

USING APPRAISAL ANALYSIS TO MAP VALUE SYSTEMS IN HIGH-STAKES
WRITING RUBRICS

by

NICOLE ELIZABETH SIFFRINN

(Under the Direction of Ruth Harman)

ABSTRACT

With the extensive use of high-stakes tests in grades K-12 and new national standards on the rise, it is critical to consider how the language used to evaluate assessment writing could be problematic for students, teachers, and evaluators. In this study, I conduct an Appraisal analysis to examine what kind of writing is valued by the state of Pennsylvania and how changes in these values could affect the interpretation of the expectations presented in the rubrics. The goal is to compare the rubrics to see if the Keystone Exam, developed in light of the Common Core initiative, will align to the new standards. My method of analysis includes quantifying instances of Appreciation, Graduation, and Engagement in both of the rubrics and then comparing the results. My findings show a lack of clarity in expectations as to what counts as good writing in both rubrics, which suggests that the Keystone Exam will not align to the Common Core State Standards.

INDEX WORDS: appraisal theory, systemic functional linguistics, writing rubric, high-stakes writing exams; assessing writing

USING APPRAISAL ANALYSIS TO MAP VALUE SYSTEMS IN HIGH-STAKES
WRITING RUBRICS

by

NICOLE ELIZABETH SIFFRINN

BA, Bloomsburg University, 2011

A Thesis Submitted to the Graduate Faculty of The University of Georgia in Partial
Fulfillment of the Requirements for the Degree

MASTER OF ARTS

ATHENS, GEORGIA

2013

© 2013

Nicole Elizabeth Siffrinn

All Rights Reserved

USING APPRAISAL ANALYSIS TO MAP VALUE SYSTEMS IN HIGH-STAKES
WRITING RUBRICS

by

NICOLE ELIZABETH SIFFRINN

Major Professor: Ruth Harman

Committee: Ruth Harman
Linda Harklau
Christy Desmet

Electronic Version Approved:

Maureen Grasso
Dean of the Graduate School
The University of Georgia
May 2013

ACKNOWLEDGEMENTS

I would like to thank Dr. Ruth Harman for offering her support and expertise throughout this process. I have especially valued the detail of her feedback. I would also like to thank Dr. Linda Harklau and Dr. Christy Desmet for agreeing so willingly to support me in this project. And finally, I would like to thank Xiaodong Zhang for helping me in my endeavors to better understand the field of Systemic Functional Linguistics.

TABLE OF CONTENTS

	Page
ACKNOWLEDGEMENTS	iv
LIST OF TABLES	vii
LIST OF FIGURES	viii
CHAPTER	
1 INTRODUCTION	1
Motivation	1
Background	2
The Research	8
2 RATIONALE	10
3 THEORETICAL AND METHODOLOGICAL FOUNDATIONS	13
Systemic Functional Linguistics	13
Appraisal Theory	14
4 METHODS	24
Data	24
Analysis	24
5 RESULTS	26
PSSA Rubric	26
Keystone Rubric	30
Similarities and Differences between the Rubrics	34

6	DISCUSSION	40
	Appraisal in the Rubrics.....	41
	Implications for the Creation of Rubrics	51
	Conclusion	53
	REFERENCES	55
	APPENDICES	59
	A PSSA Rubric	59
	B KE Rubric	60
	C PSSA Coding	61
	D KE Coding	66

LIST OF TABLES

	Page
Table 1: Lexical Choices Associated with the Five Appreciation Types	18
Table 2: Options for Graduation	20
Table 3: Options for Engagement	22
Table 4: Number and Percentage of Appreciation in the PSSA Rubric	27
Table 5: Number and Percentage of Graduation in the PSSA Rubric	28
Table 6: Number and Percentage of Engagement in the PSSA Rubric	29
Table 7: Number and Percentage of Appreciation in the KE Rubric	31
Table 8: Number and Percentage of Graduation in the KE Rubric	32
Table 9: Number and Percentage of Engagement in the KE Rubric	33
Table 10: Similarities and Differences in Structure between the PSSA and KE Rubrics	34
Table 11: Similarities and Differences in Appraisal between the PSSA and KE Rubrics	36
Table 12: Similarities and Differences in Appraisal between the PSSA and KE Rubrics by Essay Score	39

LIST OF FIGURES

	Page
Figure 1: Expanded Appraisal Framework	16

CHAPTER 1

INTRODUCTION

Motivation

The evaluation of first and second language writing competency for K-12 students has garnered a lot more attention in recent years due to the increased use of high-stakes tests under No Child Left Behind (NCLB). NCLB, which was signed into law by the Bush administration in 2002, requires that all students in grades 3 through 8, as well as one grade in high school, be tested annually in reading and mathematics (U.S. Department of Education, 2009). The expectation was and continues to be that schools will use such assessments to determine whether students are making progress toward meeting state academic content standards. Furthermore, the results are used as a means to gauge a schools' and districts' adequate yearly progress (AYP), which is an accountability measurement under the law.

Although some researchers see the tight hold on assessment as beneficial, recent research and surveys have also documented problems with high-stakes testing. For example, a 2005 survey given by the International Reading Association revealed that despite support for the general idea of NCLB, criticisms regarding the implementation of exams and their impact on instruction under the law were plentiful ("Mixed Reactions to NCLB," 2005). Some of the main criticisms include a) testing results being inadequate measurements of school performance, b) teaching to the test negatively impacting instruction, and c) the emphasis placed on reading and math taking away from other

subjects. The implementation of this law has thus impacted curricular choices and classroom practices and in regards to writing instruction, the consequences have been seemingly detrimental.

Applebee and Langer (2011), in a study on writing instruction in middle schools and high schools, found that on average, the weight placed on the writing component of high-stakes assessments is less than that of multiple choice or short answer questions. A result of this, they claim, has been a lack of explicit writing instruction in the classroom, which is critical for understanding differences in genre (Cope & Kalantzis, 1993). Yet while findings seem to vary on how writing is being taught and how much writing is actually taking place in the classroom (e.g. Applebee and Langer, 2011; Kiuahara, Graham, & Hawken, 2009), what remains constant are the arguments about what effective writing instruction looks like and how best to assess it. Given that NCLB requires each state to administer the same standards-based assessment to every student for reporting purposes (U.S. Department of Education, 2009) and that 48 out of 50 states use writing assessments as an area of measurement (Hillocks, 2002), the need for explicitness in what is expected from student writers is imperative for not only creating fair assessments, but for fostering development in the students' writing as well. In a state like Pennsylvania, this need is even greater as statewide writing assessments have been in effect since 1990 (Lumley & Yan, 2001), with little having been done to examine their impact on instruction or the students' writing.

Background

High-stakes Assessment Writing. With the extensive use of high-stakes tests across the nation, and more specifically in Pennsylvania, assessment writing has become

its own genre. Following Martin (1996), genres are dynamic social processes that are shaped by and construe social context. This means that genres represent a specific text or type of discourse, where text refers to written language and discourse refers to oral language in use (Cameron, 2001). In other words, the content and structure of the writing exam, rubric included, can be viewed as emerging from the social processes that have influenced and continue to influence the current testing era.

What makes one genre different from another are the features that work to mark the specific social and communicative context in which the genre is produced. As Halliday and Hasan (1989) note, both structural features and register features work to shape a genre. For high-stakes writing exams, these structural features, which aid in organizing a text, are realized broadly through the use of a prompt, a set of instructions, and a rubric. The register features, which aid in implying a particular communicative context and purpose, are realized through the use of vague, prescriptive language (Hillocks, 2002). For example, as Barone and Taylor (2006) explain, the prompt usually asks students to respond to a broadly-based task and the preferred result is a rote piece of formulaic writing, typically a five-paragraph essay that shows a command of all of the performance criteria present in the accompanying rubric.

While arguments about the prompts, writing demands, and rubrics have been common (e.g. Kohn, 2006; Broad, 2003; Saddler & Andrade, 2004; Spandel, 2006; Goodrich, 2001), the design of the exams has changed little, if at all, as differences in high-stakes writing tests tend only to exist in states where portfolio-based projects are allowed (McCarthy, 2008). In the majority of states, however, Pennsylvania included,

this is not the reality as writing assessments are typically timed and done on demand without the opportunity for reflection and revision.

Pennsylvania State Assessment. In Pennsylvania, the writing portion of the state assessment has come under harsh scrutiny. Lumley and Yan (2001) note that the assessment has led to numerous court cases and that some school districts have even joined forces in an attempt to stop the exam altogether. Despite this resistance, however, the writing portion of the Pennsylvania System of School Assessment (PSSA) has remained largely unaltered since its adoption in 1989. Currently, the Writing PSSA is administered in grades 5, 8, and 11 and requires students to respond to three different prompts as well as 20 multiple-choice questions. In grades 8 and 11, the exam only includes prompts related to the persuasive and informational genres, as the expository mode of writing is thought to prepare students for college and the workforce (Data Recognition Corporation, 2011).

In terms of evaluation, a mode-specific scoring rubric with five criteria is used. The criteria are focus, content, organization, style, and conventions. Interestingly, when Lumley and Yan (2001) surveyed a group of teachers in Pennsylvania about classroom practices in regards to writing instruction, many noted that the characteristics of writing set forth by the rubric – focus, content, organization, style, and conventions – received more attention than the descriptors that accompanied them. These criteria and their descriptors will be of particular importance to this study.

While rubrics have been said to provide a valid and reliable means of evaluation, their use has been and continues to be widely debated. Despite the debates, however, the Pennsylvania Department of Education does not appear to have any plans to remove them

from statewide assessments, even with full implementation of a new set of exams in early July 2013. The question thus becomes whether or not the state has made changes to the rubric for the new exam and whether these changes, if there are any, value a similar or different kind of writing.

The New Exam. By July 1, 2013, every district in Pennsylvania is expected to implement the newly developed Common Core State Standards (CCSS) (Implementing the Common Core, 2010). The CCSS, which were adopted by Pennsylvania on July 2, 2012, define the skills and knowledge that students need in English Language Arts and Mathematics to succeed in college and the workforce. The CCSS initiative began in response to the variation in standards that existed between states (Rothman, 2011). The purpose was to create national standardization so that every student, regardless of place of residence, would have the same expectations and opportunities to prepare for success upon graduating high school. Additionally, the CCSS are internationally benchmarked, which means that they are also seen as preparing students for success in the global economy.

The CCSS are seen as very different from earlier standards, which were developed individually by states and created inconsistencies in student expectations across the nation (Rothman, 2011). For example, some states, like Pennsylvania, developed their standards by grade span instead of by grade level and some states' standards were deemed too general or vague. To address such problems, developers of the CCSS set out to create "fewer, clearer, and higher" K-12 grade-level standards in English Language Arts and Mathematics driven by current research on college and career readiness.

One important aspect of the new Standards that is strongly emphasized by the creators is that the CCSS are not a national curriculum (Implementing the Common Core, 2010). This means that while they do amount to a set of nationally shared goals for preparing students for college and the workforce, it is still currently up to each state to revise their curriculum and assessments to align to the Standards. This is why Pennsylvania has decided to move forward with their newly developed Keystone Exam (KE), which directly supports the content of the CCSS.

While the KE will not be fully implemented statewide until after the July 1, 2013 deadline, the transition process from the PSSAs has been and will continue to be large. For example, instead of testing students in just four academic subjects – reading, writing, science, and mathematics – the KE will test students in 10 different areas – Algebra I, Algebra II, Geometry, Biology, Chemistry, English Composition, Literature, Civics and Government, U.S. History, and World History – and there will be more constructed-response or open-ended questions as well as additional testing opportunities, which means that students will be expected to know more, write more, and be held more accountable for passing the exam (Implementing the Common Core, 2010). Of particular concern, is that the writing portion of the assessment will remain largely the same, as a study conducted by Lane (2010) reveals that 87% of Pennsylvania's old standards for English Language Arts, off-grades included, aligned moderately or very strongly in terms of content to the CCSS. This statistic suggests that little may have been done to alter the writing exam, especially for the upper grades, as a whopping 91% of the 11th grade standards aligned to the CCSS.

The Writing Exam and the Trajectory of English Learners. Despite relatively stagnant results for performance on the Writing PSSA – approximately 70% of students have achieved proficiency over the last three years (O’Conner, Abedi, & Tung, 2012) – there is a growing concern that English language learners are going to continue to fall behind their peers. While the state defines English language learners (ELLs) as those students who need a planned, adapted, or modified instructional program because their dominant language is not English (Pennsylvania Department of Education, 2013), I use the term in a broader sense to include any English learner (EL) whose home vernacular is different from that of the school. Hence, when I use ELL, I’m referring to the state’s definition, and when I use EL I am referring to the entire student population of English learners regardless of proficiency level.

Between 1995 and 2005, the enrollment of ELLs in public schools increased by 60.8% and in Pennsylvania alone, this resulted in 100% growth (National Center for Education Statistics, 2013). Given that Pennsylvania requires all students to take the PSSAs regardless of English language proficiency and that the KEs will follow suit with this condition (2011-2012 PSSA Handbook; Implementing the Common Core, 2010), it is becoming more and more alarming that ELL students’ performance on the Writing PSSA declined 10% between the 2005 and 2009 school years and that it is projected that this achievement gap will widen for secondary school students (O’Conner, Abedi, & Tung, 2012). Of interest to this study then, is how the guidelines in the PSSA may be affecting how ELs interpret what counts as good writing and whether or not the guidelines in the KE are better suited to address the needs of this population of developing writers.

The Research

In response to the challenges that students face with high-stakes writing, especially the EL populace, educational linguists have begun looking at Systemic Functional Linguistics to aid students in understanding and developing texts (Schleppegrell, 2004; Fang and Schleppegrell, 2006). Under SFL, language is viewed as a system of choices that people learn to use for various social, academic, and work purposes. Yet while SFL has been used as an analytical tool to examine register features in writing (e.g. Harman, in press; Kress, 1994; Macken-Horarik, 2003; Schleppegrell, 2004), children's writing development (e.g. Weaver, 1996), and ELL writing development (e.g. Shulze, 2011), little research has been done using SFL to examine the language used to assess students' writing, EL or not. Martin (1996) and Hood (2004), however, have provided insights into how the Appraisal framework, situated within SFL, can be used as a means to analyze what people value in writing.

The Appraisal framework was developed to aid in understanding the language of evaluation and stance (White, 2012). J.R. Martin (2000), who has driven a lot of the research in this area, explains that Appraisal allows for the examination of how attitudes, judgments, and emotive responses are a) either explicitly presented or indirectly implied in texts and b) either presupposed or assumed. Textual analysis of Appreciation, which is a type of attitude within the Appraisal framework used to positively or negatively assess objects or artifacts, like writing, is thus a useful tool for exploring what kind of writing is valued on high-stakes assessments.

