construction-integration model, theorizes that reading comprehension involves *recalling* information from the surface and deeper levels of the text and generating a *situation model*, where background knowledge and experiences are *integrated* into the text for a specific purpose (Kintsch, 1998). Thus, according to these models, reading comprehension relies not only on the background knowledge of the reader, but also on the cognitive processes necessary to capitalize on this existing knowledge.

A number of models have also been developed to account for the cognitive processes involved in writing. One of the most influential conceptualizations of the writing process came from Hayes and Flower (1980), who outlined the individual levels of information processing that are involved in the production of text. Importantly, this model is non-linear and emphasizes the *interactions* that can occur among the writing processes (e.g., planning and translating), the task environment (e.g., assignment and audience), and the information in the writer's long-term memory (e.g., knowledge of the topic). This model was later revised to account for more recent research in cognitive science and writing (Hayes, 1996). Namely, a working memory component was added to

the model and there was a stronger emphasis on the writer's motivation to communicate during the writing process. Overall, both models had a strong impact on the field of writing research because they introduced writing as a non-linear information-processing task that is reliant on a limited capacity cognitive system.

The majority of contemporary writing models emphasize the role of lower level cognitive skills (e.g., working memory), as well as background knowledge; yet, they rarely discuss the role of the higher level skills that are intended to make use of this background knowledge (e.g., integrating knowledge or generating inferences). It is not a stretch, however, to modify the constructive and active aspects of text *comprehension* theories to apply to processes of text *production*. For instance, throughout the writing process, a writer typically *constructs* meaning through a combination of their prior knowledge (e.g., about the world, domain, language) and experience, and works to *integrate* this knowledge into a coherent structure that is appropriate for a particular audience. The similar constructs shared between text generation and production provide sufficient grounds to investigate the higher level cognitive skills that are both shared and unique to the tasks of reading comprehension and writing. Specifically, research on literacy and the reading-writing connection can be substantially enhanced with investigations that are not only based on the knowledge required to complete literacy tasks, but also on the higher level cognitive skills that are associated with accessing and efficiently using this knowledge.

Connections between Reading Comprehension and Writing Proficiency

Previous research on the reading-writing connection has focused primarily on informing pedagogy with the underlying goal of enhancing students' performance on literacy tasks.

One purpose of earlier studies in this vein was to identify and assess relationships between reading and writing. Accordingly, a number of studies have been conducted to investigate reading-writing relations from various perspectives (e.g., background knowledge, procedures, strategies, etc.). Researchers have subsequently attempted to use these findings to integrate various aspects of reading and writing instruction (Clarke, 1988; Graham & Hebert, 2011; Straw & Schreiner, 1982).

Although the research on reading-writing connections remains relatively sparse regarding the role of lower- and higher level cognitive skills, a limited number of studies have identified cognitive skills that are necessary for reading comprehension and writing tasks (both separately and in combination). Lower level skills, such as working memory capacity, have been linked to performance on both reading comprehension and writing tasks (Berninger, Cartwright, Yates, Swanson, & Abbott, 1994; Daneman & Carpenter, 1980), and a number of higher level cognitive skills have been examined in relation to reading comprehension performance (e.g., Bloom, Fletcher, van den Broek, Reitz, & Shapiro, 1990; Chabot, Zehr, Prinzo, & Petros, 1984; Hannon & Daneman, 2001; Masson & Miller, 1983; Singer & Richot, 1996). Prior research, however, has rarely (if ever) simultaneously examined the impact of lower and higher level cognitive skills on reading comprehension and writing performance. Such investigations are critical for the advancement of theoretical understandings of literacy as a whole, as well as for the development of literacy curriculum that is sensitive to similarities and differences between reading comprehension and writing processes.

Shared Knowledge. To understand how to most effectively develop an integrated pedagogy for reading comprehension and writing, educators must understand what

background knowledge is shared between each task and what knowledge is unique. In their review of the literature, Fitzgerald and Shanahan (2000) classified four knowledge bases as common factors across both reading comprehension and writing: meta-knowledge, domain knowledge, knowledge about universal text attributes, and procedural knowledge. *Meta-knowledge* involves knowing about the purpose of reading comprehension and writing, understanding the interaction between readers and writers, and monitoring one's own comprehension and knowledge. This form of knowledge has been positively linked to both reading comprehension and writing in previous studies (Langer, 1986; Shell, Colvin, & Bruning, 1995). *Domain knowledge* refers to one's prior knowledge about a given content area, as well as the knowledge that may be gained during the reading comprehension or writing processes (e.g., word meanings learned through context). *Knowledge about universal text attributes* includes a reader or writer's understanding of syntax and discourse (e.g., schemata knowledge), along with their knowledge of specific language features, such as phonemes and morphology. Finally, *procedural knowledge* that is shared across reading comprehension and writing typically involves the skills needed to construct meaning from, or with, a text (Langer, 1986). This can include primarily automatic processes, such as the recall of information from memory, or deliberate strategies, such as developing analogies or questioning given information (Kellogg, 1994).

A number of studies have investigated shared knowledge between reading comprehension and writing using assessments related to both the surface and deep levels of text (e.g., vocabulary knowledge, text cohesion, etc.; for reviews, see Galda, 1983; Shanahan, 1988; Stotsky, 1983). These studies have typically involved measuring one

aspect of reading (e.g., holistic comprehension score, vocabulary knowledge), along with one for writing (e.g., holistic writing score, organization, grammar), and reporting the correlation between the variables. The reported correlation serves as evidence for a connection between reading and writing. In a review of the literature, Tierney and Shanahan (1991) reported consistently moderate to strong correlations between various measures of reading and writing (though typically never exceeding $r = .50$).

One of the earliest and most extensive studies of the reading-writing connection was a longitudinal study that followed over 300 students from kindergarten until their high school graduation (Loban, 1963; Loban, 1967). Data were collected for each student on reading, writing, and listening abilities, along with various other aspects of their language use. Loban (1967) reported a strong relationship between reading and writing, finding that poor readers were also poor writers, while good writers were also good readers. Specifically, he reported that all students who ranked superior in writing ability ranked above age level in reading; additionally, all students who ranked illiterate in writing ranked below age level in reading.

Following the successful results of Loban's study, numerous researchers continued to investigate the shared background knowledge found across reading comprehension and writing (e.g., Abbott & Berninger, 1993; Berninger, Abbott, Abbot, Graham, & Richards, 2002; Juel, Griffith, & Gough, 1986; Shanahan, 1984). Abbott and Berninger (1993), for instance, investigated the relationships between potential latent factors underlying the writing development of over 600 elementary school students. They found correlations ranging from $r = .22$ to $r = .54$ between writing quality and reading comprehension ability. Similarly, Juel et al. (1986) conducted a longitudinal study that

followed students from first to second grade and analyzed the reading-writing relationship at multiple levels of text, including spelling, word recognition, reading comprehension, and writing ability. They concluded that the lowest-level aspects of reading and writing (e.g., word recognition and spelling) relied upon similar knowledge bases, such as phonemic awareness and vocabulary knowledge, but the higher level abilities (e.g., comprehension and production) relied upon different, yet somewhat overlapping, knowledge bases (e.g., discourse knowledge, strategic knowledge).

Although correlations between knowledge sources related to reading and writing have been fairly consistent across age and text levels (Fitzgerald & Shanahan, 2000), even higher correlations have been found by researchers using multiple measures to capture individual text aspects (e.g., vocabulary knowledge, text comprehension, and coherent text production). For example, Berninger et al. (2002) investigated connections between reading and writing as part of a larger research project comparing language by hand (reading and writing) to language by eye (listening and speaking). They administered multiple measures of each factor, yielding shared variances of up to 85% for word recognition and spelling factors, and up to 66% for surface-level text comprehension and production. Their results indicated that reading had a greater influence on writing than writing had on reading, but that both skills were strongly related across multiple grade levels.

Shared procedures and strategies. Prior research on the reading-writing connection has typically neglected to examine the specific procedures and strategies that are involved in both literacy tasks; however, this issue has begun to receive more attention. Specifically, researchers have investigated overlapping steps, procedures, and

strategies employed by students during both reading comprehension and writing tasks, such as summarizing information, developing hypotheses, and making connections about the text (Birnbaum, 1982; Martin, 1987; Shanahan, 1984; Tierney, 1983). The majority of the data from these studies involve think-aloud protocols, interviews, or classroom observations. Therefore, the results are predominantly qualitative in nature.

Ryan (1985) analyzed the verbal protocols of eight students who were above average proficiency in both reading and writing. She identified six strategies that were common to reading comprehension and writing for these students: reporting (paragraphing), conjecturing (hypothesizing, predicting), contextualizing (creating scenarios, imagining), structuring, monitoring, and revising. In a similar analysis, Kirby (1986) analyzed videotapes of five basic-level students and found that, throughout the reading comprehension and writing processes, these five students more frequently exhibited similar rather than different strategies. For example, if the students failed to plan when composing a text, they similarly failed to set goals and plan before reading. She concluded that many of the shortcomings of students while reading or writing were similar across both tasks.

In one of the largest studies of the reading and writing connection to date, Langer (1986) investigated the reading and writing behaviors of 67 students in the third, sixth, and ninth grades. She analyzed shared background knowledge, as well as monitoring, reasoning, and strategy use during students' reading and writing of stories and reports. She found that students' behaviors differed as a function of their age and the difficulty of the specific task. Additionally, she found that reading comprehension and writing called upon a number of similar cognitive processes, such as reasoning and constructing

meaning. However, while both reading comprehension and writing were associated with similar processes, there were strong differences in the frequency and nature of their use. For instance, there was a stronger focus on surface-level text issues, such as syntax, vocabulary, and mechanics when writing, as compared to reading. Langer concluded that while reading and writing are similar constructs, they have vastly different origins. For instance, readers are constrained by an author's particular word choice, rhetorical style, and content when generating meaning from a text. Writers, on the other hand, have fewer constraints and can more freely draw information from their own background knowledge of a topic.