The aim of this research is thus to conduct an Appraisal analysis of the Grade 11 persuasive writing rubrics for the PSSA and the KE. The purpose is to identify and map

the values of writing that are approved by the state. Given this task, I will ask the following questions:

- 1) What kind of writing is valued in the PSSA and KE rubrics?
- 2) How do the values change, if at all, between ratings and between the two sets of rubrics?
- 3) How do these changes affect, if at all, how the expectations in the two sets of rubrics are interpreted?

CHAPTER 2

RATIONALE

Given that rubrics have received a lot more attention in recent years due to high-stakes testing, exploring the arguments and debates about their use is foundational for understanding how and why they have become such a popular tool. The problem with this, however, as Turley and Gallagher (2008) note, is that the debates have never really moved beyond whether rubrics are “good” or “bad.” For example, without proposing or specifying any alternatives to rubrics, Mabry (1990) notes that they “operationalize” assessments by standardizing scoring, writing, and the teaching of writing, and Charney (1984) remarks that evaluations might be affected by superficial characteristics in the writing because of them. Even though these arguments, the “good” and the “bad,” are certainly necessary for identifying what is working and/or not working, the initial question of many researchers still remains: how can we un-standardize writing rubrics, draw attention away from superficial characteristics, and promote higher-order thinking, all while remaining fair and consistent? In terms of high-stakes assessments, like the PSSAs and the KEs, this question presents even more challenges since the rubrics need to be relatively short and easy to use given the thousands of essays that get graded.

While some proponents have proposed abandoning rubrics altogether because they are damaging (Wilson, 2007, 2006), it is arguably not the use of the rubric that is harmful, but rather the content of the rubric itself as the expectations that are created are not always clear to teachers, students, or evaluators. Furthermore, the type of writing that

is valued within the rubric might not be easily accessible to all members of the school population, as values tend to be conditioned by gender, ethnicity, generation, and class (Martin, 1996; Macken-Horarik, 2003). In other words, texts are productive of a great number of meanings and reading positions may shift depending on the person and his/her cultural and social experiences.

Despite the potential for multiple readings, however, the only requirements for the implementation of rubrics is that they must be reliable and valid. This means that there must be consistency in scoring and that the assessment measures what it is supposed to (Mabry, 1990). Validity can, however, be undermined for two reasons. First, writing rubrics typically have three to six different levels of performance, or criteria, as will be seen on both the PSSA and KE rubrics. While fewer choices increases the chances of consistency between scorers due to less room for disagreement, it also suggests that not all of the features of the students' writing are being taken into consideration; thus, there is a discrepancy between the score and the students' performance (Mabry, 1990).

Furthermore, Beck (2006) claims that evaluators must rely on their own knowledge to interpret what is meant by the criteria provided in rubrics and Cooper (1999) notes that while the most common criteria for evaluating persuasive writing relate to the thesis, structure, cohesion, and evidence within the essay, rubrics often fail to elaborate on what is needed to successfully meet the requirements of those domains. It is obvious then, that if teachers and evaluators have difficulty interpreting the specifics of the criteria, or if they have a different reading of the values set forth by the rubric, students, and especially those considered ELs, inevitably will too.

These challenges tend to be even more marked for first-generation ELs who have started school at the secondary level, where language demands increase as grade levels rise (National Center for Education Statistics, 2013). Nationwide, these heightened language demands have resulted in an increase in the achievement gap in English Language Arts between ELLs and non-ELLs. And in Pennsylvania, this trend is not any different, as the achievement gap in writing specifically had widened 41.1 percentage points in grade 11 by 2009 (O’Conner, Abedi, and Tung, 2012).

Using Appraisal as a means to analyze these rubrics will thus help to reveal how clear or unclear the requirements for completing the writing prompts are and offer, perhaps, a foundation for future explorations of why this achievement gap has continued to widen. Even though the structure and design of the rubrics could be problematic too in terms of difficulties understanding the layout of the document, since evaluators gauge students’ success on language-based criterion, it is critical that the focus first be on the content of the rubrics and how understandable the expectations actually are. I will, however, briefly comment on the design of the rubrics when I compare the PSSA and the KE in the findings section.

CHAPTER 3

THEORETICAL AND METHODOLOGICAL FOUNDATIONS

Systemic Functional Linguistics

SFL, which provides a framework for understanding how language is used to create and exchange meaning (Young & Fitzgerald, 2006), is a useful device for examining structural items, attitudes, and cohesion in texts. SFL differs from Chomskyan notions of transformational grammar in that it focuses on language in context as opposed to language as abstraction. It treats language as both a system and a resource for making meaning in social and cultural situations (Gibbons, 2006). Following the work of M.A.K. Halliday and his colleagues, SFL is grounded in social semiotics, whereby language users make meaning via linguistic choices, not rules.

For Halliday, language use is influenced by, as much as it influences, the context of the situation, or register, which is thought of as containing three general metafunctions: one for construing experiences, one for enacting social relations, and one for combining the previous two to create a text. In more technical terms, these processes have come to be known as the ideational (field of the discourse), interpersonal (tenor of discourse), and textual (mode of the discourse) metafunctions as they allow different types of meaning to be made through the use of different lexicogrammatical resources. As Gibbons explains:

Given a certain field, choices are made from within the resources of the experiential function of language; given a certain tenor, choices are made

from within the interpersonal resources of the language system, and given a certain mode of communication, choices are made from within the textual resources of the language system (Gibbons, 2006, p. 31).

In other words, there is a two-way relationship between the categories of the register and the structure of language and in order for participants to make sense of one another and the world around them, each part of this meaning-making system is needed.

Appraisal Theory

Overview. Because the focus in this study is the interpersonal metafunction and the Appraisal framework that is situated within it, there will be an analysis of how evaluative language, attitude, and emotion are used to create interpersonal proposals and propositions (White, 2012). There are three subtypes of Appraisal – Attitude, Engagement, and Graduation – and each type functions as a semantic resource for achieving interpersonal means.

Attitude, which refers to the positive or negative assessment of a person, place, thing, or affair, involves three semantic regions that construe emotion, ethics, or aesthetics (Martin & White, 2005). For example, I can claim that I am happy with a piece of writing, that the writer is skillful, or that a piece of writing is strong. The differences reside in what is being appraised – my own reaction to the writing, the writer's behavior, or the writing itself.

Graduation draws attention to meanings on the basis of whether the Force of an utterance is Raised or Lowered and/or whether the Focus of a categorization is Sharpened or Softened (White, 2012). For example, “*precise* evidence” indicates Raising while “*imprecise* evidence” indicates Lowering – note the positive and negative connotations at

play – and “a true piece of writing” indicates Sharpening while “ it is *sort of* a good piece of writing” indicates Softening – note the illumination and the blurring. These resources are thus concerned with scalability in regards to the appraised item. That is, Graduation allows us to comment on how strongly we feel about something or someone, like a piece of writing or the writer.

Engagement, on the other hand, is indicative of how participants negotiate and adjust their utterances, or rather how the authorial voice positions itself in regards to creating or denying a space for other voices and alternative positions (Martin & White, 2005). For example, “the writer demonstrates control of language” works to close down a space for alternative positions, as the only option is to agree that the writer has demonstrated control of language. “The writer *may* demonstrate control of language,” on the other hand, works to open up a space for alternative positions, as “may” enables various opinions to be made about whether or not the writer demonstrated control of language. Engagement is thus a resource for adopting a stance toward a certain value position.

All three of these subtypes – Appreciation, Graduation, and Engagement – work together to create evaluative language and help to reveal the author’s relationship with the reader and/or material being appraised. In other words, these resources enable a text, or rather the authorial voice in a text, to adopt stances towards the subject matter they present and those with whom they communicate. In the case of the PSSA and the KE, then, the institutional voice within the rubrics approves or disapproves of a certain type and quality of writing and these same stances and values are to be taken up or shared with those who use the rubrics.

The Appraisal system can, however, be extended beyond Attitude, Graduation, and Engagement, as each category is able to expand into more detailed subcategories to aid in identifying the specifics of interpersonal meaning in a given text. Figure 1 shows the subcategories of Attitude, Graduation, and Engagement.

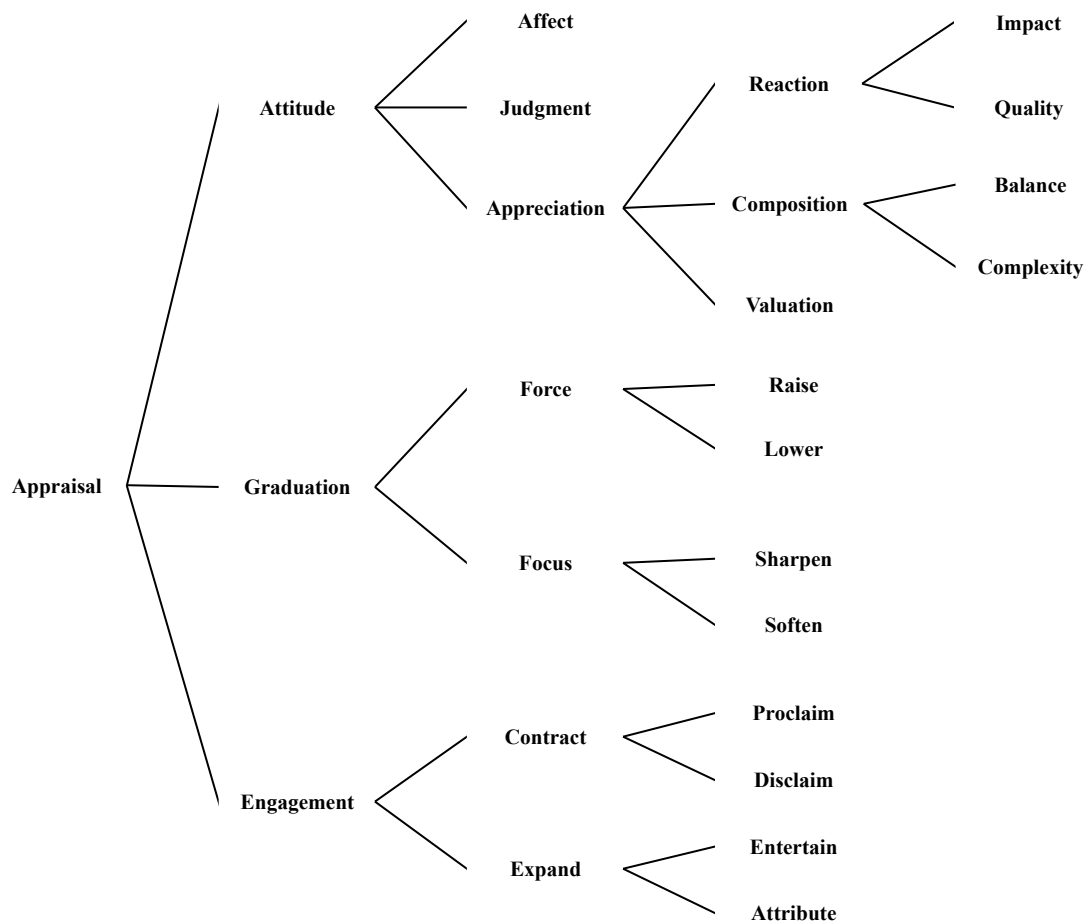


Figure 1 – *Expanded Appraisal Framework* (modified from Martin & Rose, 2003)

As can be seen in Figure 1, I have expanded every area of the framework except for Affect and Judgment, which are subcategories of Attitude. The reason for this is because unlike Appreciation, which involves *the evaluation of objects and products with*

respect to systems of social value, Affect and Judgment deal with emotions and behaviors. Appreciation thus works well in terms of analyzing the evaluative language present in the rubrics, as the appraisal is in reference to the writing, not the writer or his/her behavior. Graduation and Engagement, however, have both been extended to two additional sublevels because they provide additional information about the scalability of the values set forth by the rubric and how opened up or closed down these values are to multiple interpretations.

Appreciation. Appreciation, as was mentioned before, construes the evaluation of “things.” It is useful for analyzing the evaluative language in the rubrics as it shifts the personal realization of emotions and feelings to an institutional framework (Martin & White, 2005). This means that emotion is recontextualized so as to evade appearing as an affectual response or judgment of someone’s behavior. For example, the authorial voice in the rubrics is able to avoid any type of personal orientation towards the criteria by appraising the object of value, in this case the writing, in the institutional realm, which contains certain attitudes that one is expected to have about the type and quality of writing that is valued.

In general, Appreciation can be divided into three categories: Reaction, Composition, and Valuation (Martin & White, 2005). Reaction and Composition can further be extended to the subtypes Impact, Quality, Balance, and Complexity. Reactions that encode Impact deal with whether or not a product grabs the appraiser’s attention and Reactions that encode Quality deal with whether or not the product pleases the appraiser (see Table 1 below for concrete examples). Composition, on the other hand, encodes the evaluation of the textual components of the product. Hence, Balance encodes values of

how the product hangs together and Complexity encodes values of how hard the product is to follow. Valuation, the third subcategory, encodes values for how worthwhile the product is. Table 1, adapted from Martin and White (2005), shows examples of what kind of lexical choices occur at each of these levels. As can be seen, there are positive and negative lexical choices associated with each type of Appreciation and this continuum will aid in mapping how the values in the rubrics change, if at all, as the rating scales change.

Table 1 <i>Lexical Choices Associated with the Five Appreciation Types</i>		
Appreciation Type	Positive Lexical Choices	Negative Lexical Choices
Reaction Impact	arresting, captivating, engaging...; fascinating, exciting, moving...; lively, dramatic intense...	dull, boring, tedious ...; dry, ascetic, uninviting ...; flat, predictable, monotonous ...
Reaction Quality	okay, fine good ...; lovely, beautiful, splendid ...; appealing, enchanting, welcome ...	bad, yuk, nasty; plain, ugly, grotesque; repulsive, revolting, off-putting ...
Composition Balance	balanced, harmonious, unified ...; symmetrical, proportioned; consistent, considered, logical ...	unbalanced, discordant ...; uneven, flawed ...; contradictory, disorganized ...
Composition Complexity	simple, pure, elegant; lucid, clear, precise ...; intricate, rich, detailed ...	ornate, extravagant, byzantine ...; arcane, unclear, woolly ...; plain, monolithic, simplistic ...
Valuation Worth	penetrating, profound, deep ...; innovative, original, creative ...; timely, long awaited, landmark ...	shallow, reductive, insignificant ...; derivative, conventional ...; dated, overdue, untimely ...

The Appreciation framework also works well for this analysis because it can be interpreted metafunctionally, linking itself with the broader system of SFL (Martin & White, 2005). Reaction, which again realizes emotive language, can be oriented towards the interpersonal metafunction. In other words, analysis can show the evaluative stance toward the subject and/or reader as reaction is related to affection given a similar lexis. Composition, which focuses on the arrangement of a text, can be oriented towards the textual metafunction in that organization and order are intimately connected to theme and

cohesion. And Valuation, which deals with experiential worth, can be oriented towards the experiential metafunction, as the value of objects and artifacts is largely dependent on the institutions in which they develop from. Mapping the Appreciation types within the rubrics onto the three metafunctions thus provides more evidence as to why shifts in the type of evaluative language in the KE from the earlier rubric may cause confusion, especially given that no concrete examples are given as to what is expected.

An appraised item, for instance, for the highest score may realize a different value of Appreciation for the lowest score. Therefore, what might be considered a realization of Reaction at one level could be a realization of Composition at another, which would mean that the evaluation shifts from the interpersonal to the textual metafunction, or from the tenor of the discourse to the mode of the discourse. This shift can be seen in the PSSA rubric with “content” being appraised as “substantial” for the highest score and “adequate” for the second highest score. The term “substantial,” which deals with the amount of content, can be categorized as a realization of Composition and the term “adequate,” which deals with quality, can be categorized as a realization of Reaction. The problem with this shift is that at one level the writing is being evaluated for how much content is present and at another it is being evaluated for the quality of the content. Additionally, since there are no specifications as to what constitutes “substantial content” or “adequate content,” the rubric user must rely on his/her own disciplinary knowledge to interpret the meaning of these criteria, which creates a greater possibility that the essay score will be affected by the fact that the writer and the evaluator could very well have different readings of the expectations. This link with the metafunctions is thus critical to

understanding why a shift in Appreciation of the same item can be so problematic when trying to interpret the expectations set forth by a rubric.

Graduation. In addition to Appreciation, Graduation is an appraisal resource that can aid in determining how clear or unclear the criteria in a rubric are, as it is concerned with scalability and/or the reconstrual of categorical meanings. Graduation can thus be classified as having Force or Focus or both (Martin & Rose, 2003). While meanings dealing with Force pertain to degree of intensity or amount, meanings dealing with Focus pertain to making non-gradable items gradable. Each subtype can be expanded further to help flesh out how Force and Focus are realized (see Table 2 below).

Table 2 <i>Options for Graduation</i>			
Force		Focus	
Raise	Lower	Sharpen	Soften
A <u>great</u> writer...	A <u>lousy</u> writer...	A <u>true</u> writer...	<u>Not quite</u> a writer
An <u>excellent</u> piece of writing	A <u>worthless</u> piece of writing	<u>Exact</u> evidence	<u>Near exact</u> evidence
<u>Arresting</u> argument	<u>Dull</u> argument	<u>Very</u> authentic argument	<u>Somewhat</u> authentic argument

As can be seen in Table 2, both Force and Focus are used as resources to grade items and set them apart from one another. Force is used to show how strongly someone feels about something (Martin & White, 2005) and in this research study, the something is the writing. The strength of these evaluative feelings, which again, for the purposes of this analysis, are institutionalized, can be thought of as having volume. Thus, the appraiser can either Raise the volume or Lower it. For example, in the PSSA rubric, one of the criterion awards points for a “Sharp controlling point” and one of the criterion awards points for a “Clear controlling point.” In both instances, the pre-modification

serves to amplify the type of controlling point that is present and in both instances, the amplification serves to **raise** the volume as there is a positive connotation embedded in the lexical choices. Additionally, if these items were to be scaled, “sharp” would be classified as having a higher grading than “clear.”

Focus, on the other hand, is used to make something that is non-gradable gradable and there are two resources for doing that: Sharpening and Softening (Martin & White, 2005). This means that Focus is used to make categorical distinctions, drawing boundaries between things. When a category is **sharpened**, the specification is maximized, or up-scaled. When a category is **softened**, the specification is blurred, or downscaled. For example, in the PSSA rubric, one of the criterion awards points for “some evidence” and one of the criterion awards points for “no evidence.” The pre-modification allows a boundary to be marked between how much evidence is or is not present. “Some” thus serves to blur the boundary as no specifics are made clear and “no” serves to sharpen the boundary as it is very direct and provides an either/or option.

Engagement. The Engagement system deals primarily with authorial positioning. It provides the means for a space to be opened up or closed down with regard to other voices or alternative positions. This opening and closing is referred to as Expansion and Contraction and there is a taxonomy (see Table 3 below) for each of these subclasses (Martin & White, 2005).

Under Expansion, the authorial voice distances itself from the given proposition by inviting or entertaining a different position or multiple positions (Martin & White, 2005). In other words, the authorial voice opens up the dialogical space, which in turn lowers the interpersonal stakes for anyone who would advance a different standpoint.

Table 3 <i>Options for Engagement</i>			
Expand		Contract	
Entertain	Attribute	Proclaim	Disclaim
<p>The writer <u>may</u> shift point of view...</p> <p>Rhetorical strategies <u>may</u> be evident...</p>	<p>The writer <u>acknowledges</u> possible counterclaims...</p>	<p>The writer <u>provides</u> relevant content...</p> <p>Clear controlling point <u>presented</u> as a position...</p>	<p>The writer <u>does not</u> need additional evidence to support claims...</p>

There are two types of resources for realizing Engagement: Entertain and Attribute. When the authorial voice **entertains**, it explicitly characterizes the proposition as being only one of a range of possible stances, grounded in its own subjectivity. For example, in the KE rubric, one of the criterion reads, “The writer may use simplistic transitions ...” The use of *may* makes this proposition dialogically expansive as it entertains the possibility of the writer using simplistic transitions. When the authorial voice **attributes**, on the other hand, it again characterizes the proposition as being only one of a range of possible stances, but this time does so by bringing in an external voice. For example, one of the criteria reads, “The writer acknowledges possible counterclaims ...” The use of “acknowledge” makes this proposition dialogically expansive as it attributes the recognition of possible counterclaims to the writer. In other words, the use of a reporting verb makes it impossible to know where the authorial voice stands on this matter. Both Entertain and Attribute, however, consider or invoke dialogical alternatives.

Under Contraction, the authorial voice sets itself against any alternative positions by adopting a particular stance towards the given proposition (Martin & White, 2005). In other words, the authorial voice closes down the dialogical space, which in turn raises the

interpersonal stakes as advancing a different standpoint is not a viable option. There are two types of resources for realizing Contraction: Proclaim and Disclaim. When the authorial voice proclaims, it characterizes the proposition as being highly deserving or justifiable. This means that the authorial voice estranges itself from, represses, or discounts any alternative positions. For example, in the KE rubric, one of the criterion reads, “The writer provides relevant content ...” The use of *provides* allows the authorial voice to endorse the proposition, thereby ruling out any alternatives. When the authorial voice disclaims, on the other hand, it characterizes the proposition as being in conflict with or in rejection of some opposing position. For example, one of the criteria might read, “The writer does not need additional evidence to support claims...” The use of *does not* allows the authorial voice to deny any converse position. Both Proclaim and Disclaim, however, close down any dialogical alternatives.

Given this discussion on Appreciation, Graduation, and Engagement and how each work as a resource to create interpersonal meaning, I will discuss results of the analyses and comparison of the two rubrics in Chapter 5 and then provide rich discussion, informed by recent literature, on these findings in Chapter 6.

CHAPTER 4

METHODS

Data

The following data was collected as part of this study: the PSSA persuasive writing rubric for grade 11 and the KE persuasive writing rubric for grade 11 (see Appendix A and B for these items). All of these items are available for public access on the Pennsylvania Department of Education's website: www.pde.state.pa.us.

I selected Pennsylvania for the site of inquiry for several reasons. First, I have experienced the Pennsylvania State Standardized Assessment as both a student and a student teacher, and while I am very familiar with the scoring guidelines and how the test is administered, I have struggled to understand what the scoring criteria actually mean and have, in turn, as Ball (2011) claims of many teachers, relied on my tacit knowledge of how to write within the persuasive genre to navigate such texts. Second, since full implementation of the Common Core in the state is expected by July, I think it is imperative to examine how the values within the old rubric differ from the new rubric, if at all. Third, since a national test is currently under development and Pennsylvania is playing a role in its creation, knowing precisely what is or is not being valued in writing could shed light on future writing assessments in the state and nationwide.

Analysis

Informed by the SFL appraisal theory and empirical research articulated in the previous section, I began my examination of the rubrics by coding first for Appreciation,

then for Graduation, and then for Engagement. I placed the results into tables based on domain and essay score (see Appendix C and D). This allowed me to track and view how the values changed, if at all, between domains and essay scores. When all of the domains and essay scores were coded, I compiled the results based on Appraisal type and essay score. This allowed me to see what patterns emerged as the Appraisal type or scores changed. As will be seen in the results section, I documented the number and percentage of each of the three expanded Appraisal subsystems (Appreciation, Graduation, and Engagement) to aid in comparisons within the rubrics and between them.

CHAPTER 5

RESULTS

I will first discuss findings about the PSSA rubric and then move into findings about the KE rubric. Both of these discussions will be centered on the Appreciation, Graduation, and Engagement present and/or not present in each of the rubrics. Lastly, I will compare and contrast the findings in the PSSA and the KE.

PSSA Rubric

Overall. The following section will detail the number and percentage of Appreciation, Graduation, and Engagement in the PSSA Rubric. Each subsection includes a table that is organized by the essay score and the Appraisal category being examined. The essay score runs horizontally and the Appraisal category runs vertically. This allows one to see how, if at all, the instances of Appraisal differ between scores.

Appreciation. Table 4 shows the number and percentage of Appreciation in the PSSA Rubric. It is categorized by essay score, whereby 4 is the highest and 1 is the lowest, and Appreciation type, which includes Reaction, Composition, and Valuation. Reaction is further broken into Impact and Quality and Composition is further broken into Balance and Complexity. There is also a category for Ambiguous instances of Appreciation, whereby the Appraisal choice could not be clearly identified as Reaction, Composition, or Valuation and was thus classified as potentially having characteristics of two Appreciation types.

The table below shows that there are 77 instances of Appreciation in the PSSA Rubric. These instances account for 21.6% of all of the words. Of these 77 instances, more of them occur for the highest possible score than for all of the lower scores – 31.2% compared to 23.4% and 22.1%, respectively.

			Score 4		Score 3		Score 2		Score 1		Averages	
			#	%	#	%	#	%	#	%	#	%
Appreciation Type	Reaction	Impact	1	1.3	1	1.3	0	0	0	0	2	2.6
		Quality	0	0	3	3.9	4	5.2	1	1.3	8	10.4
	Composition	Balance	6	7.8	3	3.9	7	9.1	13	16.9	29	37.7
		Complexity	11	14.3	8	10.4	4	5.2	2	2.6	25	32.5
	Valuation		3	3.9	1	1.3	0	0	0	0	4	5.2
	Ambiguous		3	3.9	2	2.6	2	2.6	2	2.6	9	11.7
	Totals		24	31.2	18	23.4	17	22.1	18	23.4	77	21.6

The table also shows that Composition is the type of Appreciation that is used the most in the rubric. Of all the instances of Appreciation, it occurs 70.2% of the time. 37.7% of the Composition deals with Balance and 32.5% deals with Complexity. Furthermore, it can be seen that the instances of Composition Balance are greater for the lowest possible score in the rubric and that those of Composition Complexity are greater for the highest possible score.

There are only four instances of Valuation, three of which occur in Score 4 and one of which occurs in Score 3. Similarly, there are only two instances of Reaction Impact, one of which occurs in Score 4 and one of which occurs in Score 3. Reaction Quality, on the other hand, makes eight appearances throughout the rubric, all but one of which are found in Score 3 and Score 2.

Graduation. The following table shows the number and percentage of Graduation in the PSSA Rubric. It is categorized by essay score and Graduation type, which includes Force and Focus. Force is broken into Raising and Lowering and Focus is broken into Sharpening and Softening. Totals are present for each of the categories and subcategories.

			Score 4		Score 3		Score 2		Score 1		Averages	
			#	%	#	%	#	%	#	%	#	%
Graduation	Force	Raise	24	31.2	12	15.6	4	5.2	3	3.9	43	55.8
		Lower	0	0	6	7.8	13	16.9	15	19.5	34	44.2
	Totals		24	31.2	18	23.4	17	22.1	18	23.4	77	21.6
	Focus	Sharpen	22	30.6	15	20.8	12	16.7	12	16.7	61	84.7
		Soften	0	0	2	2.8	3	4.2	6	8.3	11	15.3
	Totals		22	30.6	17	23.6	15	20.8	18	25	72	20.2

Table 5 shows that there are 77 instances of Force, 43 of which entail Raising and 34 of which entail Lowering. The percentage of Raising decreases as the score decreases – 31.2%, 15.6%, 5.2%, and 3.9% – and the percentage of Lowering increases as the score increases – 0%, 7.8%, 16.9%, and 19.5. There are twice as many instances of Raising for score 4 than there are for Score 3, and there are three times as many instances of raising for Score 3 than there are for Score 2. There are not any instances of Lowering for Score 4, and of the 44.2% of Lowering that occurs throughout the entire rubric, 36.4% of it occurs in Score 2 and Score 1.

The table also shows that there are 72 instances of Focus, 61 of which entail Sharpening and 11 of which entail Softening. The difference between the two is 68.4%. The percentage of Sharpening decreases as the score decreases, all except for the

percentage of Sharpening for Score 1, as it remains the same as the percentage of Sharpening for Score 2 – 16.7%. The percentage of Softening, on the other hand, increases as the score increases – 0%, 2.8%, 4.2%, and 8.3%. There are not any instances of Softening for Score 4, and of the 15.3% of Softening that occurs throughout the entire rubric, 8.3% of the instances occur for Score 1.

Engagement. Table 6 shows the number and percentage of Engagement in the PSSA Rubric. It is categorized by essay score and Engagement type, which includes Contraction and Expansion. Contraction is broken into the subcategories Proclaim and Disclaim, and Expansion is broken into the subcategories Entertain and Attribute. Totals are present for each of the categories and subcategories.