Working memory. From the cognitive perspective, research has predominantly investigated the role of lower level skills, such as working memory, on both reading comprehension and writing performance tasks. Significant correlations have consistently been reported between individuals' working memory scores and their performance on higher level cognitive tasks such as following directions, note taking, and learning to write computer programs (Engle, 2002). Specifically, researchers have pointed to the role of working memory in the comprehension (Daneman & Carpenter, 1980; Engle, Cantor, & Carullo, 1992; Yuill, Oakhill, & Parkin, 1989) and production of text (Berninger & Swanson, 1994; Kellogg, 2008; McCutchen, 1996). For instance, in a longitudinal study of 7-year-old children, Oakhill, Cain, and Bryant (2003) showed that scores on working memory tasks accounted for significant variance in scores on reading comprehension tasks. Similarly, Berninger et al. (1994) identified significant relations between the working memory capacity and writing ability of fourth, fifth, and sixth graders. Only a relatively smaller number of studies have *simultaneously* investigated the role of working

memory capacity on both reading comprehension and writing tasks. In one such study, Babayigit and Stainthorp (2011) followed two groups of children (2nd grade and 4th grade) for one year (into the 3rd and 5th grades, respectively). Students' completed a battery of assessments that were intended to measure the component processes of reading fluency, spelling accuracy, reading comprehension and narrative text writing. Results of this study indicated that working memory was moderately related to both reading comprehension and writing at both time points; however, working memory capacity failed to make unique contributions when other measures, such as vocabulary, were considered.

Importantly, the majority of these studies utilized working memory tasks that required participants to utilize some skills that are intrinsically related to reading comprehension and writing ability (e.g., vocabulary knowledge). Studies that have used other measures of working memory (i.e., non-reading-based measures of working memory), on the other hand, have typically failed to report similar correlations between working memory scores and literacy performance (Baddeley, Logie, Nimmo-Smith, & Brereton, 1985; Dixon, LeFevre, & Twilley, 1988). Thus, the specific role of working memory on reading comprehension and writing proficiency remains unclear among educational and cognitive researchers.

Overall, these studies are only a small subset of the large body of literature revealing the strength of connections between reading comprehension and writing. Based on these studies and extensive reviews of the literature (Nelson & Calfee, 1998; Shanahan & Tierney, 1990; Tierney & Shanahan, 1991), it is clear that both literacy skills draw upon similar procedures and knowledge bases. What many of these studies lack,

however, is a focus on the shared *higher level* cognitive skills between reading comprehension and writing. Thus, it is still relatively unclear how higher level cognitive skills relate to performance across both literacy tasks.

Higher level cognitive skills. Research on the relations between reading comprehension and writing has placed less of an emphasis on the role of higher level cognitive skills. Although there is a large body of literature that examines the role of such processes for reading comprehension (e.g., Hannon & Daneman, 2001), much less is known about how these processes play a similar or different role in the writing process. Reading comprehension researchers have attempted to isolate the *component processes* (i.e., individual processes) that allow a reader to develop meaningful and coherent representations of the texts that they are reading (i.e., using knowledge about the world and in the text to develop a deep understanding of a concept). In multiple studies, a cognitive component processes task developed by Hannon and Daneman (2001) has emerged as a strong method of predicting students' reading comprehension abilities, accounting for over 60% of the variance in standardized reading comprehension measures (Daneman & Hannon, 2001; Hannon, 2012; Hannon & Daneman, 2006). This component processes test provides measures of four cognitive abilities: access of prior knowledge from long-term memory, integration of prior knowledge with new information in a text, making inferences based on information in the text, and the ability to recall new information from memory. Research has demonstrated the relationship of these four components to text *comprehension*, but not to text *production*. In light of the constructive theories of comprehension and production, however, it may be reasoned that

these component processes might reflect important similarities and differences between the higher level cognitive skills underlying both literacy tasks.

The ability to access prior knowledge from long-term memory has been previously tied to achievement on reading comprehension tests. Chabot et al. (1984), for instance, investigated the relations between word recognition, lexical access, and semantic memory access to reading comprehension scores. They found that reading scores were most highly influenced by students' ability to quickly access their semantic memory. Similarly, the integration of prior knowledge with new information has been found to be a considerable source of individual differences in reading comprehension ability (Singer & Richot, 1996). Numerous taxonomies for inferences in reading comprehensions have been proposed, with little consensus among researchers (Graesser et al., 1994). However researchers have rarely contested that the ability to make inferences based upon information in the text can lead to improved performance in comprehension (Bloom et al., 1990). Finally, Masson and Miller (1983) found that the ability to remember the information provided in a text is directly related to comprehension of that text. Overall, these components have strong empirical ties to proficiency in reading comprehension. Thus, due to the similar *constructive* and *meaning-making* nature of both reading and writing, these components may be an area of shared variance between both processes.

CURRENT STUDY

The current study investigates how students' lower- and higher level cognitive skills relate to their performance on reading comprehension and writing tasks. Specifically, we examine the degree to which cognitive skills that are associated with the construction of

meaning contribute to both reading comprehension and writing performance.

Importantly, this study is specifically focused on the reading *comprehension* process.

Therefore, we do not aim to make claims about more general and low-level reading tasks, such as the process of decoding or processing syntax. In this study, we first confirm the relationship of cognitive skills to reading comprehension performance and then examine the degree to which these skills play similar or unique roles in writing proficiency.

METHOD

Participants

The participants were 108 undergraduate college students from a university located in southwestern United States. The students were between 18 and 37 years of age ($M=19.75$), and the majority of the participants were either in their first or second year of college. Of the 108 students, 48.1% were female; 53.7% were Caucasian, 22.2% were Hispanic, 10.2% were Asian, 3.7% were African-American, and 9.3% reported “other.” Additionally, 25.9% of the students reported that they were second language speakers of English, while 74.1% reported that they were native speakers. Importantly, all students were enrolled in regular classes at the university; therefore, the English as a Second Language (ESL) students had strong enough English skills to perform regular, non-ESL coursework. The participants were recruited through the psychology department participant pool and given credit in their Introductory Psychology course for participation in the study. Seven participants had missing Aospa data due to computer failure. The data for these participants were analyzed using pairwise deletion.

Materials

Demographics Questionnaire. The participants completed a questionnaire including questions about their age, year in college, gender, ethnicity, and native language.

Writing Performance Assessment. Each student wrote a timed (25-minute), prompt-based, argumentative essay.

Essay prompt. The essays were written in response to a prompt written in the style of the Scholastic Aptitude Test (SAT). The prompt and instructions are provided below:

You will now have 25 minutes to write an essay on the prompt below. The essay gives you an opportunity to show how effectively you can develop and express ideas. You should, therefore, take care to develop your point of view, present your ideas logically and clearly, and use language precisely.

While serious thinking about important matters may disturb people in the short term, it benefits them immeasurably in the long term. Only by confronting unpleasant truths and by weighing both sides of complex issues can people understand the facts—whether in history, politics, literature, or their own lives—and make appropriate decisions. People may find it difficult, or uncomfortable, to think seriously about important matters, but not doing so means that they are leading lives without meaning or purpose.

Does every individual have an obligation to think seriously about important matters, even when doing so may be difficult?

Plan and write an essay in which you develop your point of view on this issue. Support your position with reasoning and examples taken from your reading, studies, experience, or observations

Writing score. Writing ability was assessed using expert human scores. Two expert human raters scored each essay independently using a 6-point rating scale developed for the SAT. The rating scale was used to holistically assess the quality of the essays and had a minimum score of 1 and a maximum score of 6. Raters were first trained to use the rubric with a small sample of similar essays. A Cronbach's Alpha (α) was used to assess inter-rater reliability between raters. When the raters reached an $\alpha = .70$, the ratings were considered reliable and the raters scored a larger subsection of the corpus. The final inter-rater reliability for the raters for the essays was $\alpha > .89$. Raters were then given the opportunity to adjudicate any ratings wherein the disagreement between the raters was greater than one. After adjudication, exact accuracy between raters was 67% and adjacent accuracy was 100%. Average scores between the raters were calculated for each essay to provide a holistic score.

Reading Comprehension Ability. Reading comprehension ability was assessed with the Gates-MacGinitie (4th ed.) reading skill test (form S) level 10/12 (MacGinitie & MacGinitie, 1989). The comprehension test comprises 48 multiple-choice questions that assess students' reading comprehension ability across short passages. Each passage is associated with two to six questions. The questions assess shallow text comprehension as well as deeper level comprehension that require the reader to make inferences about the text. The participants were administered the standard instructions, including two practice questions, and given 20 minutes to complete the test.

Vocabulary Knowledge. The vocabulary section of the Gates-MacGinitie (4th ed.) reading test (form S) level 10/12 (MacGinitie & MacGinitie, 1989) was used to assess the participants' vocabulary knowledge. The test is comprised of 45 simple sentences, each with an underlined vocabulary word. For each underlined word, participants are asked to select the most closely related word from a list of five choices. The sentences are designed to suggest the vocabulary word's part of speech but provide no contextual information about the word's meaning. Participants were administered the standard instructions, including two practice questions, and given 10 minutes to complete the test.

Working Memory Capacity. Participants completed the Automated Operation Span (Aospan; Unsworth, Heitz, Schrock, & Engle, 2005) task to provide a measure of their working memory capacity. We chose to use this non-verbal, mathematics-based version of the complex span task to ensure that any variance that was shared between working memory scores and the other measures was due to an executive control component, rather than verbal or reading comprehension-related skills. In this task, participants are asked to remember a series of letters while performing simple math problems. In each trial, participants view an equation on the computer screen and are instructed to press a button after solving the equation. Participants are then presented with a potential solution and asked to judge its accuracy. Feedback is presented to the participants and a random letter then appears to be recalled at a later time. After solving a set of three to seven math equations, participants are presented with 12 letters and asked to select the letters they recognize in the correct order. Feedback on both math and letter recall accuracy is then presented to the participants. Participants

are asked to maintain their equation accuracy at approximately 85% throughout the task. The Aospan task required approximately 15-20 minutes to complete and concluded after 75 trials. A participant's Aospan Total reflected the total number of letters correctly recognized and correctly ordered.

Component Processes Task. Participants completed Hannon and Daneman's (2001) component processes task, which is intended to measure individual differences in four cognitive component processes of reading comprehension: the ability to access prior knowledge from long-term memory, to integrate accessed prior knowledge with new text information, to make inferences based on information provided in the text, and to recall the new text information from memory.

The participants are explicitly instructed to use their world knowledge throughout the task. The task consists of six three-sentence paragraphs, each accompanied by 18 true-false statements. The sentences in the paragraphs are variably composed of nonsense terms (e.g., MIRT, COFT) and real terms (e.g., WATERMELON, OSTRICH). Each sentence links two terms together using two to four comparative features (e.g. A MIRT resembles an OSTRICH but is larger and has a longer neck.). Overall, each paragraph comprises three nonsense terms, two real terms, and two to four semantic features. An example paragraph is provided below:

A MIRT resembles an OSTRICH but is larger and has a longer neck.

A COFT resembles a ROBIN but is smaller and has a longer neck.

A FLIP resembles a COFT but is smaller, has a longer neck, and nests on land.

For each trial, participants are asked to study the sentences one at a time as they appear on the screen. After each paragraph, participants are asked true or false statements

about the information provided in the sentences. Each true or false statement is one of four types: text memory statements, text inferencing statements, knowledge access statements, and knowledge integration statements. Accuracy is measured as the percentage of correct responses for each overall statement type (text memory, text inferencing, knowledge access, and knowledge integration).

Text memory statements assess information that is explicitly mentioned in the paragraph without calling upon the use of outside world knowledge (e.g. A MIRT is larger than an OSTRICH). Text inferencing statements assess inferences about information provided explicitly in the paragraph without the use of world knowledge (e.g. A FLIP has a larger neck than a ROBIN). Knowledge access statements assess participants' ability to access prior world knowledge but require no information presented in the paragraph (e.g. A BLUEJAY lives in Canada, whereas an OSTRICH typically doesn't). Knowledge access statements are divided into two types: low and high. Low-knowledge access statements test access to a fact that is not explicitly presented in the paragraph, but included two real terms and a feature included in the given paragraph (e.g., An OSTRICH has a longer neck than a ROBIN). High-knowledge access statements also include two real terms and test access to a fact that is not presented in the paragraph (e.g., A ROBIN lives in Canada, whereas a PENGUIN typically doesn't). However, they only include one real term presented in the paragraph. The other real term and the semantic feature are not presented in the paragraph. Finally, knowledge integration statements assess participants' ability to integrate prior world knowledge with the information provided in the paragraphs (e.g. A PENGUIN is larger than a COFT). Knowledge integration statements are divided

into three types that increase in difficulty level: low, medium, and high. Low-knowledge integration statements include one nonsense term, one real term, and a semantic feature, all of which are presented in the paragraph (e.g., A MIRT has a longer neck than a ROBIN). In the medium-knowledge integration statements, the nonsense term and the semantic feature are presented in the paragraph, but the real term is not (e.g., A MIRT is larger than a BLUEJAY). Finally, the high-knowledge integration statements contain a nonsense term that is in the paragraph, but also contain a real term and a semantic feature that are not presented in the paragraph (e.g., Like PENGUINS, MIRTS can't fly).

Procedure

The study comprised one laboratory session that lasted approximately 2 hours. The duration and order of these tasks are as follows; a demographics questionnaire (approximately 5 minutes), timed-essay (approximately 25 minutes), vocabulary test (approximately 10 minutes), and reading comprehension test (approximately 20 minutes), component processes task (approximately 30) and Aospan task (approximately 20 minutes). The Aospan and components processes tasks were presented using E-Prime software (Psychology Software Tools, Pittsburgh, PA), all other measure were presented using an online survey. In the case of the persuasive essay task, participants were not allowed to proceed before the 25 minutes had elapsed. However, for all other tasks, participants moved on to the subsequent task as soon as they had completed the task. For the Gates-MacGinitie vocabulary and comprehension tests, the maximum time allowed to complete the tasks was 10 and 20 minutes, respectively.

RESULTS

Correlation and regression analyses were conducted to examine the relations among students' reading comprehension scores, writing scores, and the six individual difference measures (i.e., vocabulary knowledge, working memory capacity, and the four cognitive component processes). We first confirm relations between students' scores on the collected measures and their reading comprehension ability. We then investigate how and whether students' scores on these measures were correlated with, and predictive of, their performance on a persuasive essay writing task.

Descriptive and Correlation Analyses

Table 5 presents the means, standard deviations, and ranges for students' scores on the collected measures (i.e., Gates-MacGinitie Reading Comprehension Test, Persuasive Essay, Gates-MacGinitie Vocabulary Test, Component Processes Task, and AOSPAN Task), as well as the Pearson correlations among all of the measures. Multicollinearity between the 6 predictor variables (Component Processing Task, AOSPAN, and vocabulary knowledge) was assessed using a threshold of $r > .90$ (Tabachnick & Fidell, 2001). None of the 6 variables demonstrated multicollinearity, therefore, none were removed from the current analyses. It is important to note, however, that some of the measures were strongly correlated; thus, caution should be exercised when interpreting the results of the analyses. As revealed in Table 5, AOSPAN Total was the only score not significantly correlated with all of the other collected measures. Because students' reading comprehension, writing, vocabulary, and component processes scores all exhibited significant relations, it can be inferred that both reading

comprehension and writing ability rely on a related set of vocabulary knowledge and higher level cognitive skills.

Table 5.

Descriptive Statistics and Pearson Correlations for Measures Related to Reading and Writing

Variables	1	2	3	4	5	6	7	8	<i>M</i>	<i>SD</i>	Range
1. Reading Comprehension	1.00								29.40	9.99	10-48
2. Writing Score	.57**	1.00							3.02	1.14	1-6
3. Vocabulary Knowledge	.79**	.55**	1.00						30.89	9.52	6-45
4. <i>CP</i> : Text Memory ^a	.55**	.25**	.40**	1.00					64.84	14.35	35.71-97.62
5. <i>CP</i> : Text Inferencing ^a	.51*	.25**	.34**	.83**	1.00				61.16	15.24	25.00-88.89
6. <i>CP</i> : Knowledge Access ^a	.68**	.41**	.68**	.53**	.39**	1.00			80.33	12.50	46.67-98.33
7. <i>CP</i> : Knowledge Integration ^a	.58**	.31**	.51**	.81**	.66**	.71**	1.00		68.45	14.54	43.75-98.96
8. AOSPAN Total	.10	.01	.12	.22*	.14	.14	.20*	1.00	56.44	11.79	23-75

Notes: *CP* = component process; ^aStatistics for the component scores are reported as percentages

* $p < .05$. ** $p < .001$

Reading comprehension scores were strongly correlated with writing scores at $r = .57, p < .001$, yielding a large effect size (see Cohen, 1988), which is in line with previous literature that has reported correlations between reading and writing scores ranging from $r = .20$ to $r = .50$ (Fitzgerald & Shanahan, 2000). The moderate to strong correlation typically found between reading and writing suggests that the two processes share some degree of commonality; however, because perfect correlations have not been

reported between the two, this also implies that they are separable constructs and potentially exhibit unique relations with the collected measures.

Reading Comprehension. Consistent with previous research, students' performance on the reading comprehension test was most strongly related to their vocabulary knowledge (Perfetti, 1985). In other words, the successful comprehension of texts was more difficult for students who possessed less explicit knowledge of the words in a given text. In addition to vocabulary knowledge, participants' ability to comprehend passages was related to several higher level cognitive abilities, providing confirmatory evidence for their role in the comprehension process.

The components most highly related to reading comprehension scores were the ability to access prior knowledge from long-term memory and the ability to integrate this knowledge with new information. Thus, in addition to the need for explicit word knowledge, comprehension also depended on students' ability to access and integrate this prior knowledge with new information. In addition, reading comprehension was significantly related to the remaining two component scores: drawing text-based inferences based on a text and recall of information explicitly presented in the text. Therefore, aside from the ability to access and integrate prior knowledge, the ability to remember, understand, and make deductions based on a given text was strongly related to students' scores on reading comprehension tests. AOSPAN was the only variable that was not significantly related to reading comprehension scores, suggesting that variations in the students working memory capacity were not driving differences in their ability to comprehend texts. Overall, the results suggest that prior knowledge and the higher level

cognitive skills associated with the use of this knowledge are strongly related to reading comprehension proficiency.

Writing Proficiency. Correlations were calculated between the collected measures and students' essay scores. Results from these analyses indicate that the writing process shares some knowledge sources and higher level cognitive abilities associated with reading comprehension performance.

The correlations indicate that students' essay scores were most strongly related to their vocabulary knowledge. Thus, similar to reading comprehension, the writing process is largely dependent on word knowledge. Students who knew fewer words on the vocabulary test tended to have greater difficulty producing quality essays. On the other hand, unlike the reading comprehension process, the correlation between vocabulary knowledge and writing was only moderate and, therefore, more weakly related to essay scores than reading comprehension scores.