Table 6 Number and Percentage of Engagement in the PSSA Rubric												
			Score 4		Score 3		Score 2		Score 1		Averages	
			#	%	#	%	#	%	#	%	#	%
Engagement	Contract	Proclaim	9	27.3	7	21.2	7	21.2	7	21.2	30	90.9
		Disclaim	0	0	0	0	0	0	3	9.1	3	9.1
	Totals		9	27.3	7	21.2	7	21.2	10	30.3	33	9.3
	Expand	Entertain	3	20	3	20	3	20	6	40	15	100
		Attribute	0	0	0	0	0	0	0	0	0	0
	Totals		3	20	3	20	3	20	6	40	15	4.2

The table above shows that there are 33 instances of Contraction. Of the 33 instances of Contraction, 90.9% of them can be categorized as Proclaiming. The only instances of Disclaiming, of which there are three, occur in Score 1. The amount of Proclaiming is relatively steady across scores, with the only difference being for Score 4, where there are two more instances of Proclaiming – 9 compared to 7. Table 4 also shows that there are 15 instances of Expansion, all of which can be classified as

Entertainment. Nine of these instances occur for Score 4, 3, and 2, which amounts to three instances per score. The remaining six instances occur for Score 1. These six instances account for 40% of the Entertainment. Interestingly, a little over twice as much Contraction is present than Expansion – 33 instances compared to 15.

Summary. As was seen, the amount of Appreciation, Graduation, and Engagement varied by essay score. While some of the distributions, like for Raising and Lowering, were stable, others, like for Reaction Impact, were scattered. While I will discuss the significance of these findings in Chapter 6, the following section will report on the findings from the KE rubric.

Keystone Rubric

Overall. The following section will detail the number and percentage of Appreciation, Graduation, and Engagement in the Keystone Rubric. Each subsection includes a table that is organized by the essay score and the Appraisal category being examined. The essay score runs horizontally and the Appraisal category runs vertically. This allows one to see how, if at all, the instances of Appraisal differ between scores.

Appreciation. The table below shows the number and percentage of Appreciation in the Keystone Rubric. It is categorized by essay score, whereby 4 is the highest and 0 is the lowest, and Appreciation type, which includes Reaction, Composition, and Valuation. Reaction is further broken into Impact and Quality and Composition is further broken into Balance and Complexity. There is also a category for Ambiguous instances of Appreciation, whereby the Appraisal choice could not be clearly identified as Reaction, Composition, or Valuation and was thus classified as potentially having characteristics of two Appreciation types.

			Score 4		Score 3		Score 2		Score 1		Score 0		Averages	
			#	%	#	%	#	%	#	%	#	%	#	%
Appreciation Type	Reaction	Impact	0	0	0	0	0	0	0	0	0	0	0	0
		Quality	0	0	0	0	2	3.2	0	0	0	0	2	3.2
	Composition	Balance	5	8.1	2	3.2	6	9.7	9	14.5	9	14.5	31	50
		Complexity	7	11.3	3	4.8	3	4.8	3	4.8	1	1.6	17	27.4
	Valuation		5	8.1	3	4.8	2	3.2	1	1.6	0	0	11	17.7
	Ambiguous		1	1.6	0	0	0	0	0	0	0	0	1	1.6
Totals			18	29.0	8	12.9	13	21	13	21	10	16.1	62	13.7

Table 7 shows that there are 62 instances of Appreciation in the Keystone Rubric. Of these 62 instances, 77.4% of them deal with Composition, 50% of which is classified as Composition Balance and 27.4% of which is classified as Composition Complexity. The instances of Composition Complexity decrease as the essay scores decrease and the instances of Composition Balance increase as the essay scores increase, with the exception of Score 3.

The table also shows that Valuation comprises the third largest amount of Appreciation, totaling 17.7%. The instances of Valuation decrease as the essay scores decrease. In regards to the remaining categories of Appreciation, there are only two instances of Reaction, both of which deal with Quality, and there is only one Ambiguous case, which can be seen occurring in Score 4. There are more than twice as many instances of Appreciation for Score 4 than there are for all of the lower scores.

Graduation. The following table shows the number and percentage of Graduation in the Keystone Rubric. It is categorized by essay score and Graduation type, which includes Force and Focus. Force is broken into Raising and Lowering and Focus

is broken into Sharpening and Softening. Totals are present for each of the categories and subcategories.

Table 8 Number and Percentage of Graduation in the Keystone Rubric														
			Score 4		Score 3		Score 2		Score 1		Score 0		Averages	
			#	%	#	%	#	%	#	%	#	%	#	%
Graduation	Force	Raise	18	29	8	12.9	1	1.6	1	1.6	0	0	28	45.2
		Lower	0	0	0	0	12	19.4	12	19.4	10	16.1	34	54.8
	Totals		18	29	8	12.9	13	21	13	21	10	16.1	62	13.7
	Focus	Sharpen	16	25.8	7	11.3	8	12.9	8	12.9	9	14.5	48	77.4
		Soften	2	3.2	1	1.6	5	8.1	5	8.1	1	1.6	14	22.6
	Totals		18	29	8	12.9	13	21	13	21	10	16.1	62	13.7

Table 8 shows that there are 62 instances of Force and Focus. Within Force, there are more instances of Lowering than there are Raising and within Focus, there are more instances of Sharpening than there are Softening. As can be seen, most of the Raising within Force occurs for Score 4 and Score 3 – 41.9% – and all of the Lowering occurs in Scores 2, 1, and 0. There is not any Raising in Score 0 and there is not any lowering in Score 4 or 3.

The table also shows that there is 54.8% more Sharpening than there is Softening. Score 4 contains the most Sharpening with 25.8% and the remaining scores contain relatively stagnant instances of Sharpening, with 7, 8, and 9 instances respectively. This means that there are nearly twice as many instances of Sharpening in Score 4 than there are in any of the other scores. In terms of Softening, Score 2 and 1 contain the most instances at 16.2% combined. The remaining three scores combine to account for 6.4% of the Softening

Engagement. Table 9 shows the number and percentage of Engagement in the Keystone Rubric. It is categorized by essay score and Engagement type, which includes Contraction and Expansion. Contraction is broken into the subcategories Proclaim and Disclaim, and Expansion is broken into the subcategories Entertain and Attribute. Totals are present for each of the categories and subcategories.

Table 9 shows that there are 55 instances of Contraction and 23 instances of Expansion. Of the 55 instances of Contraction, 70.9% of them deal with Proclaiming, and of the 23 instances of Expansion, 87% of them deal with Entertaining. The only instances of Disclaiming occur in the lowest two scores, with 12 out of the 16 present in Score 0. The largest number of Proclaiming occurs in Score 4 accounting for 23.6% of it.

			Score 4		Score 3		Score 2		Score 1		Score 0		Averages	
			#	%	#	%	#	%	#	%	#	%	#	%
Engagement	Contract	Proclaim	13	23.6	8	14.5	6	10.9	8	14.5	4	7.3	39	70.9
		Disclaim	0	0	0	0	0	0	4	7.3	12	21.8	16	29.1
	Totals		13	23.6	8	14.5	6	10.9	12	21.8	16	29.1	55	12.1
	Expand	Entertain	1	4.3	1	4.3	9	39.1	8	34.8	1	4.3	20	87
		Attribute	1	4.3	1	4.3	1	4.3	0	0	0	0	3	13
	Totals		2	8.7	2	8.7	10	43.5	8	34.8	1	4.3	23	5.1

Table 9 also shows that for Expansion, there are 74% more instances of Entertainment than there are Attribution. Entertainment is the highest for Score 2 at 39.1% and Score 1 at 34.8%. Attribution only occurs three times, once in Scores 4, 3, and 2. As a whole, Expansion is the highest for Score 2 at 43.5% and Score 1 at 34.8%, and Contraction is the highest for Score 0 at 29.1%.

Summary. As was seen, the amount of Appreciation, Graduation, and Engagement varied by essay score just like it did in the PSSA. While some of the distributions, like for Valuation, were stable, others, like for Reaction, were scattered. While I will discuss the significance of these findings in Chapter 6, the following section will compare the results of the PSSA and KE rubrics.

Similarities and Differences Between the Rubrics

Overall. The following section will detail the similarities and differences between the PSSA and KE rubrics. The first section will focus briefly on the structure of the rubrics and the second section will focus on the instances of Appraisal in the rubrics.

Structure. The table below shows the similarities and differences in structure between the PSSA and KE rubrics. The table highlights specifics pertaining to the domains, scoring, and organization of each of the rubrics.

Table 10 <i>Similarities and Differences in Structure Between the PSSA and KE Rubrics</i>		
	PSSA	KE
Domains	Focus Content Development Organization Style	Thesis/Focus Content Organization Style
Scoring	Scores range from 4 to 1	Scores range from 4 to 0
Organization	Scores run vertically from 4 to 1 with the domains appearing to the right of the score and the requirements appearing to the right of the domain	Requirements are in bullet point format in a table that features the scores running horizontally and the domains running vertically

As Table 10 shows, there are four domains in the PSSA and the KE rubrics and these domains have a similar wording. The only difference is that for the KE rubric, “Thesis” is added to the Focus domain and “Development” is eliminated from the

Content Domain. The scoring, however, while still out of 4 for each of the rubrics, does differ in that the lowest score one can receive on the PSSA is 1 while the lowest score one can receive on the KE is 0.

The table also shows that the organization of each of the rubrics differs significantly (refer to Appendix A and B to view the rubrics). In the PSSA rubric, the scores run vertically with the domains and their requirements appearing in each of the boxes pertaining to the score. Thus, the requirements for the same domain are not side-by-side and do not fall directly on top of one another. Instead, to compare the requirements for the same domains, one must sift through or try to ignore the requirements of all of the other domains. The KE rubric, on the other hand, is set up as a table with the essay score appearing horizontally across the top of the table and the domains appearing vertically along the left-hand side of the table. This setup enables the requirements for the same domain to appear side-by-side, which allows the rubric user to read from left to right what is expected at each score within the same domain.

Additionally, the requirements in the PSSA rubric are not laid out using complete sentences while the requirements in the KE rubric are. Underneath each of the scores on the KE rubric, the words “At this score point the writer –” are written and then the requirements are listed in bullet point format in each of the domains to complete the sentence. There are no bullet points within the domains in the PSSA rubric.

Appraisal overall. The table below shows the similarities and differences in Appraisal between the PSSA and KE rubrics. The table highlights specifics pertaining to the instances of Appreciation, Graduation, and Engagement in the rubrics.

Table 11 <i>Similarities and Differences in Appraisal Between the PSSA and KE Rubrics</i>		
Appraisal Type	PSSA	KE
Appreciation by Percentage	77 Instances (21.6% of the total words)	62 Instances (13.7% of the total words)
	Reaction Impact: 2.6% Reaction Quality: 10.4% Composition Balance: 37.7% Composition Complexity: 32.5% Valuation: 5.2% Ambiguous: 11.7%	Reaction Impact: 0.0% Reaction Quality: 3.2% Composition Balance: 50.0% Composition Complexity: 27.4% Valuation: 17.7% Ambiguous: 1.6%
Graduation	77 Instances of Force	62 Instances of Force
	Raising: 55.8% Lowering: 44.2%	Raising: 45.2% Lowering: 54.8%
Engagement	77 Instances of Focus	62 Instances of Focus
	Sharpening: 84.7% Softening: 15.3%	Sharpening: 77.4% Softening: 22.6%
Engagement	33 Instances of Contraction	55 Instances of Contraction
	Proclaiming: 90.9% Disclaiming: 9.1%	Proclaiming: 70.9% Disclaiming: 29.1%
Engagement	15 Instances of Expansion	23 Instances of Expansion
	Entertaining: 100% Attributing: 0.0%	Entertaining: 87% Attributing: 13%

Table 11 shows that there are 77 instances of Appraisal in the PSSA rubric and there are 62 instances of Appraisal in the KE rubric. In the PSSA rubric, these instances account for 21.6% of the total words and in the KE rubric, these instances account for 13.7% of the total words. For the specific types of Appreciation, there is more Reaction Impact, Reaction Quality, Composition Complexity, and Ambiguity in the PSSA rubric than there is in the KE rubric. The entire category of Composition, however, contains the greatest amount of Appreciation in both of the rubrics.

For Graduation Force, the table shows that there is more Raising in the PSSA rubric than in the KE and there is more Lowering in the KE rubric than in the PSSA. The difference between Raising and Lowering in the PSSA rubric is 11.6% and the difference between Raising and Lowering in the KE rubric is 9.6%. Even though one rubric features more Raising or more Lowering than the other, the difference between both of these subcategories is not great. For Graduation Focus, there is more Sharpening in the PSSA rubric than in the KE and there is more Softening in the KE rubric than in the PSSA. The difference between Sharpening and Softening in the PSSA rubric is 69.4% and the difference between Sharpening and Softening in the KE rubric is 54.8%. Again, these differences are not that much different for each of the rubrics.

The table also shows that for Engagement, there are more instances of Contraction than Expansion in the rubrics. Comparatively, however, there are more instances of Contraction and Expansion in the KE rubric than in the PSSA. Within the category of Contraction, the PSSA features 20% more Proclaiming than the KE and the KE features 20% more Disclaiming than the PSSA. Within the category of Expansion, 100% of the instances in the PSSA can be categorized as Entertaining, while 87% of them can be in the KE. This means that the KE rubric does feature some Attributing.

Appraisal by score. The table below shows the similarities and differences in Appraisal between the PSSA and KE rubrics. The table highlights specifics pertaining to the instances of Appreciation, Graduation, and Engagement by essay score.

As can be seen in Table 12, the Appreciation values between the rubrics differ by score just as they differ by overall percentage. For example, in the PSSA rubric, Composition Balance is greatest for the highest score and in the KE rubric, Composition

Balance is greatest for the lowest score. Interestingly, the percentage of Composition Balance falls in both rubrics for the second highest score and then rises again for third highest score. For Composition Complexity, however, the instances in both rubrics can be seen decreasing as the score decreases.

In terms of Graduation, Table 12 shows that in both the PSSA and the KE rubrics, the percentage of Raising decreases as the score increases, while the percentage of Lowering increases, for the most part, as the score decreases. The table also shows that Sharpening is most prevalent for the highest score in both of the rubrics, while the distribution of Softening varies. Interestingly, the greatest amount of Softening occurs in the lowest score for the PSSA while the greatest amount of Softening occurs in the highest score for the PSSA.

Table 12 also shows that with Engagement, the amount of Proclaiming is greatest for the highest score in both rubrics, while the amount of Disclaiming is greatest for the lowest score in both rubrics. In terms of Expansion, the KE features more Entertaining for the middle scores than the PSSA does, and it also features Attribution for the top three highest scores whereas the PSSA does not feature any.