Second, the remaining measures exhibited correlations comparable, albeit weaker, to those observed with reading comprehension scores. Specifically, the ability to recognize previously viewed information, the ability to make inferences about a given text, the ability to access prior knowledge from memory, and the ability to integrate this knowledge with new information were all significantly correlated with essay scores. Thus, although reading comprehension and writing scores were both related to vocabulary knowledge and the higher level cognitive skills, writing scores exhibited much weaker relationships to the measures overall. Finally, and also similar to reading comprehension, was the correlation between essay scores and AOSPAN. In this study,

variations in students' working memory scores did not significantly relate to variations in their performance on either the comprehension or essay writing tasks.

Accordingly, it may be the case that reading comprehension and writing are similar cognitive processes, but the writing process is more reliant on additional, knowledge sources and higher level cognitive abilities (which were not assessed in this study). Overall, the results of the correlation analyses suggest that the higher level cognitive skills related to the reading comprehension process may play a similar role in the writing process, albeit a somewhat weaker one.

Regression Analyses

Reading Comprehension. To determine whether the higher level cognitive skills were predictive of students' reading comprehension performance, a linear regression analysis was conducted to predict reading comprehension from the four component processes scores. This analysis yielded a significant model $F(4, 101) = 28.06, p < .001; R^2 = .53,$ ² with two significant predictors: text inferencing and knowledge access (see Table 6). Results of this analysis suggest that students' performance on the reading comprehension assessment was strongly influenced by their ability to make inferences about the text and their ability to access prior knowledge about the given topics.

A follow-up hierarchical linear regression analysis was conducted to assess whether the component scores accounted for unique variance in students' reading comprehension scores over and above vocabulary knowledge. Therefore, vocabulary knowledge was entered as the first block of a regression analysis and the four component scores were entered into a second block.

² Note that a portion of the variance in the final R^2 for all of the regression analyses is explained by non-significant variables.

Table 6.

*Linear Regression Analysis for Component Processes Predicting Reading**Comprehension Scores and Writing Scores*

Variable	Reading Comprehension Scores				Writing Scores			
	B	SE B	B	t	B	SE B	B	t
Constant	-18.58	4.66		-3.99**	-.22	.70		-.31
<i>CP</i> : Text Memory	.05	.11	.07	.44	-.01	.02	-.10	-.51
<i>CP</i> : Text Inferencing	.17	.08	.26	2.08*	.02	.01	.20	1.24
<i>CP</i> : Knowledge	.46	.08	.58	5.84**	.04	.01	.42	3.26*
Access								
<i>CP</i> : Knowledge	-.04	.10	-.05	-.38	-.01	.01	-.05	-.25
Integration								
	$R^2 = .53$				$R^2 = .19$			

Notes: *CP* = component process; * $p < .05$. ** $p < .001$

This analysis yielded two significant models. The first model confirmed that vocabulary knowledge significantly predicted reading comprehension scores, $F(1, 99) = 169.43, p < .001; R^2 = .63$ (see Table 7). Further, the second model revealed that the component scores accounted for significant variance above and beyond the variance accounted for in the first block, $F(5, 95) = 47.25, p < .001; R^2 = .71$. One of the component scores was significant (knowledge access) and another was marginally significant (text inferencing). The results of this regression analysis suggest that the ability to comprehend text is strongly predicted by students' word knowledge, as well as higher level cognitive skills associated with the access of prior knowledge and the generation of inferences.

Writing Proficiency.

Table 7.

Linear Regression Analysis for Measures Predicting Reading Comprehension Scores and Writing Scores

Variable	Reading Comprehension Scores				Writing Scores			
	B	SE B	B	ΔR	B	SE B	B	ΔR
Model 1				.63**				.30**
Vocabulary Knowledge	.07	.01	.55**		.07	.01	.55**	
Model 2				.08**				.01
Vocabulary Knowledge	.62	.08	.59**		.06	.01	.48**	
CP: Text Memory	.10	.10	.12		-.01	.02	-.06	-.30
CP: Text Inferencing	.31	.18	.17		.03	.03	.13	.82
CP: Knowledge Access	.25	.12	.19*		.02	.02	.10	.72
CP: Knowledge	-.05	.08	-.06		-.01	.01	-.05	-.31
Integration								
				$R^2 = .72$				$R^2 = .31$

Notes: CP = component process; * $p < .05$. ** $p < .001$

A linear regression analysis was conducted to determine whether the higher level component processes were predictive of students' writing scores. As Table 7 reveals, the model was significant, $F(4, 101) = 5.72, p < .001; R^2 = .19$, with one significant predictor: knowledge access. The results of this analysis indicate that writing performance had one similar predictor as the reading comprehension analysis (i.e., knowledge access). Thus, the reading comprehension and writing processes both rely on a similar cognitive ability to successfully access prior knowledge about a given topic.

A follow-up hierarchical regression analysis was conducted to examine the degree to which the component scores could predict essay scores over and above vocabulary knowledge. Writing scores were regressed onto vocabulary knowledge in a first block, yielding a significant model, $F(1, 99) = 42.37, p < .001; R^2 = .30$ (see Table 7). The four component scores were then entered into the second block of the model. This yielded a significant model, $F(5, 95) = 8.48, p < .001$; however, the block did not provide significant unique variance above the first block. Overall, the results of this analysis suggest that, although all of the collected measures were correlated with writing proficiency, students' writing performance was most strongly predicted by their vocabulary knowledge.

Overall, the results of the regression analyses confirm the correlation analyses indicating that both reading comprehension and writing were strongly predicted by vocabulary knowledge. The regression analyses also suggest that, aside from knowledge access, individuals' higher level cognitive abilities accounted for little additional variance over vocabulary knowledge. Therefore, despite the significant correlations between these collected measures and reading comprehension and writing scores, only vocabulary knowledge predicted significant, unique variance in both skills.

DISCUSSION

In this study, we examined whether students' reading comprehension and writing scores were similarly or differently related to lower and higher level cognitive skills. The results of the current study indicate that reading comprehension proficiency was strongly related to both vocabulary knowledge and the higher level cognitive skills. Further, writing ability was moderately associated with a subset of the measured variables, namely

vocabulary knowledge and the ability to access prior knowledge (although this did not predict unique variance above the vocabulary scores). The only individual difference measure that was not significantly correlated with either reading comprehension or writing ability was working memory capacity. This was likely due to the working memory measure that was used in this study (i.e., a non-verbal or reading-comprehension related measure). Therefore, it may be the case that working memory capacity is largely unrelated to reading comprehension and writing ability, at least with normal, adult populations. However, future research is needed to make strong claims regarding these relationships. Overall, our results support the hypothesis that reading comprehension and writing ability share common knowledge sources and higher level cognitive skills. However, they also suggest that the writing process is much less reliant on these measured variables than reading comprehension; therefore, the writing process may depend on additional, unmeasured factors, such as writing strategies, writing genres, and grammar.

A valuable contribution of this study is the explicit comparison of reading comprehension and writing performance in terms of higher level cognitive skills. Although these cognitive constructs have been thoroughly examined in the reading comprehension literature, far fewer studies have focused on how they may overlap with students' writing proficiency. Because of this dearth of cognitive studies related to writing proficiency, relatively little is known about the higher level cognitive skills that students employ when they engage in the *process* of text production. As mentioned earlier in this paper, most contemporary cognitive models of the writing process focus on the role of lower level cognitive skills in their examination of text production. However,

the current study revealed that working memory (our measure of lower level cognitive skills) was unrelated to either reading comprehension or writing scores. This result supports the argument that researchers should place a stronger emphasis on the investigation of higher level cognitive skills when studying and developing models of the text comprehension and production processes. Additionally, such studies can help researchers develop more sophisticated conceptualizations of how text comprehension and production overlap, which can inform theoretical models of literacy more generally. The results of this study also provide crucial information that may inform effective pedagogical practices aimed at improving literacy performance. We hypothesize that the shared components identified between the reading comprehension and writing processes can be leveraged to provide more beneficial interventions for struggling students.

Reading-Writing Connections

The results of this study confirm a strong link between the reading comprehension and writing processes (i.e., scores on these measures were correlated at $r = .57$) and identify higher level cognitive skills that are shared between the two literacy activities.

Correlation and regression analyses confirmed that the four cognitive component processes were strongly predictive of reading comprehension scores, accounting for over half of the variance. However, only knowledge access (and marginally, text inferencing) accounted for significant variance over and above vocabulary knowledge. These results support the notion that reading comprehension relies on both prior word knowledge and higher level cognitive processes. While vocabulary knowledge was most highly related to reading comprehension scores, the ability to comprehend text was additionally influenced by higher level cognitive skills, such as the ability to access prior knowledge and make

inferences about a given text. Overall, these analyses confirm prior literature suggesting that successful text comprehension is largely influenced by an individual's degree of word knowledge, prior knowledge, and their ability to make inferences about a text. Future studies will need to further tease apart these relations by examining differential relations between the individual difference measures and literal and inference-generation questions on reading comprehension assessments. Such analyses will provide more fine-grained information about the role of lower- and higher level cognitive skills on students' text comprehension.

Beyond our analysis of reading comprehension scores, we further investigated the degree to which the higher level cognitive skills similarly or differentially impacted writing performance. The results revealed that some of the higher level cognitive skills that were related to reading comprehension performance were also related to students' writing ability. Specifically, the correlation analysis revealed that writing performance was significantly related to all four of the component processes (i.e., text memory, text inferencing, knowledge access, and knowledge integration). In an initial regression analysis, however, knowledge access was the only component process that significantly predicted writing performance, and this significance was washed out once vocabulary knowledge was added to the model. Importantly, these results provide insight into the cognitive and knowledge factors related to the writing process. Additionally, they suggest that variation in students' performance on writing assessments may be largely attributed to individual differences in their vocabulary knowledge and, to a lesser extent, in their ability to successfully activate relevant prior knowledge. Overall, the results of the current study fall in line with previous research, suggesting that vocabulary knowledge

was strongly related to both the ability to comprehend and produce text (e.g., Abbott & Berninger, 1993; Fitzgerald & Shanahan, 2000). However, they also suggest that the two literacy tasks are related to higher level cognitive skills associated with the use of this knowledge.

A limitation of the current study relates to the assessments used to measure both reading comprehension and vocabulary knowledge. First, there currently exists some debate as to whether the Gates-MacGinitie Reading Test accurately evaluates the cognitive skills (specifically related to inferencing) that are required for deep comprehension (Cutting & Scarborough, 2006; van den Broek & Espin, 2012). Therefore, it is possible that the relations between the scores on the component processes task (i.e., those related to higher level cognitive skills) and the reading comprehension task would be much higher if we were to have used a more sophisticated measure of deep comprehension. Future studies should consider the limitations of the Gates-MacGinitie reading test when investigating comprehension – particular when assessing the higher level processes that are related to comprehension. Additionally, although both the reading comprehension and vocabulary assessments were technically separate measures, they were developed as subscales of a holistic measure of reading skill (i.e., there are comprehension and vocabulary sections of the test). While the vocabulary test did not involve comprehension of long passages, the task itself did require students to read simple phrases and sentences. Therefore, it is not a *pure* measure of vocabulary. It is therefore likely that the relations between vocabulary knowledge and the reading comprehension and writing tasks were slightly inflated. To address these issues, future studies should utilize multiple assessments of both reading comprehension and writing.

This will reduce variance attributed to test-level effects and allow researchers to draw stronger conclusions about the relations between higher level skills and reading comprehension and writing. Despite this limitation, however, the results suggest that both reading and writing are similarly related to sources of vocabulary knowledge and the four measured component processes; yet, writing is drawing on other areas of knowledge or processes that were not measured in the current study.

Implications for the Classroom

The results of the current study have important implications for the development of effective literacy pedagogy. First, these findings indicate that word knowledge is critical for the successful completion of both reading comprehension and writing tasks.

Therefore, as the depth of students' vocabulary knowledge increases, their performance on both reading comprehension and writing tasks increases as well. In terms of pedagogy, this result suggests that teachers should consider placing a strong emphasis on vocabulary instruction in the classroom. Unfortunately, vocabulary interventions tend to have limited success (Elleman, Lindo, Morphy, & Compton, 2009; Mol, Bus, & deJong, 2009; Pearson, Hiebert, & Kamil, 2007). Many vocabulary instruction studies have shown improvements in students' decoding ability and in their ability to derive word meanings from context; however, they rarely reveal improvements in students' reading or writing ability (Baumann, Edwards, Boland, Olejnik, & Kame'enui, 2003; Elleman et al., 2009; Mol et al., 2009; Pearson et al., 2007). These results, combined with those reported here, imply that a more successful approach may be to provide vocabulary instruction in the context of the targeted skill, such as reading comprehension or writing. The current findings further indicated that reading comprehension and writing performance are both

influenced by students' abilities to access their prior knowledge. This implies that it may not be enough to know words or concepts. Students must learn skills and strategies that allow them to access and use their knowledge (McNamara & Scott, 1999; McNamara, 2004).

Although this study provided some significant insights into the processes involved in literacy skills, there are many questions that remain to be answered. First, this study provided little information on the role that development might play on the reading-writing connection. Future studies including multiple age groups will provide valuable information regarding whether the patterns of relations change as students develop their knowledge and cognitive skills. Second, the number of measures that could be included in this study was relatively limited. Studies including a wider range of individual difference measures (e.g., motivation, strategy knowledge, or attention control) will provide a more complete picture of reading comprehension and writing processes and the relations between them. Finally, this study did not tease apart differences between native and non-native speakers of English. Certainly, the reading comprehension and writing processes are different in the first and second languages. Therefore, future work is needed to investigate how language differences influence these literacy processes, as well as what individual differences contribute to proficiency in these skills.

CONCLUSIONS

The results of this study and future studies on this topic have the potential to have a strong influence on the design and implementation of literacy curricula. Although reading comprehension and writing are clearly different tasks, research indicates that they overlap in terms of their purposes, processes, and sources of knowledge. Research investigating

the role of cognition in the reading-writing connection, therefore, may provide insight into additional processes and knowledge sources needed to reach proficiency in both reading comprehension and writing. One crucial question to address is whether there are specific cognitive and knowledge-based components of literacy proficiency. To answer this question, research is needed that investigates how multiple factors influence students' proficiency across a variety of literacy tasks. If researchers can identify these components of literacy, educators may be able to determine the skills and knowledge sources that should be most strongly emphasized in the classroom, and the appropriate situations to provide reading and writing instruction in combination versus separately.

Chapter 4

GENERAL DISCUSSION

Previous research on the writing process has largely supported the idea that lower level cognitive skills and language knowledge play an important role in the production of high-quality texts (Abbott & Berninger, 1993; Fitzgerald & Shanahan, 2000). Indeed, results from a number of studies have identified specific individual differences, such as vocabulary knowledge and working memory capacity, that can predict students' performance on writing tasks (e.g., Abbott & Berninger, 1993; Berninger & Swanson, 1994; Fitzgerald & Shanahan, 2000; Hayes, 1996; Kellogg, 2001; 2008; McCutchen, 1996). Often lacking, however, is a focus on the role of *higher level* cognitive skills during the writing process. Thus, while it is clear that students require some degree of knowledge (e.g., of the domain and of the language, etc.) to produce high-quality texts, it is relatively unclear what strategic and metacognitive skills must be developed in order for students to successfully leverage this knowledge during writing tasks. This decreased focus on higher level skills poses a critical problem if researchers aim to accurately model the writing process, and develop interventions that most effectively enhance students' writing performance.

The two studies presented in this thesis project examined the role of higher level cognitive skills during the writing process. In Study 1 (Chapter 2), students' metacognitive skills were examined through a systematic analysis of their criteria for self-assessment. Specifically, the evaluative misalignment hypothesis was proposed, which posits that students assess their own writing according to a different set of criteria than their teachers. Results from automated text analyses revealed that there were, indeed, differences in the essay ratings provided by the students and their teachers. In

particular, teacher ratings were more strongly predicted by the linguistic features when compared to the student ratings, and teachers' ratings were significantly correlated with a larger number of indices than were the students' ratings.

The findings from Study 1 suggest that students may struggle with writing because their criteria for quality writing do not match those of their teachers. Consequently, students may produce texts that fail to meet the goals and standards set by their teachers, and students may also not understand *why* their teachers have evaluated their essays in a certain way. *Evaluative misalignment* may exacerbate the intrinsically difficult nature of writing assessment. Raters must not only consider the overall purpose and argument strength of the text; they must also make decisions about what linguistic features to attend to (e.g., syntax, vocabulary strength) and the degree to which these various essay components should be emphasized (Huot, 1996; Meadows & Billington, 2005). For expert human raters (i.e., teachers and other trained, professional raters), these evaluations are guided by a number of linguistic text properties, ranging from the complexity of students' syntactic constructions to the semantic and rhetorical features of the essay (Freedman, 1979; McNamara, Crossley, & McCarthy, 2010; Witte & Faigley, 1981). Additionally, expert raters place different relative weights on each of these linguistic properties; therefore, they may be more or less attuned to certain features depending on various other properties of the text. These considerations may be more or less automatic for expert raters given their expertise in writing.

By contrast, students may lack sufficient strategies and knowledge to rate their own essays based on these sets of linguistic features (Donovan & Smolkin, 2006; Graham, 2006; Wong, 1999). In general, these results point to the importance of higher

level *metacognitive* skills during the writing process. Although previous research suggests that prompting students to self-assess can improve their writing (Graham & Perin, 2007; Hillocks, 1986), the findings from the current study suggest that these self-assessments may not necessarily be accurate. Thus, students may need to be provided with explicit instruction on both the criteria that teachers utilize to assess writing, as well as (and perhaps most importantly) the strategies needed to make decisions about which linguistic properties to attend to.

Study 2 similarly examined the role of higher level cognitive skills in the writing process. In this study, we employed an individual differences approach to investigate how variations in certain higher level cognitive skills were related to students' performance on writing tasks. The rationale for Study 2 was that the text production process has (theoretically, at least) a number of similarities with the process of comprehending texts – namely, both tasks are communicative, language-based activities, which involve the active construction of meaning. Given these similarities, our goal was to examine the degree to which lower and higher level cognitive skills that are commonly associated with reading comprehension skills relate to performance on writing tasks.

The results of Study 2 confirmed previous research, indicating that students' scores on a reading comprehension test were strongly related to their word knowledge, as well as a number of higher level cognitive skills. Further, these students' scores on the writing task were moderately associated with a subset of these measured variables, most strongly, vocabulary knowledge and the ability to access prior world knowledge. A strong contribution of this study was the explicit examination of the relationships between students' higher level cognitive skills and their writing performance. Although

these cognitive skills have been previously examined in research on reading comprehension, substantially less information is known about how these skills contribute to writing performance. The results of Study 2, therefore, support the notion that researchers should place a stronger focus on the role of higher level cognitive skills in their investigations of the writing process. Studies such as these can help to inform models of the writing process, as well as the development of interventions aimed to improve students' writing performance.

Overall, the results of both Study 1 and Study 2 indicate that higher level cognitive skills are important components of the writing process. Specifically, Study 1 focused on the role of metacognition in the writing process, revealing that students assess their essays based on substantially different criteria than their teachers. Study 2 investigated higher level cognitive skills that are most typically associated with reading comprehension. These results revealed that the ability to recognize text from memory, access prior knowledge, integrate prior knowledge with new knowledge, and make inferences all relate to students' performance on writing tests. Obviously, the results from this thesis project are not sufficient to provide a comprehensive understanding of the role that higher level skills play in the writing process. Future studies are needed to further the hypotheses and theoretical perspectives discussed in these two manuscripts.