Table 12 <i>Similarities and Differences in Appraisal Between the PSSA and KE Rubrics by Essay Score</i>		
Appreciation	<p>Percentages are given from left to right from highest to lowest score.</p> <p>Reaction Impact: 1.3%, 1.3%, 0%, 0%</p> <p>Reaction Quality: 0%, 3.9%, 5.2%, 1.3%</p> <p>Composition Balance: 7.8%, 3.9%, 9.1%, 16.9%</p> <p>Composition Complexity: 14.3%, 10.4%, 5.2%, 2.6%</p> <p>Valuation: 3.9%, 1.3%, 0%, 0%</p> <p>Ambiguous: 3.9%, 2.6%, 2.6%, 2.6%</p>	<p>Percentages are given from left to right from highest to lowest score.</p> <p>Reaction Impact: 0%, 0%, 0%, 0%, 0%</p> <p>Reaction Quality: 0%, 0%, 3.2%, 0%, 0%</p> <p>Composition Balance: 8.1%, 3.2%, 9.7%, 14.5%, 14.5%</p> <p>Composition Complexity: 11.3%, 4.8%, 4.8%, 4.8%, 1.6%</p> <p>Valuation: 8.1%, 4.8%, 3.2%, 1.6%, 0%</p> <p>Ambiguous: 1.6%, 0%, 0%, 0%</p>
Graduation	<p>Force - Raising: 31.2%, 15.6%, 5.2%, 3.9%</p> <p>Force - Lowering: 0%, 7.8%, 16.9%, 19.5%</p> <p>Focus - Sharpening: 30.6%, 20.8%, 16.7%, 16.7%</p> <p>Focus - Softening: 0%, 2.8%, 4.2%, 8.3%</p>	<p>Force - Raising: 29%, 12.9%, 1.6%, 1.6%, 0%</p> <p>Force - Lowering: 0%, 0%, 19.4%, 19.4%, 16.1%</p> <p>Focus - Sharpening: 25.8%, 11.3%, 12.9%, 12.9%, 14.5%</p> <p>Focus - Softening: 29%, 12.9%, 21%, 21%, 16.1%</p>
Engagement	<p>Contraction - Proclaim: 27.3%, 21.2%, 21.2%, 21.2%</p> <p>Contraction - Disclaim: 0%, 0%, 0%, 9.1%</p> <p>Expansion - Entertain: 20%, 20 %, 20%, 40%</p> <p>Expansion - Attribute: 0%, 0%, 0%, 0%</p>	<p>Contraction - Proclaim: 23.6%, 14.5%, 10.9%, 14.5%, 7.3%</p> <p>Contraction - Disclaim: 0%, 0%, 0%, 7.3%, 21.8%</p> <p>Expansion - Entertain: 4.3%, 4.3 %, 39.1%, 34.8%, 4.3%</p> <p>Expansion - Attribute: 4.3%, 4.3%, 4.3%, 0%, 0%</p>

CHAPTER 6

DISCUSSION

Under NCLB, high-stakes testing has become a more prevalent part of K-12 education, and arguments about how to best assess students' writing competence under such circumstances have intensified. With a new Common Core aligned exam on the horizon in Pennsylvania, the need to examine the language of the assessment criteria on the old test, the PSSA, and weigh it against the criteria on the new test, the KE, is urgent. Given this time-sensitive task, I began this research by posing the following questions: 1) What kind of writing is valued in the PSSA and KE rubrics? 2) How do the values change, if at all, between ratings and between the two sets of rubrics? and 3) How do these changes affect how the expectations in the rubrics are interpreted?

The findings previously discussed suggest that the criteria in the PSSA and KE persuasive writing rubrics are alarmingly similar, which means that the writing portion of the Keystone Exam will not align to the Common Core State Standards. I contend here that the frequency of the Appraisal resources can indicate the value that test makers place on certain items in the rubric. I will discuss the significance of the results in terms of having a relatively high occurrence or a relatively low occurrence in relation to individual scores and overall criteria. This means that the higher the frequency of an Appraisal resource, the more value that is placed on a given item within the students' essay.

In general, the findings show that students are largely awarded for how their essay hangs together (Composition: Balance) and for how difficult or easy it is to follow

(Composition: Complexity). The findings also show that there is more positive (Gradation: Raise) appraisal for the higher scores and more negative (Graduation: Lower) appraisal for the lower scores. This positive and negative Graduation is, in addition, connected to the Sharpening (e.g. when the specification of an appraised item is maximized) that occurs for the highest and lowest scores and the Softening (e.g. when the specification of an appraised is blurred) that occurs for the scores that exist in the middle of the spectrum. This distribution suggests that the expectations for the scores in the middle are more open to opinion and interpretation since the appraisal resources used serve to make the already vague criteria even more inexplicit by exploiting terms like “some,” which opens up a space for the rubric user to decide exactly what counts as “some” given that specifics are not provided. Despite this open space, however, the amount of Contraction (e.g. when the dialogical space closes so as to reject or deny any alternative positions) present in each of the rubrics suggest that overall, there is less room for different stances in the rubrics, especially within the highest and lowest scores. The following sections will detail what these findings mean.

Appraisal in the Rubrics

The model of writing competence suggested by the Pennsylvania State Standardized Assessment provides an interesting comparison to the Keystone Exam. Both feature similar, broad domains – Focus, Content, Organization, and Style – which hold students accountable for general characteristics of good writing. Subsequently, these general characteristics and their effects will be described in reference to the Appreciation, Graduation and Engagement in each of the rubrics below.

Appreciation. As the Appreciation analysis showed, the Balance and Complexity of the essays was the most prevalent piece of appraisal in both rubrics, which suggests that rather than awarding points for effectiveness in responding to a specific rhetorical context, students are rewarded points mostly for the textual components of their essays. These findings seem to reiterate the assumption, as put forth by McNamara, Crossley, and McCarthy (2010), that cohesion is seen in the writing rubric as intimately linked with comprehension, sufficient or insufficient evidence, and the overall quality of an essay. The problem, however, is that the textual component is only one part of the writing.

Most interesting in the PSSA and KE rubrics is that the instances of Composition Balance are greater for the lowest possible score (7.9% compared to 16.9%; 8.1% compared to 14.5%) while the instances of Composition Complexity are greater for the highest possible score (14.3% compared to 2.6%; 11.3% compared to 1.6%). This distribution suggests that better qualities of writing are recognized for being easy to follow while poorer qualities of writing are recognized for not hanging together. For example, terms such as “clear” and “precise” are used to appraise how unproblematic it is to follow the writer’s position within the higher scores, whereas terms such as “underdeveloped” and “undeveloped” are used to appraise how discordant the presentation of the writer’s position is within the lower scores. The distinction between the two, however, is not made clear in the rubrics, as there are not any indicators of how this cohesion or coherence is created.

One speculation for this distribution is that there is a difference between cohesion and coherence that many people are not aware of (O’Reilly and McNamara 2007). While

cohesion refers to specific cues that can be identified in a text (e.g. causal relationships such as because or therefore, referential overlap such as repetition of words or concepts), coherence exists in the mind of the reader. In other words, readers with a high knowledge of the topic can make inferences while reading, which may make a text seem more coherent than it is because they are not affected by its cohesion. Halliday and Hasan (1976) acknowledge this notion as well, explaining that there are additional aspects of textual meaning that cannot be denoted by cohesion. For example, word frequency, syntactic complexity, and lexical diversity contribute to perceptions of coherence in the reader's mind.

Given these differences, it seems that cohesion may actually be present in the requirements of the Organization domain in the PSSA and KE as there is reference to logical order and transitions, and in the Style domain as there is reference to control of language and sentence structures. This does not account for the appearance of Composition in all of the domains, however. Additionally, even though these features of cohesion and coherence are present, there are not any explanations as to how or why such features are necessary to respond to the specific rhetorical context of the genre, and as Beck and Jeffery (2007) make clear, explicitness in the importance of features is critical for distinguishing one type of writing from another. This similarity between the two rubrics in terms of a lack of specification thus does not bode well for the KE's supposed alignment to the Common Core State Standards.

Aside from the high prevalence of appraisal dealing with Composition in the rubrics, the PSSA also features a higher percentage of Reaction Quality (10.4%), which again deals with whether or not the appraised item is liked. This percentage raises some

concern, as Reaction in general is a realization of affection, which means that it elicits an emotional response. In other words, Reaction has little to do with any specific features in the text and instead, as Mabry (1990) explains, addresses how well something should be done, not what should be done, which is typical of the language in rubrics and is often used as a means to obscure already vague expectations. Given that all but one of the instances of Reaction Quality occurs in Score 3 and Score 2, it seems as if this type of Appraisal may help to blur expectations at both of these levels. For example, “adequate” (Score 3) and “inadequate” (Score 2) are used to describe the writer’s content, but given that there are no specifications as to what counts as “adequate” or “inadequate,” the appraisal stands merely as an affectual critique. Quite possibly, this ambiguity could be due to not knowing how to specifically critique an essay that meets some but not all of the requirements (Hillocks, 2002), as it is much easier to identify the quality of an essay when it is at one end of the scoring spectrum or the other.

This ambiguousness is seen again in the KE rubric with Valuation instead of Reaction. Valuation, which is used to appraise how worthwhile a product is, is related to cognition or opinion. This means that like Reaction, Valuation has an attitudinal quality. For example, claiming that there is an “effective style and tone” within a piece of writing implies a judgment and given that there are no explanations as to what creates this effectiveness, this instance of Valuation lacks credibility as it is simply the considered opinion put forth by the authorial voice in the rubric. Alarming in the KE, this Valuation is seen occurring 17.7% of the time, 8.1% of which is in the criteria for the highest score. Even though this distribution does not bode well for informing the writer of what he/she needs to do in order to complete a task, it does make sense as an essay

with a higher score is more likely to receive praise for being worthwhile or effective (Hillocks, 2002). However, even strong writers need to know what is expected of them and these high instances of Valuation seem to suggest otherwise as the criteria for the highest score does not characterize specifically what top-notch writing looks like.

Of additional concern is the percentage of Ambiguous cases (11.7%) in the PSSA rubric. Recall that Ambiguous cases were classified as having characteristics of two or more Appreciation types. Thus, a term such as “substantiated,” as will be discussed below, could be classified as either Valuation or Composition. Interestingly, almost all of the Ambiguous instances are on account of the word “substantiated,” which is seen occurring in each of the scores. But what exactly does substantiated mean? Merriam-webster.com defines substantiated as “1) to give substance or form to, embody or 2) to establish by proof or competent evidence, verify.” Given this definition, there is an inclination to categorize it with Valuation as it seems to connote something that is justifiable or appropriate, especially when appraising the word “argument,” as it does in the rubric. There is, however, also an inclination to categorize it with Composition, as it has a textual quality. The question is how does this ambiguity affect how the expectations are interpreted?

In a 2006 study on writing perceptions, Beck claimed that individuals must rely on their own intellect of what good writing is in order to interpret rubrics and requirements. For developing writers and second language writers, this reliance may be culturally ambiguous, as not every group has the same cultural values. In other words, when specifics are not provided in the rubric as to what features or rhetorical strategies are needed to create a “substantiated argument,” these populations may fail to complete

the task because their interpretation of what counts as a “substantiated argument” may be vastly different from the state’s interpretation. The language within the rubric, in this case at least, can thus be seen as privileging U.S. born people. Furthermore, given the ambiguity of the phrase, some writers and evaluators may take it to mean justifying a claim while others may take it to mean shaping or structuring a claim. The difference in interpretation thus creates more room for misalignment between the students’ writing and the evaluation of it.

Overall, analysis reveals that the use of Appreciation in both the PSSA and KE rubrics shows an attempt to focus on the features in the persuasive genre with use of terms such as “a well-defined introduction” and “includes a clear position” present. However, realizations of Appreciation only provide a means to evaluate items, not specify characteristics of good writing. For example, as shown in the quantitative analysis in Chapter 5 and in my discussion above, the focus on Balance and Composition, both used as appreciation of the textual components of writing, evaluate how well the writing hangs together or how difficult or easy it is to follow, not specify what allows it to hang together or makes it easy to follow. As a result, even though the majority of Appreciation values are in reference to the evaluation of the textual components of writing, they do not denote, for example, the use of logical connectors. It thus seems that Appreciation is used in place of specification in both rubrics and this use serves to broaden the vagueness of the criteria.

The disturbing similarity between both the PSSA and KE in terms of Appreciation is also critical, as the KE is supposed to be aligned with the newly instated Common Core State Standards, which are seen as creating “clearer” expectations for all students in

regards to what specifically they should know and be able to do in order to succeed in college and the workforce (Rothman, 2011). For example, the large amount of Reaction (e.g. used to express emotion) in the PSSA and the large amount of Valuation (e.g. used to express judgment) in the KE creates a vagueness in the rubrics as they elicit an interpretation from the rubric user as to what counts as “good” or “effective” writing since no specifications are provided. This interpretation is problematic because without any particulars, there is more room for misalignment between what the state expects from writers and evaluators and what they actually receive from writers and evaluators. Furthermore, this makes scoring the essays more flawed and thus decreases the validity and reliability of the rubrics.

Graduation. As the Graduation analysis showed, more Raising occurred for the higher scores in both rubrics and more Lowering occurred for the lower scores, which does not come as a surprise as Raising is often associated with positivity and Lowering is often associated with negativity (Martin & White, 2005). The same is true of Sharpening and Softening of which there was a similar distribution in both the PSSA and the KE.

A potential complication with this positivity and negativity is that there are discipline-specific attitudes one is expected to have towards certain practices (Martin, 2000). In other words, the positive or negative connotations associated with the appraised items in the rubrics involve an awareness that is garnered from being socialized into the discipline and genre. Thus, any misalignment from students or evaluators in terms of attitudinal expectations, or even from myself as a researcher, could affect the reading of the rubrics and alter the scalability of such realizations. For example, the use of “relevant” in the rubrics compared to “specific” could be perceived as somewhat

problematic. While both appraise the term “content” – relevant content and specific content – they appear in different scores. Without looking at the scores and their placement, my initial thought would be that to have “specific content” is better than to have “relevant content.” However, this is not the case in the PSSA rubric, as “relevant” appears for Score 4 and “specific” appears for Score 3. And even though after looking at the rubric I am able to note this, I am still not sure what constitutes the difference between why “relevant” is awarded more points than “specific” because the rubric does not indicate this.

Interesting to note is that English has more resources for grading items positively than it does for grading them negatively (Martin & Rose, 2003), which would explain why there is a lot more Raising and Sharpening in each of the rubrics than there is Lowering and Softening. Also, given that the lower graded items seem to be more obvious with the use of prefixes like “un-“ and “in-” and key terms such as “minimal” and “no,” making students and evaluators aware of these indicators could help in deciphering the scalability of items, at least for the highest and lowest scores.