For example, in a more recent study, we have begun to examine the *interactions* between students' lower- and higher level skills during the writing process. As previously mentioned, writing researchers have placed a particularly strong emphasis on the role of individuals' working memory capacity in their production of high quality texts. However, the link between working memory capacity and writing skill has failed to be consistently

supported by the literature. In this recent study, we examine the hypothesis that the role of working memory capacity in the writing process is modulated by the development and use of strategic skills (i.e., inferencing skills). We then employ a computational linguistic approach to test this hypothesis by relating lower- and higher level linguistic properties of the students' essays to measures of their working memory capacity. The results suggest that working memory scores are most strongly related to lower level properties of students' essays. Most importantly, however, they indicate that these relationships between the essay properties and working memory scores are reduced for students with high inferencing skills. Hence, working memory capacity is related to the properties of less skilled students' writing, but primarily for lower level aspects of the writing such as word choice, whereas working memory plays a lesser role for more skilled students, whose writing is characterized more by higher level aspects such as cohesion.

This is one example of the type of future research that can be conducted in this area. However, there are countless research questions that can be asked to develop a stronger understanding of these higher level skills within the context of the writing process. For instance, how might the influence of higher level cognitive skills differ according to various populations and contexts? Additionally, are the same skills important for persuasive writing as for other forms of writing, such as creative writing? Overall, the studies in the current thesis project point towards a strong role of higher level skills in the writing process, and provide a strong foundation on which to develop future research. Ideally, once researchers have identified the specific roles that these higher level skills play in the writing process, their findings can be used to develop interventions that effectively improve students' performance on a variety of writing tasks.

REFERENCES

- Abbott, R., & Berninger, V. W. (1993). Structural equation modeling of relationships among developmental skills and writing skills in primary- and intermediate-grade writers. *Journal of Educational Psychology, 85*, 478-508.
- Andrade, H., & du Boulay, B. (2003). The role of rubric-referenced self-assessment in learning to write. *Journal of Educational Research, 97*, 21-34.
- Andrade, H., Du, Y., & Wang, X. (2008). Putting rubrics to the test: The effect of a model, criteria generation, and rubric-referenced self-assessment on elementary school students' writing. *Educational Measurement: Issues and Practices, 27*, 3-13.
- Andrews, R., Rogerson, C., Beverton, S., Freeman, A., Locke, T., Low, G., Robinson, A., & Zhu, D. (2006). The effect of grammar teaching on writing development. *British Educational Research Journal, 32*, 39-55.
- Babayigit, S., & Stainthorp, R. (2011). Modeling the relationships between cognitive-linguistic skills and literacy skills: New insights from a transparent orthography. *Journal of Educational Psychology, 103*, 169-189.
- Baddeley, A. D., Logie, R., Nimmo-smith, I., & Brereton, N. (1985). Components of fluent reading. *Journal of Memory and Language, 24*, 119-131.
- Baer, J. D., & McGrath, D. (2007). The reading literacy of U.S. fourth-grade students in an international context: Results from the 2001 and 2006 Progress in International Literacy Study (PIRLS). National Center for Educational Statistics, Institute of Education Sciences, U.S. Department of Education.
- Baumann, J. F., Edwards, E. C., Boland, E. M., Olejnik, S., & Kame'enui, E. J. (2003). Vocabulary tricks: Effects of instruction in morphology and context on fifth-grade students' ability to derive and infer word meanings. *American Educational Research Journal, 40*, 447-494.
- Bereiter, C., & Scardamalia, M. (1987). *The psychology of written composition*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Berninger, V. W., Abbott, R. D., Abbott, S. P., Graham, S., & Richards, T. (2002). Writing and reading: Connections between language by hand and language by eye. *Journal of Learning Disabilities, 35*, 39-56.
- Berninger, V. W., Cartwright, A. C., Yates, C. M., Swanson, L., & Abbott, R. D. (1994). Developmental skills related to writing and reading acquisition in the intermediate grades: Shared and unique functional systems. *Reading and Writing, 6*, 161-196.

- Berninger, V. W., & Swanson, H. L. (1994). Modifying Hayes and Flower's model of skilled writing to explain beginning and developing writers. In E. C. Butterfield (Ed.), *Children's writing: Toward a process theory of the development of skilled writing* (pp. 57-81). Greenwich, CT: JAL.
- Birnbaum, J. C. (1982). The reading and composing behavior of selected fourth- and seventh-grade students. *Research in the Teaching of English, 16*, 241-260.
- Bloom, C. P., Fletcher, C. R., van den Broek, P., Reitz, L., & Shapiro, B. P. (1990). An on-line assessment of causal reasoning during comprehension. *Memory and Cognition, 18*, 65-71.
- Camara, W. J. (2003). *Scoring the essay on the SAT writing section*. New York: College Entrance Examination Board.
- Chabot, R. J., Zehr, H. D., Prinzo, O. V., & Petros, T. V. (1984). The speed of word recognition subprocesses and reading achievement in college students. *Reading Research Quarterly, 19*, 147-161.
- Chenoweth, A., & Hayes, J. (2001). Fluency in writing: Generating text in L1 and L2. *Written Communication, 18*, 80-98.
- Chenoweth, A., & Hayes, J. (2003). The inner voice in writing. *Written Communication, 20*, 99-118.
- Cho, K., & Schunn, C. D. (2007). Scaffolded writing and rewriting in the discipline: A web-based reciprocal peer review system. *Computers and Education, 48*, 409-426.
- Clarke, L. K. (1988). Invented versus traditional spelling in first graders' writings: Effects on learning to spell and read. *Research in the Teaching of English, 22*, 281-309.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.
- Corden, R. (2007). Developing reading-writing connections: The impact of explicit instruction of literary devices on the quality of children's narrative writing. *Journal of Research in Childhood Education, 21*, 269-289.
- Couzijn, M. (1999). Learning to write by observation of writing and reading processes: Effects on learning and transfer. *Learning and Instruction, 9*, 109-142.
- Crossley S. A., Kyle, K., Allen, L. K., & McNamara, D. S. (2014). The importance of grammar and mechanics in writing assessment and instruction: Evidence from data mining. In J. Stamper, Z. Pardos, M. Mavrikis, & B. M. McLaren (Eds.), *Proceedings*

of the 7th International Conference on Educational Data Mining (pp. 300-303). London, UK.

- Crossley, S. A. & McNamara, D. S. (2010). Cohesion, coherence, and expert evaluations of writing proficiency. In S. Ohlsson & R. Catrambone (Eds.), *Proceedings of the 32nd Annual Conference of the Cognitive Science Society* (pp. 984-989). Austin, TX: Cognitive Science Society.
- Crossley, S. A., & McNamara, D. S. (2011). Understanding expert ratings of essay quality: Coh-Metrix analyses of first and second language writing. *International Journal of Continuing Engineering Education and Life-Long Learning*, 21, 170-191.
- Crossley, S. A., & McNamara, D. S. (2012). Predicting second language writing proficiency: The roles of cohesion and linguistic sophistication. *Journal of Research in Reading*, 35, 115-136.
- Crossley, S. A., Weston, J., McLain-Sullivan, S. T., & McNamara, D. S. (2011). The development of writing proficiency as a function of grade level: A linguistic analysis. *Written Communication*, 28, 282-311.
- Cutting, L. E., & Scarborough, H. S. (2006). Prediction of reading comprehension: Relative contributions of word recognition, language proficiency, and other cognitive skills can depend on how comprehension is measured. *Scientific Studies of Reading*, 10, 277-299.
- Daneman, M., & Carpenter, P. A. (1980). Individual differences in working memory and reading. *Journal of Verbal Learning and Verbal Behavior*, 19, 450-466.
- Daneman, M., & Hannon, B. (2001). Using working memory theory to investigate the construct validity of multiple-choice reading comprehension tests such as the SAT. *Journal of Experimental Psychology: General*, 130, 208-223.
- Dixon, P., LeFevre, J., & Twilley, L. C. (1988). Word knowledge and working memory as predictors of reading skill. *Journal of Educational Psychology*, 80, 465-472.
- Donovan, C. A., & Smolkin, L. B. (2006). Children's understanding of genre and writing development. In C. A. MacArthur, S. Graham, & J. Fitzgerald (Eds.), *Handbook of writing research* (pp. 131-143). New York: Guilford.
- Dunning, D., Johnson, K., Ehrlinger, J., & Kruger, J. (2003). Why people fail to recognize their own incompetence. *Current Directions in Psychological Science*, 12, 83-86.
- Elleman, A. M., Lindo, E. J., Morphy, P., & Compton, D. L. (2009). The impact of vocabulary instruction on passage-level comprehension of school-age children: A meta-analysis. *Journal of Research on Educational Effectiveness*, 2, 1-44.

- Engle, R. W. (2002). Working memory capacity as executive attention. *Current Directions in Psychological Science, 11*, 19-23.
- Engle, R. W., Cantor, J., & Carullo, J. J. (1992). Individual differences in working memory and comprehension: A test of four hypotheses. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 18*, 972-992.
- Falchikov, N., & Boud, D. (1989). Student self-assessment in higher education: A meta-analysis. *Review of Educational Research, 59*, 395-430.
- Fellbaum, C. (Ed.). (1998). *WordNet: An electronic lexical database*. Cambridge, MA: MIT Press.
- Fitzgerald, J. & Shanahan, T. (2000). Reading and writing relations and their development. *Educational Psychologist, 35*, 39-50.
- Flower, L. S., & Hayes, J. (1981). A cognitive process theory of writing. *College Composition and Communication, 32*, 365-387.
- Freedman, S. W. (1979). How characteristics of student essays influence teachers' evaluations. *Journal of Educational Psychology, 71*, 328-338.
- Galbraith, D. (1999). Writing as a knowledge-constituting process. In M. Torrance & D. Galbraith (Eds.), *Knowing what to write* (pp. 139-169). Amsterdam, NL: Amsterdam University Press.
- Galda, L. (1983). Research in response to literature. *Journal of Research and Development in Education, 16*, 1-7.
- Gesier, S., & Studley, R. (2001). *UC and SAT: Predictive validity and differential impact of the SAT I and SAT II at the University of California*. Oakland, CA: University of California.
- Gernsbacher, M. A. (1997). Coherence cues mapping during comprehension. In J. Costermans & M. Fayol (Eds.), *Processing interclausal relationships in the production and comprehension of text* (pp. 3-21). Mahwah, NJ: Erlbaum.
- Gill, A. J., French, R. M., Gergle, D., & Oberlander, J. (2008). The language of emotion in short blog texts. *Proceedings of CSCW* (pp. 299-302). New York, NY: ACM Press.
- Graesser, A. C. & McNamara, D. S. (2011). Computational analyses of multilevel discourse comprehension. *Topics in Cognitive Science, 2*, 371-398.
- Graesser, A. C., McNamara, D. S., Louwerse, M., & Cai, Z. (2004). Coh-Metrix: Analysis of text on cohesion and language. *Behavior Research Methods, Instruments, & Computers, 36*, 193-202.

- Graesser, A., Singer, M., & Trabasso, T. (1994). Constructing inferences during narrative text comprehension. *Psychological Review*, *101*, 371-395.
- Graham, S. (2006). Strategy instruction and the teaching of writing: A meta-analysis. In C. MacArthur, S. Graham, & J. Fitzgerald (Eds.), *Handbook of writing research* (pp. 187-207). New York: Guilford Press.
- Graham, S., & Hebert, M. (2011). Writing-to-read: A meta-analysis of the impact of writing and writing instruction on reading. *Harvard Educational Review*, *81*, 710–744.
- Graham, S. & Perin, D. (2007). A meta-analysis of writing instruction for adolescent students. *Journal of Educational Psychology*, *99*, 445-476.
- Halliday, M. A. K., & Hasan, R. (1976). *Cohesion in English*. London: Longman.
- Hancock, J. T., Curry, L. E., Goorha, S., & Woodworth, M. (2008). On lying and being lied to: A linguistic analysis of deception in computer-mediated communication. *Discourse Processes*, *45*, 1-23.
- Hancock, J. T., Landrigan, C., & Silver, C. (2007). Expressing emotion in text-based communication. In *Proceedings of the SIGCHI conference on human factors in computing systems* (pp. 929-932). New York, NY: ACM.
- Hannon, B. (2012). Understanding the relative contributions of lower-level word processes, higher-level processes, and working memory to reading comprehension performance in proficient adult readers. *Reading Research Quarterly*, *47*, 125-152.
- Hannon, B., & Daneman, M. (2001). A new tool for measuring and understanding the individual differences in the component processes of reading comprehension. *Journal of Educational Psychology*, *93*, 103-128.
- Hannon, B., & Daneman, M. (2006). What do tests of reading comprehension ability such as the VSAT really measure?: A componential analysis In A. V. Mittel (Ed). *Focus on Educational Psychology*. Nova Science Publications (pp. 105-146).
- Harris, K. R., Graham, S., Mason, L. H., & Friedlander, B. (2008). *Powerful writing strategies for all students*. Baltimore, MD: Brookes.
- Hayes, J. R. (1996). A new framework for understanding cognition and affect in writing. In C. M. Levy & L. S. Ransdell (Eds.), *The science of writing: Theories, methods, individual differences and applications* (pp. 1-27). Hillsdale, NJ: Erlbaum.
- Hayes, J. R., & Flower, L. S. (1980). Identifying the organization of writing processes. In L. Gregg & E. R. Steinberg (Eds.), *Cognitive processes in writing* (pp. 3-30). Hillsdale, NJ: Lawrence Erlbaum.

- Hillocks, G. (1984). What works in teaching composition: A meta-analysis of experimental treatment studies. *American Journal of Education*, 93, 133-170.
- Hillocks, G. (1986). *Research on written composition: New directions for teaching*. Urbana, IL: Eric Clearinghouse on Reading and Communication Skills.
- Huot, B. (1996). Toward a new theory of writing assessment. *College Composition and Communication*, 47, 549-566.
- Juel, C., Griffith, P., & Gough, P. (1986). A longitudinal study of the changing relationships of word recognition, spelling, reading comprehension, and writing from first to second grade. *Journal of Educational*, 78, 243-255.
- Kellogg, R. T. (1994). *The psychology of writing*. New York: Oxford University Press.
- Kellogg, R. T. (1996). A model of working memory in writing. In M. Levy & S. Ransdell (Eds.), *The science of writing: Theories, methods, individual differences, and applications* (pp. 57-72). Mahwah, NJ: Lawrence Erlbaum.
- Kellogg, R. T. (2001). Competition for working memory among writing processes. *American Journal of Psychology*, 114, 175-191.
- Kellogg, R. T. (2008). Training writing skills: A cognitive developmental perspective. *Journal of Writing Research*, 1, 1-26.
- Kellogg, R. T., Whiteford, A. P., Turner, C. E., Cahill, M., & Mertens, A. (2013). Working memory in written composition: An evaluation of the 1996 model. *Journal of Writing Research*, 5, 159-190.
- Kieft, M., Rijlaarsdam, G., Galbraith, D., & Van den Bergh, H. (2007). The effects of adapting a writing course to students writing strategies. *British Journal of Educational Psychology*, 77, 565-578.
- Kintsch, W. (1998). *Comprehension: A paradigm for cognition*. Cambridge, England: Cambridge University Press.
- Kirby, K. (1986). *Reading and writing processes of selected high-risk college freshmen*. Unpublished doctoral dissertation, University of Georgia.
- Kos, R., & Maslowski, C. (2001). Second graders' perceptions of what is important in writing. *The Elementary School Journal*, 101, 567-585.
- Kruger, J., & Dunning, D. (1999). Unskilled and unaware of it: How difficulties in recognizing one's own incompetence lead to inflated self-assessments. *Journal of Personality and Social Psychology*, 77, 1121-1134.

- Landauer, T. K., McNamara, D. S., Dennis, S., & Kintsch, W. (Eds.). (2007). *LSA: A road to meaning*. Mahwah, NJ: Erlbaum.
- Langer, J. A. (1986). *Children reading and writing: Structures and strategies*. Norwood, NJ: Ablex.
- Light, R. J. (2001). *Making the most of college: Students speaking their minds*. Cambridge: Harvard University Press.
- Lin, S., Monroe, B., & Troia, G. A. (2007). Development of writing knowledge in grades 2-8: A comparison of typically developing writers and their struggling peers. *Reading and Writing Quarterly, 23*, 207-230.
- Lipstein, R., & Renninger, K. A. (2006). "Putting things into words": 12-15 year-old students' interest for writing. In P. Boscolo & S. Hidi (Eds.), *Motivation and writing: Research and school practice*. New York: Kluwer Academic/Plenum.
- Loban, W. D. (1963). *The language of elementary school children*. Urbana, IL: National Council of Teachers of English.
- Loban, W. D. (1967). *Language ability – grades 10, 11, and 12: Final report*. Urbana, IL: National Council of Teachers of English.
- Louwerse, M. M. (2001). An analytic and cognitive parameterization of coherence relations. *Cognitive Linguistics, 12*, 291-315.
- MacGinitie, W. H., & MacGinitie, R. K. (1989). *Gates MacGinitie reading tests*. Chicago: Riverside.
- Martin, S. (1987). *The meaning-making strategies reported by provident readers and writers*. Paper presented at the meeting of the National Reading Conference, St. Petersburg, FL.
- Masson, M. E. J., & Miller, J. A. (1983). Working memory and individual differences in comprehension and memory of text. *Journal of Educational Psychology, 75*, 314-318.
- McCutchen, D. (1996). A capacity theory of writing: Working memory in composition. *Educational Psychology Review, 8*, 299-325.
- McNamara, D. S. (2004). SERT: Self-explanation reading training. *Discourse Processes, 38*, 1-30.
- McNamara, D. S. (2013). The epistemic stance between the author and the reader: A driving force in the cohesion of text and writing. *Discourse Studies, 15*, 575-592.

- McNamara, D. S., Cai, Z., & Louwrese, M. M. (2007). Optimizing LSA measures of cohesion. In T. K. Landauer, D. S. McNamara, S. Dennis, & W. Kintsch (Eds.), *Handbook of latent semantic analysis* (pp. 379-400). Mahwah, NJ: Erlbaum.
- McNamara, D. S., Crossley, S. A., & McCarthy, P. M. (2010). Linguistic features of writing quality. *Written Communication, 27*, 57-86.
- McNamara, D. S., & Graesser, A. C. (2012). Coh-Metrix: An automated tool for theoretical and applied natural language processing. In P. M. McCarthy & C. Boonthum-Denecke (Eds.), *Applied natural language processing and content analysis: Identification, investigation, and resolution* (pp. 188-205). Hershey, PA: IGI Global.
- McNamara, D. S., Jacovina, M. E., & Allen, L. K. (in press). Higher order thinking in comprehension. In P. Afflerbach (Ed.), *Handbook of Individual Differences in Reading: Text and Context*. Routledge.
- McNamara, D. S., Louwrese, M. M., McCarthy, P. M., & Graesser, A. C. (2010). Coh-Metrix: Capturing linguistic features of cohesion. *Discourse Processes, 47*, 292-330.
- McNamara, D. S., & Magliano, J.P. (2009). Self-explanation and metacognition: The dynamics of reading. In D. J. Hacker, J. Dunlosky, and A. C. Graesser (Eds.), *Handbook of metacognition in education* (pp. 60-81). New York: Routledge.
- McNamara, D. S., & Scott, J. L. (1999). Training self-explanation and reading strategies. In M. Hahn & S. C. Stoness (Eds.), *Proceedings of the Twenty-First Annual Conference of the Cognitive Science Society* (pp. 387-392). Hillsdale, NJ: Erlbaum.
- Meadows, M. & Billington, L. (2005). *Review of the literature on marking reliability*. Report for the Qualifications and Curriculum Authority. London: National Assessment Agency.
- Mol, S., Bus, A., & deJong, M. (2009). Interactive book reading in early education: A tool to stimulate print knowledge as well as oral language. *Review of Educational Research, 79*, 979-1007.
- Myers, J. L., & O'Brien, E. J. (1998). Accessing the discourse representation during reading. *Discourse Processes, 26*, 131-157.
- National Assessment of Educational Progress. (2007). *The Nation's Report Card: Writing 2007*. Retrieved Nov. 20, 2010, nces.ed.gov/nationsreportcard/writing/
- National Assessment of Educational Progress. (2009). *The Nation's Report Card: Reading 2009*. Retrieved Nov. 5, 2012, [nces.ed.gov/nationsreportcard/reading.](http://nces.ed.gov/nationsreportcard/reading/)