It is also important to take notice of where exactly Softening occurs in the rubrics, as Softening serves to blur the focus of the categorization. In the PSSA, the instances of Softening increase as the scores decrease. Additionally, there are not any instances of Softening for Score 4. In the KE, the scores in the middle, Score 2 and Score 1, contain almost all the instances of Softening in the rubric (16.2%). The question is what effect does this have on the rubric?

Given that Softening serves as a broadening mechanism, it makes sense that terms like “some” and “a variety of” are used more often in the criteria for the scores in the

middle of the spectrum, as it is difficult to define what exactly constitutes an average piece of writing as opposed to a very strong or very weak piece of writing. The Softening is thus used to further obscure what exactly is needed or not needed to be awarded points at these levels and as a consequence, opens up a larger space for interpretation as the rubric users must decide for themselves what exactly counts as and differentiates between “some evidence” and “minimal evidence.”

Overall, the presence and distribution of Graduation in each of the rubrics suggests that students, teachers, and evaluators must be socialized into the discipline in order to understand the requirements in the rubrics. The realizations of positive and negative values are not always clear cut and this difficulty in scaling certain items makes delineating between expectations at different levels a complicated task. Furthermore, the use of Softening creates additional obscurity on top of the already vague guidelines. It thus seems that Graduation is used along with Appreciation in place of specification, especially in the scores that exist in the middle of the rubrics.

Engagement. As the Engagement (e.g. how much room there is for alternative positions) analysis showed, both rubrics featured more Contraction than Expansion, which suggests that there is less room for any alternative positions when interpreting the guidelines. However, the majority of the Contraction occurred for the highest and lowest scores in both of the rubrics and the majority of the Expansion occurred for the middle scores. While Engagement in the PSSA was more evenly distributed than in the KE, the overall spread of Contraction and Expansion seems to reiterate the difficulties in interpreting the guidelines for an average piece of writing, as the criteria in the higher and lower scores appear to be less negotiable than the criteria in the middle scores.

The problem is that even when the criteria in the rubric Contracts, it does not make the expectations any clearer. For example, the rubric reads, “The writer provides relevant content ...” and “The writer displays some evidence ...” While “provides” and “displays” signify that the authorial voice is proclaiming what the writer has done by closing down the space for any opposing positions, the Contraction does not aid in interpreting what exactly constitute “relevant content” or “some evidence.” The contraction is only there for the purposes of being able to check, “yes, the writer provided this” or “no, the writer did not provide this.”

In terms of Expansion having a higher rate of occurrence for the middle scores, the explanation is actually quite straightforward. Just as Softening obscures the expectations for the middle scores, Expansion serves to open up these criteria for further interpretation. For example, the use of “may” is highly prevalent in the criteria for the middle scores in each of the rubrics and as Martin and White (2005) explain, modal auxiliaries allow the authorial voice to make assessments of likelihood, thereby indicating that a given position is but one of a range of possible positions. Thus, for the criteria in the middle scores, there is more dialogical space for students, teachers, and evaluators to posit whether or not, for instance, “simplistic expressions” were used. As a result, this Expansion makes the rubric less valid and less reliable as there is more room for conflicting viewpoints.

Overall, the Engagement in the rubrics has shown that while there is less space for alternative positions in interpreting whether a piece of writing does or does not exemplify something, it does not aid in specifying what the requirements within the criteria are. In addition, for the middle scores where the dialogical space is opened up, there is more

room for interpretation, which seems to reiterate the idea that it is easier to more definitively assess writing that is either really strong or really weak.

Summary. This discussion has shown the complications that can come along with reading interpersonal meaning as the interpretation of how items are evaluated can be altered by one's culture, immersion into a given discipline, and lack of specifications as to what makes an evaluation valid. Given that the rubrics used the Appraisal resources extensively and oftentimes in place of specifying what the writer actually needs to do, there is a need for more explicitness in the guidelines. As Mabry (1990) notes, rubrics are artifacts that represent what the test taker should know, but in the case of the PSSA and the KE, this knowing comes along with being socialized into the discipline-specific modes of writing, as what constitutes the terms used in the rubrics is typically not made clear and is often obscured by the use of Appreciation, Graduation, and Engagement. Drawing attention to specific structural and register features in genres of writing thus becomes critical and rubrics, if crafted in a way that help the test taker realize how the text's organization and grammatical and lexical features work to imply a particular purpose, could help foster the teaching of writing as well as first and second language writing development.

Implications for the Creation of Rubrics

Given that my SFL analysis of the rubrics clearly shows that language learners, teachers, and evaluators could easily be confounded by the very vague and multiple interpretations of the rubrics, this section proposes an SFL-informed approach to designing rubrics. I propose this especially in light of the new Common Core State Standards that call for text complexity that is not only related to Balance and Complexity,

but that instead highlights how content and form interconnect at the ideational, interpersonal, and textual level. Given the multiple interpretations of what counts as good writing according to the PSSA and KE and the lack of specification as to what writers need to do to complete the tasks, I think genre-based rubrics would offer a sound alternative, as they have the potential to create clearer expectations for all students, teachers, and evaluators. Genres, which are dynamic social processes that are shaped by and construe social context, represent a specific text or type of discourse (Martin, 1996). Hence, the expository texts which were examined in this research include certain features that allow them to be classified as such. For example, in order for a piece of writing to be expository, it must contain an argument that provides supporting evidence and weighs different views. To achieve this, writers must use nominalizations to name the supporting evidence, clause organization strategies to show causal relationships and attitudes, and modality and attitudinal connectors to guide and build upon the argument (Schleppegrell, 2000). The KE rubric, which is a reflection of the Common Core State Standards, does this to a certain extent, but the fact that particular linguistic and rhetorical features work to create meaning needs to be more prevalent.

While genre-based pedagogies have been implemented in classrooms in order to facilitate students' writing (e.g. Cope & Kalantzis, 1993), the assessment criteria in such cases are still not made clear and often create a discrepancy between what is expected and what is thought to be expected (Bardine, Bardine, & Deegan, 2000). A detailed examination of what the expository genre entails is thus needed in order to create a valid rubric, or in this case, to modify an existing one. For this to work, however, the rubric needs to be as specific as possible without being excessive. Since rubrics inherently

standardize writing by quantifying it and reducing it to set a variables (Hillocks, 2002), the goal of the designer should be to be as explicit and comprehensive as possible. For example, instead of stating that the student should provide “clear and substantiated evidence,” the designer should explain what exactly counts as “clear and substantiated evidence” and how certain features work to do this as it relates to the expository genre. This would benefit the students in two ways. First, a detailed rubric sets up clear expectations so that students know exactly what needs to be done in order to successfully complete the task. Second, the evaluator and the teacher are also very aware of the expectations and can thus provide the student with more meaningful feedback, pointing directly to the structural elements and features present or not present in the students’ writing and how they are or are not working to create the type of meaning desired. As Hirsch (1977) notes, “We cannot get reliable, independent agreement in the scoring of writing samples unless we also get widespread agreement about the qualities of good writing.” A detailed, genre-based rubric thus has the potential to create clear expectations for all parties involved, thereby increasing the validity and reliability of such assessments, as the structure and features of a particular genre of writing would have the same discipline-specific expectations across the board. In addition, a genre-based rubric has the potential to help prepare students to succeed at writing tasks at the post-secondary level, as the genres represented in the Common Core based exams, like the KE, are those associated with particular fields of study and being in college and the workforce.

Conclusion

The text-based examination that I presented in this paper motivates us to focus on creating rubrics with greater clarity in what is expected of student writers. Given the

alarming similarities between the two rubrics, it does not appear that the KE will align to the Common Core State Standards. An obvious next step would be to weigh the CCSS against the KE to examine where this probable misalignment occurs, but of greater importance, I would argue, is to illuminate and explicate the structural and register features of particular genres of writing, as such a practice would enable teachers to better prepare students for writing in different contexts and for different purposes. In addition, this explicit attention to features could also increase the consistency of scoring writing in the classroom and on high-stakes exams, as writing competence would be measured specifically by what features the writer did or did not include in his/her essay to aid in responding to the specific communicative context. Using insights from Systemic Functional Linguistics along with the quantitative methods above thus has the potential to shed light on new links between writing development, instruction, and assessment.

REFERENCES

- 2011-2012 PSSA. *Resource materials*. Retrieved from http://www.portal.state.pa.us/portal/serve.pt/community/pennsylvania_system_of_school_assessment_%28pssa%29/8757/resource_materials/507610
- Applebee, A. N., Langer, J. A. (2011). A snapshot of writing instruction in middle schools and high schools. *English Journal*, 100 (6), 14-27.
- Ball, D. (2000). Bridging practices: Inter-twining content and pedagogy in teaching and learning to teach. *Journal of Teacher Education*, 51, 241-247.
- Bardine, B.A., Bardine, M.S., and Deegan, E.F. (2000). Beyond the red pen: Clarifying our role in the response process students. *English Journal*, 90(1), 94-101.
- Barone, D., & Taylor, J. (2006). *Writing K-8*. Thousand Oaks, CA: Corwin Press.
- Beck, S. (2006). Subjectivity and intersubjectivity in the teaching and learning of writing. *Research in the Teaching of English*, 40 (4), 413-460.
- Beck, S. and Jeffery, J. (2007). Genres of high-stakes writing assessments and the construct of writing competence. *Assessing Writing*, 12, 60-79.
- Broad, B. (2003). *What we really value: Beyond rubrics in teaching and assessing writing*. Logan: Utah State UP.
- Cameron, D. (2001). *Working with spoken discourse*. Los Angeles: SAGE.
- Charney, D. (1984). The validity of using holistic scoring to evaluate writing: A critical overview. *Research in the teaching of American english*, 18 (01), 65-81.
- Cooper, C. (1999). What we know about genres, and how it can help us assign and evaluate writing. In C. Cooper & L. Odell (eds.). *Evaluating writing: The role of teachers' knowledge about text, learning, and culture*. Urbana, IL: NCTE. 23-52.
- Cope, B., & Kalantzis, M. (Eds.) (1993). *The powers of literacy: A genre approach to teaching writing*. Pittsburgh, PA: University of Pittsburgh Press.
- Data Recognition Corporation. (2011). *Technical report for the 2011 pennsylvania system of school assessment*. Retrieved from: http://www.portal.state.pa.us/portal/server.pt/community/technical_analysis/7447.

- Fang, Z. & Schleppegrell, M. (2008). *Reading in secondary content areas: A language based pedagogy*. Ann Arbor: The University of Michigan Press.
- Gibbons, P. (2006). *Bridging discourses in the esl classroom*. New York: Continuum.
- Goodrich Andrade, H. (2001). The effects of instructional rubrics on learning to write. *Current Issues in Education*, 4 (4). Available: <http://cie.ed.asu.edu/volume4/number4/>.
- Halliday, M.A.K. & Hasan, R. (1976). *Cohesion in english*. London: Longman.
- Halliday, M.A.K., & Hasan, R. (1989). *Language, context and text: A social semiotic perspective*. Oxford, UK: Oxford University Press.
- Harman, R. (Summer 2013). Literary intertextuality in genre-based pedagogies: Building lexical cohesion in fifth-grade L2 writing. In Special Issue, (Heidi Byrnes, Editor) *Journal of Second Language Writing*.
- Hillocks, G. (2002). *The testing trap: How state writing assessments control learning*. New York: Teachers College Press.
- Hirsch, E. D., Jr. (1977). *The philosophy of composition*. Chicago: University of Chicago Press.
- Hood, S. 2004. *Appraising research: Taking a stance in academic writing*. University of Technology Sydney PhD. Thesis.
- Implementing the common core standards in pennsylvania: Reflections from the field. *Common Core State Standards*. Retrieved from http://www.portal.state.pa.us/portal/server.pt/community/current_initiatives/19720/common_core_state_standards/792440
- Kiuhara, S. A., Graham, S., Hawken, L.S. (2009). Teaching writing to high school students: A national survey. *Journal of Educational Psychology*, 101 (1), 136-160.
- Kohn, A. (2006). The trouble with rubrics. *English Journal*, 95(4), 12-15.
- Kress, G. (1994). *Learning to write* (2nd ed.). London: Routledge.
- Lane, S. (2010, June). Evaluation of the alignment between PA academic standards and the common core standards. Retrieved from http://www.portal.state.pa.us/portal/server.pt/community/current_initiatives/19720/common_core_state_standards/792440
- Lumley, D. R., & Yan, W. (2001). The impact of state mandated large-scale writing

- assessment policies in pennsylvania. *U.S. Department of Education*. Web. 19 Jan. 2013.
- Mabry, L. (1990). Writing to the rubric: Lingering effects of traditional standardized testing on direct writing assessment. *The Phi Delta Kappan*, 80 (9), 673-679.
- Macken-Horarik. (2003). Appraisal and the special instructiveness of narratives. 23 (2), 285-312.
- Martin, J.R. (1996). Evaluating disruption: Symbolising theme in junior secondary narrative. In *Literacy and Society*. Eds. Ruqaiya Hasan and Geoff Williams. London: Longman.
- Martin, J.R. (2000). Beyond exchange: Appraisal systems in english. In S. Hunston and G. Thompson (eds.). *Evaluation in text: Authorial stance and the construction of discourse*. Oxford: Oxford University Press. 142-175.
- Martin and Rose (2003). *Working with discourse: Meaning beyond the clause*. London: Continuum.
- Martin, J.R., and White, P.R.R. (2005). *Language of evaluation*. London: Palgrave Macmillan.
- McCarthy, S.J. (2008). The impact of no child left behind on teachers' writing instruction. *Written communication*, 25(4), 462-505.
- McNamara, D.S., Crossley, S.A., and McCarthy, P.M. (2010). Linguistic features of writing. *Written Communication*, 27(1), 57-86.
- Mixed reactions to NCLB. (2005, February). *Reading Today*, 22(4), 1-4.
- National Center for Education Statistics. English language learners. *Fast facts*. U.S. Department of Education. Web. 18 Jan. 2013.
- O'Conner, R., Abedi, J., and Tung, S. (2012). *A descriptive analysis of enrollment and achievement among english language learner students in pennsylvania*. (Issues and Answers Report, REL 2012–No. 127). Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory Mid-Atlantic. Retrieved from <http://ies.ed.gov/ncee/edlabs>.
- O'Reilly, T. and McNamara, D.S. (2007). The impact of science knowledge, reading skill, and reading strategy knowledge on more traditional "high-stakes" measures or high school students' science achievement. *American Educational Research Journal*, 44. 161-196.