- National Assessment of Educational Progress. (2011). The Nation's Report Card: Writing 2011. Retrieved Nov. 5, 2012, nces.ed.gov/nationsreportcard/writing.
- The National Commission on Writing. (2004). *Writing: A ticket to work. Or a ticket out.* College Board.
- National Governors Association & Council of Chief School Officers. (2010). *Common core state standards.* Retrieved from <http://www.corestandards.org>
- Nelson, N., & Calfee, R. C. (1998). The reading-writing connection. In N. Nelson & R. C. Calfee (Eds.), *Ninety-seventh yearbook of the National Society for the Study of Education* (Part II, pp. 1-52). Chicago: National Society for the Study of Education.
- Nelson, M. M., & Schunn, C. D. (2009). The nature of feedback: How different types of peer feedback affect writing performance. *Instructional Science, 37*, 375-401.
- Oakhill, J. V., Cain, J., & Bryant, P. E. (2003). The dissociation of word reading and text comprehension: Evidence from component skills. *Language and Cognitive Processes, 18*, 443-468.
- O'Reilly, T., & McNamara, D. S. (2007). Reversing the reverse cohesion effect: Good texts can be better for strategic, high-knowledge readers. *Discourse Processes, 43*, 121-152.
- Parodi, G. (2007). Reading-writing connections: Discourse-oriented research. *Reading and Writing, 20*, 225-250.
- Pearson, P. D., Hiebert, E. H., & Kamil, M. L. (2007). Vocabulary assessment: What we know and what we need to learn. *Reading Research Quarterly, 42*, 282-296.
- Pennebaker, J. W., Booth, R. J., & Francis, M. E. (2007). Linguistic Inquiry and Word Count: LIWC [computer software]. Austin, TX.
- Pennebaker, J. W., Mayne, T. J., & Francis, M. E. (1997). Linguistic predictors of adaptive bereavement. *Journal of Personality and Social Psychology, 72*, 863-871.
- Perfetti, C. A. (1985). *Reading ability.* New York: Oxford University Press.
- Porter, L. R. (1997). *Creating virtual classroom: Distance learning with the Internet.* New York, NY: Wiley.
- Powell, P. (2009). Retention and writing instruction: Implications for access and pedagogy. *College Composition and Communication, 60*, 664-682.
- Rogers, L. A., & Graham, S. (2008). A meta-analysis of single subject design writing intervention research. *Journal of Educational Psychology, 100*, 879-906.

- Roscoe, R. D., & McNamara, D. S. (2013). Writing pal: Feasibility of an intelligent writing strategy tutor in the high school classroom. *Journal of Educational Psychology, 105*, 1010-1025.
- Roscoe, R. D., Crossley, S. A., Weston, J. L., & McNamara, D. S. (2011). Automated assessment of paragraph quality: Introductions, body, and conclusion paragraphs. In R. C. Murray & P. M. McCarthy (Eds.), *Proceedings of the 24th International Florida Artificial Intelligence Research Society (FLAIRS) Conference* (pp. 281-286). Menlo Park, CA: AAAI Press.
- Ross, J. A., Rolheiser, C., & Hogaboam-Gray, A. (1999). Effects of self-evaluation training on narrative writing. *Assessing Writing, 6*, 107-132.
- Rus, V., Lintean, M., Graesser, A. C., & McNamara, D. S. (2009). Assessing student paraphrases using lexical semantics and word weighting. In V. Dimitrova, R. Mizoguchi, B. du Boulay, & A. C. Graesser (Eds.), *Artificial intelligence in education: Building learning systems that care; From knowledge representation to affective modeling* (pp. 165-172). Amsterdam, The Netherlands: IOS Press.
- Ryan, S. M. (1985). An examination of reading and writing strategies of selected fifth grade students. In J. Niles & R. Lalik (Eds.), *Issues in literacy: A research perspective* (Thirty-fourth yearbook of the National Reading Conference, pp. 386-390), Rochester, NY: National Reading Conference.
- Saddler, B. & Graham, S. (2007). The relationship between writing knowledge and writing performance among more and less skilled writers. *Reading and Writing Quarterly, 23*, 3, 231-248.
- Schoonen, R., & de Glopper, K. D. (1996). Writing performance and knowledge about writing. In G. Rijlaarsdam, H. V. D. Bergh, & M. Couzjin (Eds.), *Theories, models and methodology in writing research* (pp. 87-107). Amsterdam: Amsterdam University Press.
- Shanahan, T. (1984). Nature of the reading-writing relation: An exploratory multivariate analysis. *Journal of Educational Psychology, 76*, 466-477.
- Shanahan, T. (1988). The reading-writing relationships: Seven instructional principles. *The Reading Teacher, 41*, 636-647.
- Shanahan, T., & Lomax, R. (1986). An analysis and comparison of theoretical models of the reading-writing relationship. *Journal of Educational Psychology, 78*, 116-123.
- Shanahan, T., & Tierney, R. J. (1990). Reading-writing relationships: Three perspectives. In J. Zutell & S. McCormick (Eds.), *Literacy theory and research: Analyses from multiple paradigms* (Thirty-ninth yearbook of the National Reading Conference, pp. 13-34). Chicago: National Reading Conference.

- Sharp, D. B. (2007). *Learn to write*. ISA Careers website. Retrieved Sept. 20, 2011, <http://www.isa.org/Template.cfm?Section=Careers&Template=/ContentManagement/ContentDisplay.cfm&ContentID=5328>
- Shell, D. F., Colvin, C., & Bruning, R. H. (1995). Self-efficacy, attribution, and outcome expectancy mechanisms in reading and writing achievement: Grade-level and achievement-level difficulties. *Journal of Educational Psychology, 87*, 368-398.
- Singer, M., & Richot, K. F. M. (1996). The role of working memory capacity and knowledge access in text inference processing. *Memory and Cognition, 24*, 733-743.
- Stotsky, S. (1983). Research on reading/writing relationships: A synthesis and suggested directions. *Language Arts, 60*, 627-642.
- Straw, S. B., & Schreiner, R. (1982). The effect of sentence manipulation on subsequent measures of reading and listening. *Reading Research Quarterly, 17*, 335-352.
- Tabachnick, B. G., & Fidell, L. S. (2001). *Using multivariate statistics, fourth edition*. Needham Heights, MA: Allyn & Bacon.
- Tierney, R. J. (1983). *Analyzing composing behavior: Planning, aligning, revising*. Paper presented at the 33rd annual National Reading Conference, Austin, TX.
- Tierney, R. J., & Shanahan, T. (1991). Research on the reading-writing relationship: Interactions, transactions, and outcomes. In R. Barr, M. L. Kamil, P. Mosenthal, & P. D. Pearson (Eds.), *The handbook of reading research* (Vol. 2, pp. 246-280). New York: Longman.
- Torrance, M., & Galbraith, D. (2006). The processing demands of writing. In C. MacArthur, S. Graham, & J. Fitzgerald (Eds.), *Handbook of writing research* (pp. 67-80). New York: The Guilford Press.
- Tousignant, M., & DesMarchais, J. E. (2002). Accuracy of student self-assessment ability compared to their own performance in a problem-based learning medical program: A correlation study. *Advances in Health Sciences Education, 7*, 19-27.
- Unsworth, N., Heitz, R. P., Schrock, J. C., & Engle, R. W. (2005). An automated version of the operation span task. *Behavior Research Methods, 37*, 498-505.
- van den Broek, P. W., & Espin, C. A. (2012). Connecting cognitive theory and assessment: Measuring individual differences in reading comprehension. *School Psychology Review, 41*, 315-325.
- Wilson, M. D. (1988). The MRC Psycholinguistic Database: Machine Readable Dictionary, Version 2. *Behavior Research Methods, Instruments, & Computers, 20*, 6-11.

- Witte, S. P., & Faigley, L. (1981). Coherence, cohesion, and writing quality. *College Composition and Communication*, 32, 189-204.
- Wong, B. (1999). Metacognition in writing. In R. Gallimore, L. P. Bernheimer, D. L. MacMillan, D. L. Speech, & S. Vaughn (Eds.), *Developmental perspectives on children with high-incidence disabilities* (pp. 183-198). Mahwah, NJ: Erlbaum.
- Yuill, N., Oakhill, J., & Parkin, A. (1989). Working memory, comprehension ability and the resolution of text anomaly. *British Journal of Psychology*, 80, 351-361.
- Zwaan, R. A., Langston, M. C., & Graesser, A. C. (1995). The construction of situation models in narrative comprehension: An event-indexing model. *Psychological Science*, 6, 292-297.