- Pennsylvania Department of Education. English as a second language. Commonwealth of Pennsylvania. Web. 2013. Pennsylvania System of School Assessment. (2000). Writing assessment handbook. Harrisburg, PA: Pennsylvania Department of Education.
- Rothman, Robert. (2011). *Something in common: The common core standards and the next chapter of american education*. Cambridge, MA: Harvard Education Press.
- Saddler, B. & Andrade, H. (2004). The writing rubric. *Educational leadership*, 62(2), 48-52.
- Schleppegrell, M. J. (2000). How SFL can inform writing instruction: The grammar of expository essays. *Revista Canaria de Estudios Ingleses*, 40, 171-188.
- Schulze, J. (2011). Writing to persuade: A systemic functional view. *Gist Education and Learning Research Journal*, 5, 127-157.
- Spandel, V. (2006). Speaking my mind: In defense of rubrics. *The English Journal*, 96(1), 19-22.
- Turley, E.D., & Gallagher, C.W. (2008). On the uses of rubrics: Reframing the great rubric debate. *English Journal*, 97(4), 87-92.
- U.S. Department of Education. (2009). Standards and assessments peer review guidance: Information and examples for meeting requirements of the no child left behind act of 2001 [PDF document]. Retrieved from <http://www.2.ed.gov/admins/lead/account/saa.html>
- Weaver, C. (1996). *Teaching grammar in context*. Portsmouth, NH: Boynton/Cook.
- White, P.R.R. (2012). *The appraisal website*. Verbosity Enterprises. Web. 15 Jan. 2013. <http://grammatics.com/appraisal/>
- Wilson, M. (2006). Rethinking rubrics in writing assessment. Portsmouth, NH: Heinemann.
- Wilson, M. (2007). "Why I won't be using rubrics to respond to students' writing." *The English Journal*, 96(4), 62-66.
- Young, L. & Fitzgerald, B. (2006). *The power of language*. London: Equinox

APPENDICES

APPENDIX A: PSSA WRITING RUBRIC

4	FOCUS	Sharp, distinct controlling point presented as a position and made convincing through a clear, thoughtful, and substantiated argument with evident awareness of task and audience.
	CONTENT DEVELOPMENT	Substantial, relevant, and illustrative content that demonstrates a clear understanding of the purpose. Thoroughly elaborated argument that includes a clear position consistently supported with precise and relevant evidence. Rhetorical (persuasive) strategies are evident.
	ORGANIZATION	Effective organizational strategies and structures, such as logical order and transitions, to develop a position supported with a purposeful presentation of content.
	STYLE	Precise control of language, stylistic techniques, and sentence structures that creates a consistent and effective tone.
3	FOCUS	Clear controlling point presented as a position and made convincing through a credible and substantiated argument with general awareness of task and audience.
	CONTENT DEVELOPMENT	Adequate, specific and/or illustrative content that demonstrates an understanding of the purpose. Sufficiently elaborated argument that includes a clear position supported with some relevant evidence. Rhetorical (persuasive) strategies may be evident.
	ORGANIZATION	Organizational strategies and structures, such as logical order and transitions, to develop a position supported with sufficient presentation of content.
	STYLE	Appropriate control of language, stylistic techniques, and sentence structures that creates a consistent tone.
2	FOCUS	Vague evidence of a controlling point presented as a position that may lack a credible and/or substantiated argument with an inconsistent awareness of task and audience.
	CONTENT DEVELOPMENT	Inadequate, vague content that demonstrates a weak understanding of the purpose. Insufficiently elaborated argument that includes an underdeveloped position supported with little evidence.
	ORGANIZATION	Inconsistent organizational strategies and structures, such as logical order and transitions, to develop a position with inadequate presentation of content.
	STYLE	Limited control of language and sentence structures that creates interference with tone.
1	FOCUS	Little or no evidence of a controlling point presented as a position that lacks a credible and/or substantiated argument with minimal awareness of task and audience.
	CONTENT DEVELOPMENT	Minimal evidence of content that demonstrates a lack of understanding of the purpose. Unelaborated argument that includes an undeveloped position supported with minimal or no evidence.
	ORGANIZATION	Little or no evidence of organizational strategies and structures, such as logical order and transitions, to develop a position with insufficient presentation of content.
	STYLE	Minimal control of language and sentence structures that creates an inconsistent tone.

APPENDIX B: KE WRITING RUBRIC

ENGLISH COMPOSITION

V

MODULE 2

KEYSTONE PERSUASIVE SCORING GUIDELINES

Scoring Domain	Score Point 4 At this score point, the writer—	Score Point 3 At this score point, the writer—	Score Point 2 At this score point, the writer—	Score Point 1 At this score point, the writer—	Score Point 0 At this score point, the writer—
Thesis/Focus	<ul style="list-style-type: none"> establishes and sustains a precise claim or position displays a clear understanding of task, purpose, and audience 	<ul style="list-style-type: none"> establishes a claim or position displays an understanding of task, purpose, and audience 	<ul style="list-style-type: none"> provides an inconsistent claim or position displays a limited understanding of task, purpose, and audience 	<ul style="list-style-type: none"> provides vague or indistinct claim or position displays a minimal understanding of task, purpose, and audience 	<ul style="list-style-type: none"> provides no evidence of claim or position displays no understanding of task, purpose, and audience <p>OR</p> <ul style="list-style-type: none"> does not respond to prompt
Content	<ul style="list-style-type: none"> provides relevant content and specific and effective supporting details that demonstrate a clear understanding of purpose uses sophisticated transitional words, phrases, and clauses to link ideas and create cohesion considers possible counterclaims (alternate or opposing arguments) 	<ul style="list-style-type: none"> provides relevant content and effective supporting details uses transitional words, phrases, and clauses to link ideas acknowledges possible counterclaims (alternate or opposing arguments) 	<ul style="list-style-type: none"> provides insufficient content and ineffective supporting details may use simplistic and/or illogical transitional expressions may not acknowledge possible counterclaims (alternate or opposing arguments) 	<ul style="list-style-type: none"> provides minimal content uses few or no transitional expressions to link ideas does not acknowledge possible counterclaims (alternate or opposing arguments) 	<ul style="list-style-type: none"> provides little to no content does not use transitions to link ideas <p>OR</p> <ul style="list-style-type: none"> does not respond to prompt
Organization	<ul style="list-style-type: none"> chooses sophisticated organizational strategies appropriate for task, purpose, and audience presents fair and relevant evidence to support claim or position includes a clear and well-defined introduction, body, and conclusion that support or reinforce the argument 	<ul style="list-style-type: none"> chooses appropriate organizational strategies for task, purpose, and audience presents relevant evidence to support claim or position includes a clear introduction, body, and conclusion that support the argument 	<ul style="list-style-type: none"> displays some evidence of organizational strategies presents insufficient evidence to support claim or position may not include an introduction, body, and/or conclusion 	<ul style="list-style-type: none"> displays little evidence of organizational strategies presents little or no evidence to support claim or position may not include an identifiable introduction, body, and/or conclusion 	<ul style="list-style-type: none"> displays no evidence of organizational strategies presents no evidence to support claim or position does not include an identifiable introduction, body, and/or conclusion <p>OR</p> <ul style="list-style-type: none"> does not respond to prompt
Style	<ul style="list-style-type: none"> uses consistently precise language and a wide variety of sentence structures chooses an effective style and tone, and maintains a consistent point of view 	<ul style="list-style-type: none"> uses precise language and a variety of sentence structures chooses an appropriate style, tone, and point of view 	<ul style="list-style-type: none"> uses imprecise language and a limited variety of sentence structures may choose an inappropriate style or tone, and may shift point of view 	<ul style="list-style-type: none"> uses simplistic or repetitious language and sentence structures demonstrates little or no understanding of tone or point of view 	<ul style="list-style-type: none"> uses repetitious language and simple sentence structures demonstrates no understanding of style, tone, or point of view <p>OR</p> <ul style="list-style-type: none"> does not respond to prompt

APPENDIX C: PSSA CODING

FOCUS (field)

- 4: Sharp, distinct controlling point **presented** as a position (field)
and **made convincing** through a **clear**, thoughtful, and substantiated argument (field)
with **evident awareness** of task and audience (genre)

CONTRACTp
EXPANDe/CONTRACTp
CONTRACTp

TEXT		APPRAISAL TYPE	GRADUATION	COMMITMENT
APPRAISAL ITEM	APPRAISED ITEM			
sharp	controlling point	C:C	Raise/Sharpen	+3
distinct		V	Raise/Sharpen	
convincing		R:I	Raise/Sharpen	
clear	argument	C:C	Raise/Sharpen	+3
thoughtful		R:Q / V	Raise/Sharpen	
substantiated		R:Q / V	Raise	
evident	awareness of task and audience	C:C	Raise/Sharpen	+1

- 3: **Clear** controlling point **presented** as a position (field)
and **made convincing** through a **credible** and substantiated argument (field)
with general awareness of task and audience (genre)

CONTRACTp/CONTRACTp
EXPANDe/CONTRACTp

TEXT		APPRAISAL TYPE	GRADUATION	COMMITMENT
APPRAISAL ITEM	APPRAISED ITEM			
clear	controlling point	C:C	Raise/Sharpen	+2
convincing		R:I	Raise/Sharpen	
credible	argument	R:Q / V	Raise/Sharpen	+2
substantiated		R:Q / V	Raise	
general	awareness of task and audience	C:C	Lower/Sharpen	+1

- 2: **Vague evidence** of a controlling point **presented** as a position (field)
that **may lack** a **credible** and/or substantiated argument (field)
EXPANDe/CONTRACTp/CONTRACTp
with an inconsistent awareness of task and audience (genre)

CONTRACTp/CONTRACTp

TEXT		APPRAISAL TYPE	GRADUATION	COMMITMENT
APPRAISAL ITEM	APPRAISED ITEM			
vague	evidence of a controlling point	C:C	Lower/Sharpen	+1
(may) lack	a credible and/or substantiated argument	C:B	Lower/Softening	(M) +1
credible	argument	R:Q / V	Raise/Sharpen	+2
substantiated		R:Q / V	Raise	
inconsistent	awareness of task and audience	C:B	Lower/Sharpen	+1

- 1: **Little or no evidence** of a controlling point **presented** as a position (field)
EXPANDe/CONTRACTd/CONTRACTp
that **lacks** a **credible** and/or substantiated argument (field)
with **minimal awareness** of task and audience (genre)

CONTRACTp/CONTRACTp
EXPANDe

TEXT		APPRAISAL TYPE	GRADUATION	COMMITMENT
APPRAISAL ITEM	APPRAISED ITEM			
little no	evidence of a controlling point	C:B C:B	Lower/Soften Lower/Sharpen	+2
lacks	a credible and/or substantiated argument	C:B	Lower/Sharpen	+1
credible substantiated	argument	R:Q / V R:Q / V	Raise/Sharpen Raise/Sharpen	+2
minimal	awareness of task and audience	C:B	Lower/Soften	+1

CONTENT (genre)

- 4: **Substantial**, relevant, and illustrative content (field)
 that **demonstrates** a **clear understanding** of the purpose (field)
 Thoroughly elaborated argument (genre)
 that **includes** a **clear** position (genre)
 consistently supported with precise and relevant evidence (genre)
 Rhetorical (persuasive) strategies **are evident** (genre)

EXPANDe
 CONTRACTp/CONTRACTp
 CONTRACTp/CONTRACTp
 CONTRACTp

TEXT		APPRAISAL TYPE	GRADUATION	COMMITMENT
APPRAISAL ITEM	APPRAISED ITEM			
substantial relevant illustrative	content	C:B C:B C:C	Raise Raise/Sharpen Raise/Sharpen	+3
clear	understanding of the purpose	C:C	Raise/Sharpen	+1
thoroughly elaborated	argument	C:C C:C	Raise/Sharpen Raise/Sharpen	+2
clear	position	C:C	Raise/Sharpen	+1
consistently	supported	C:B	Raise/Sharpen	+1
precise relevant	evidence	C:C C:B	Raise/Sharpen Raise/Sharpen	+2
evident	rhetorical strategies	C:C	Raise/Sharpen	+1

- 3: Adequate, specific, and/or illustrative content (field)
 that **demonstrates** an understanding of the purpose (field)
 Sufficiently elaborated argument (genre)
 that **includes** a **clear** position (genre)
 supported with **some** relevant evidence (genre)
 Rhetorical (persuasive) strategies **may be evident** (genre)

CONTRACTp
 CONTRACTp/CONTRACTp
 EXPANDe
 EXPANDe

TEXT		APPRAISAL TYPE	GRADUATION	COMMITMENT
APPRAISAL ITEM	APPRAISED ITEM			
adequate specific illustrative	content	R:Q C:C C:C	Lower/Sharpen Raise/Sharpen Raise/Sharpen	+3
sufficiently elaborated	argument	R:Q C:C	Lower/Sharpen Raise/Sharpen	+2
clear	position	C:C	Raise/Sharpen	+1
some relevant	evidence	C:C C:B	Lower/Soften Raise/Sharpen	+2
(may be) evident	rhetorical strategies	C:C	Raise/Soften	(M) +1

- 2: Inadequate, vague content (field)
 that **demonstrates** a weak understanding of the purpose (field)
 Insufficiently elaborated argument (genre)
 that **includes** an underdeveloped position (genre)
 supported with **little evidence** (genre)

CONTRACTp

CONTRACTp
EXPANDe

TEXT		APPRAISAL TYPE	GRADUATION	COMMITMENT
APPRAISAL ITEM	APPRAISED ITEM			
inadequate vague	content	R:Q C:C	Lower/Sharpen Lower/Sharpen	+2
weak	understanding of the purpose	R:Q	Lower/Sharpen	+1
insufficiently elaborated	argument	R:Q C:C	Lower/Sharpen Raise/Sharpen	+2
underdeveloped	position	C:C	Lower/Sharpen	+1
little	evidence	C:B	Lower/Soften	+1

- 1: **Minimal evidence** of content (field)
 that **demonstrates a lack of understanding** of the purpose (field)
 Unelaborated argument (genre)
 that **includes** an undeveloped position (genre)
 supported with **minimal** or **no evidence** (genre)

EXPANDe
CONTRACTp/CONTRACTpCONTRACTp
EXPANDe/CONTRACTd

TEXT		APPRAISAL TYPE	GRADUATION	COMMITMENT
APPRAISAL ITEM	APPRAISED ITEM			
minimal	evidence of content	C:B	Lower/Soften	+1
lack	of understanding of the purpose	C:B	Lower/Sharpen	+1
unelaborated	argument	C:C	Lower/Sharpen	+1
undeveloped	position	C:C	Lower/Sharpen	+1
minimal no	evidence	C:B C:B	Lower/Soften Lower/Sharpen	+2

ORGANIZATION (mode)

- 4: **Effective** organizational strategies and structures (mode)
 such as logical order and transitions (mode)
 to develop a position (field)
 supported with a purposeful **presentation** of content (field)

EXPANDe/CONTRACT

CONTRACTp

TEXT		APPRAISAL TYPE	GRADUATION	COMMITMENT
APPRAISAL ITEM	APPRAISED ITEM			
effective	organizational strategies and structures	V	Raise/Sharpen	+1
logical	order	C:B	Raise/Sharpen	+1
purposeful	presentation of content	R:Q / V	Raise/Sharpen	+1

- 3: Organizational strategies and structures (mode)
 such as logical order and transitions (mode)
 to develop a point (field)
 supported with sufficient **presentation** of content (field)

CONTRACTp

TEXT		APPRAISAL TYPE	GRADUATION	COMMITMENT
APPRAISAL ITEM	APPRAISED ITEM			
logical	order	C:B	Raise/Sharpen	+1
sufficient	presentation of content	R:Q	Lower/Sharpen	+1

- 2: Inconsistent organizational strategies and structures (mode)
 such as logical order and transitions (mode)
 to develop a position (field)
 with inadequate **presentation** of content (field)

CONTRACTp

TEXT		APPRAISAL TYPE	GRADUATION	COMMITMENT
APPRAISAL ITEM	APPRAISED ITEM			
inconsistent	organizational strategies and structures	C:B	Lower/Sharpen	+1
logical	order	C:B	Raise/Sharpen	+1
inadequate	presentation of content	R:Q	Lower/Sharpen	+1

- 1: **Little or no evidence** of organizational strategies and structures (mode)
 such as logical order and transitions (mode)
 to develop a position (field)
 with insufficient **presentation** of content (field)

EXPANDe/CONTRACTd

CONTRACTp

TEXT		APPRAISAL TYPE	GRADUATION	COMMITMENT
APPRAISAL ITEM	APPRAISED ITEM			
little	evidence of organizational strategies and structures	C:B	Lower/Soften	+2
no		C:B	Lower/Sharpen	
logical	order	C:B	Raise/Sharpen	+1
insufficient	presentation of content	R:Q	Lower/Sharpen	+1

STYLE (tenor)

- 4: Precise control of language, stylistic techniques, and sentence structures (mode)
 that creates a consistent and effective tone (tenor)

TEXT		APPRAISAL TYPE	GRADUATION	COMMITMENT
APPRAISAL ITEM	APPRAISED ITEM			
precise	control of language, stylistic techniques, and sentence structures	C:C	Raise/Sharpen	+1
consistent effective	tone	C:B V	Raise/Sharpen Raise/Sharpen	+2

- 3: Appropriate control of language, stylistic techniques, and sentence structures
 that creates a consistent tone

TEXT		APPRAISAL TYPE	GRADUATION	COMMITMENT
APPRAISAL ITEM	APPRAISED ITEM			
appropriate	control of language, stylistic techniques, and sentence structures	V	Lower/Sharpen	+1
consistent	tone	C:B	Raise/Sharpen	+1

- 2: **Limited control** of language and sentence structures
 that creates interference with tone

EXPANDe

TEXT		APPRAISAL TYPE	GRADUATION	COMMITMENT
APPRAISAL ITEM	APPRAISED ITEM			
limited	control of language, stylistic techniques, and sentence structures	C:B	Lower/Soften	+1
interference	tone	C:B	Lower	+1

1: **Minimal control** of language and sentence structures
that creates an inconsistent tone

EXPANDe

TEXT		APPRAISAL TYPE	GRADUATION	COMMITMENT
APPRAISAL ITEM	APPRAISED ITEM			
minimal	control of language, stylistic techniques, and sentence structures	C:B	Lower/Soften	+1
inconsistent	tone	C:B	Lower/Sharpen	+1

APPENDIX D: KE CODING

THESIS/FOCUS (field)

- 4: **Establishes** and **sustains** a precise claim or position
Displays a **clear** understanding of task, purpose, and audience

CONTRACTp/CONTRACTp
 CONTRACTp/CONTRACTp

TEXT		APPRAISAL TYPE	GRADUATION	COMMITMENT
APPRAISAL ITEM	APPRAISED ITEM			
precise	claim or position	C:C	Raise/Sharpen	+1

- 3: **Establishes** a claim or position
Displays an understanding of task, purpose, and audience

CONTRACTp
 CONTRACTp

TEXT		APPRAISAL TYPE	GRADUATION	COMMITMENT
APPRAISAL ITEM	APPRAISED ITEM			
	claim or position			+0

- 2: **Provides** an inconsistent claim or position
Displays a **limited** understanding of task, purpose, and audience

CONTRACTp
 CONTRACTp/EXPANDe

TEXT		APPRAISAL TYPE	GRADUATION	COMMITMENT
APPRAISAL ITEM	APPRAISED ITEM			
inconsistent	claim or position	C:B	Lower/Sharpen	+1

- 1: **Provides** **vague** or indistinct claim or position
Displays a **minimal** understanding of task, purpose, and audience

CONTRACTp/CONTRACTp
 CONTRACTp/EXPANDe

TEXT		APPRAISAL TYPE	GRADUATION	COMMITMENT
APPRAISAL ITEM	APPRAISED ITEM			
vague	claim or position	C:C	Lower/Sharpen	+2
indistinct		V	Lower/Sharpen	

- 0: **Provides** **no** evidence of claim or position
Displays **no** understanding of task, purpose or audience
 Or **does not** respond to prompt

CONTRACTp/CONTRACTd
 CONTRACTp/CONTRACTd
 CONTRACTd

TEXT		APPRAISAL TYPE	GRADUATION	COMMITMENT
APPRAISAL ITEM	APPRAISED ITEM			
no	evidence of claim or position	C:B	Lower/Sharpen	+1

CONTENT (genre)

- 4: **Provides** relevant content and specific and effective supporting details
 that **demonstrate** a **clear** understanding of purpose
Uses sophisticated transitional words, phrases, and clauses
 to link ideas and create cohesion
Considers **possible** counterclaims (alternate or opposing arguments)

CONTRACTp
 CONTRACTp/CONTRACTp
 CONTRACTp
 EXPANDa/EXPANDe

TEXT		APPRAISAL TYPE	GRADUATION	COMMITMENT
APPRAISAL ITEM	APPRAISED ITEM			
relevant	content	C:B	Raise/Sharpen	+1
specific	supporting details	C:C	Raise/Sharpen	+2
effective		V	Raise/Sharpen	
clear	understanding of purpose	C:C	Raise/Sharpen	+1
sophisticated	words, phrases, and clauses	V	Raise/Sharpen	+1

- 3: **Provides** relevant content and effective supporting details
Uses transitional words, phrases, and clauses to link ideas
Acknowledges possible counterclaims (alternate or opposing arguments)

CONTRACTp
 CONTRACTp
 EXPANDa/EXPANDe

TEXT		APPRAISAL TYPE	GRADUATION	COMMITMENT
APPRAISAL ITEM	APPRAISED ITEM			
relevant	content	C:B	Raise/Sharpen	+1
effective	supporting details	V	Raise/Sharpen	+1

- 2: **Provides** insufficient content and ineffective supporting details
May use simplistic and/or illogical transitional expressions
May not acknowledge possible counterclaims (alternate or opposing arguments)

CONTRACTp
 EXPANDe
 EXPANDe/EXPANDa/EXPANDe

TEXT		APPRAISAL TYPE	GRADUATION	COMMITMENT
APPRAISAL ITEM	APPRAISED ITEM			
insufficient	content	R:Q	Lower/Sharpen	+1
ineffective	supporting details	V	Lower/Sharpen	+1
(may) simplistic	transitional expressions	C:C	Lower/Sharpen	+2
(may) illogical		C:B	Lower/Sharpen	

- 1: **Provides minimal** content
Uses few or no transitional expressions to link ideas
Does not acknowledge possible counterclaims (alternate or opposing arguments)

CONTRACTp/EXPANDe
 EXPANDe/CONTRACTd
 CONTRACTd/EXPANDe

TEXT		APPRAISAL TYPE	GRADUATION	COMMITMENT
APPRAISAL ITEM	APPRAISED ITEM			
minimal	content	C:B	Lower/Soften	+1
few	transitional expressions	C:B	Lower/Soften	+2
no		C:B	Lower/Sharpen	

- 0: **Provides little to no** content
Does not use transitions to link ideas
Or does not respond to prompt

CONTRACTp/EXPANDe/CONTRACTd
 CONTRACTd
 CONTRACTd

TEXT		APPRAISAL TYPE	GRADUATION	COMMITMENT
APPRAISAL ITEM	APPRAISED ITEM			
little	content	C:B	Lower/Soften	+2
no		C:B	Lower/Sharpen	
does not use	transitions	C:B	Lower/Sharpen	+1

ORGANIZATION (mode)

- 4: Chooses sophisticated organizational strategies appropriate for task, purpose, and audience

Presents fair and relevant evidence

CONTRACTp

to support claim or position

Includes a clear and well-defined introduction, body, and conclusion that support or reinforce the argument

CONTRACTp/CONTRACTp

TEXT		APPRAISAL TYPE	GRADUATION	COMMITMENT
APPRAISAL ITEM	APPRAISED ITEM			
sophisticated	organizational strategies	V	Raise/Sharpen	+2
appropriate		V	Raise/Sharpen	
fair	evidence	R:Q/V	Raise/Sharpen	+2
relevant		C:B	Raise/Sharpen	
clear	introduction, body, and conclusion	C:C	Raise/Sharpen	+2
well-defined		C:C	Raise/Sharpen	

- 3: Chooses appropriate organizational strategies for task, purpose, and audience

Presents relevant evidence to support claim or position

CONTRACTp

Includes a clear introduction, body, and conclusion that support the argument

CONTRACTp/CONTRACTp

TEXT		APPRAISAL TYPE	GRADUATION	COMMITMENT
APPRAISAL ITEM	APPRAISED ITEM			
appropriate	organizational strategies	V	Raise/Sharpen	+1
relevant	evidence	C:B	Raise/Sharpen	+1
clear	introduction, body, and conclusion	C:C	Raise/Sharpen	+1

- 2: **Displays some** evidence of organizational strategies
Present insufficient evidence to support claim or position
May not include an introduction, body, and/or conclusion

CONTRACTp/EXPANDe

CONTRACTp

EXPANDe

TEXT		APPRAISAL TYPE	GRADUATION	COMMITMENT
APPRAISAL ITEM	APPRAISED ITEM			
some	evidence of organizational strategies	C:B	Lower/Softening	+1
insufficient	evidence	R:Q	Lower/Sharpen	+1
(may not) include	introduction, body, and conclusion	C:B	Lower/Softening	+M

- 1: **Displays little** evidence of organizational strategies
Presents little or no evidence to support claim or position
May not include an identifiable introduction, body, and/or conclusion

CONTRACTp/EXPANDe

CONTRACTp/EXPANDe/CONTRACTd

EXPANDe

TEXT		APPRAISAL TYPE	GRADUATION	COMMITMENT
APPRAISAL ITEM	APPRAISED ITEM			
little	evidence of organizational strategies	C:B	Lower/Softening	+1
little	evidence	C:B	Lower/Softening	+2
no		C:B	Lower/Sharpen	
(may not) include an identifiable	introduction, body, and conclusion	C:C	Raise/Sharpen	M +1

- 0: **Displays no** evidence of organizational strategies
Presents no evidence to support claim or position
Does not include an identifiable introduction, body, and/or conclusion
Or **does not** respond to prompt

CONTRACTd
CONTRACTd
CONTRACTd
CONTRACTd

TEXT		APPRAISAL TYPE	GRADUATION	COMMITMENT
APPRAISAL ITEM	APPRAISED ITEM			
no	evidence of organizational strategies	C:B	Lower/Sharpen	+1
no	evidence	C:B	Lower/Sharpen	+1
(does not) include	introduction, body, and conclusion	C:B	Lower/Sharpen	+1

STYLE (tenor)

- 4: **Uses** consistently precise language
and a wide variety of sentence structures
Chooses an effective style and tone
and **maintains** a consistent point of view

CONTRACTp

CONTRACTp

TEXT		APPRAISAL TYPE	GRADUATION	COMMITMENT
APPRAISAL ITEM	APPRAISED ITEM			
consistently	language	C:B	Raise/Sharpen	+2
precise		C:C	Raise/Sharpen	
wide	sentence structures	C:B	Raise/Soften	+2
variety		C:C	Raise/Soften	
effective	style and tone	V	Raise/Sharpen	+1
consistent	point of view	C:B	Raise/Sharpen	+1

- 3: **Uses** precise language
and a variety of sentence structures
Chooses an appropriate style, tone, and point of view

CONTRACTp

TEXT		APPRAISAL TYPE	GRADUATION	COMMITMENT
APPRAISAL ITEM	APPRAISED ITEM			
precise	language	C:C	Raise/Sharpen	+1
variety	sentence structures	C:C	Raise/Soften	+1
appropriate	style, tone, and point of view	V	Raise/Sharpen	+1

- 2: **Uses** imprecise language
and a **limited** variety of sentence structures
May choose an inappropriate style or tone,
and **may** shift point of view

CONTRACTp
EXPANDe
EXPANDe
EXPANDe

TEXT		APPRAISAL TYPE	GRADUATION	COMMITMENT
APPRAISAL ITEM	APPRAISED ITEM			
imprecise	language	C:C	Lower/Sharpen	+1
limited	sentence structures	C:B	Lower/Soften	+2
variety		C:C	Raise/Soften	
(may choose) inappropriate	style or tone	V	Lower/Sharpen	M+1
(may shift)	point of view	C:B	Lower/Soften	M+1

- 1: **Uses** simplistic or repetitious language and sentence structures
Demonstrates little or **no** understanding of tone or point of view

CONTRACTp
CONTRACTp/EXPANDe/CONTRACTd

TEXT		APPRAISAL TYPE	GRADUATION	COMMITMENT
APPRAISAL ITEM	APPRAISED ITEM			
simplistic	language and sentence structures	C:C	Lower/Sharpen	+2
repetitious		C:B	Lower/Sharpen	
little understanding	tone or point of view	C:B	Lower/Soften	+2
no understanding		C:B	Lower/Sharpen	

0: Uses repetitious language

CONTRACT_p

and simple sentence structures

Demonstrates no understanding of style, tone, or point of view

CONTRACT_d

Or **does not** respond to prompt

CONTRACT_d

TEXT		APPRAISAL TYPE	GRADUATION	COMMITMENT
APPRAISAL ITEM	APPRAISED ITEM			
repetitious	language	C:B	Lower/Sharpen	+2
simple	sentence structures	C:C	Lower/Sharpen	
no understanding	tone or point of view	C:B	Lower/Sharpen	+1