

Research Memorandum ETS RM-11-01

The CBAL Summative Writing Assessment: A Draft Eighth-Grade Design

Paul Deane

Mary Fowles

Douglas Baldwin

Hilary Persky

March 2011

The CBAL Summative Writing Assessment: A Draft Eighth-Grade Design

Paul Deane, Mary Fowles, Douglas Baldwin, and Hilary Persky
ETS, Princeton, New Jersey

As part of its nonprofit mission, ETS conducts and disseminates the results of research to advance quality and equity in education and assessment for the benefit of ETS's constituents and the field.

To obtain a PDF or a print copy of a report, please visit:

http://www.ets.org/research/contact.html

Technical Review Editor: Joel Tetreault

Technical Reviewers: Drew Gitomer and Isaac Bejar

Copyright © 2011 by Educational Testing Service. All rights reserved.

ETS, the ETS logo, GRE, and LISTENING. LEARNING. LEADING. are registered trademarks of Educational Testing Service (ETS).



Abstract

This memorandum describes the process and results of developing draft summative writing assessments within the CBAL Research Initiative (Cognitively Based Assessments of, for, and as Writing). It outlines four designs—each focused on a different skill—at the 8th grade level, reviews in detail the content of each design, and presents the motivation for the design decisions. It also briefly sketches initial results from preliminary pilots, where available, and connects each design to the 2008 CBAL Writing Competency model.

Key words: CBAL, summative writing assessment, writing competency model, six-trait scoring

Table of Contents

	Page
1. Background	1
1.1. Leading Ideas From the Literature Review	2
Interference effects	3
Lack of strategic writing skills	3
Insufficient topic-specific knowledge	4
Weak content reasoning and research skills	4
Unformed or rudimentary rhetorical goals	5
Managing the transition to knowledge-transforming strategies	6
1.2. Leading Ideas From the Pedagogical Literature	6
1.3. Developments in 2008	9
2. Revisions to the Competency Model	10
2.1. The Competency Model at the End of 2007	10
2.2. The Revised Competency Model	14
Explanation of the writing competency model	14
Distinctives.	22
3. Creating Test Specifications	23
3.1. General Test Specification	24
Individual forms	25
Each year's sequence of periodic accountability assessments	26
3.2. Eighth Grade Test Specification	26
The annual sequence.	26
Test specification #1: Expository/review	27
Discussion	28
Test specification #2: Persuasive memorandum	31
Discussion	31
Possible revisions	32
Test specification #3: Expository/report.	35
Possible revisions	36
4. General Discussion and Conclusions	39

4.1. The Test Design	39
4.2. Scoring and Dimensionality	40
4.3. Use of Behavior Features	42
5. Conclusions and Future Directions	42
References	44
Notes	51
List of Appendices	

List of Tables

Page
Table 1 - How the Six-Trait Scoring Model Maps Onto the CBAL Competency Model22
Table 2 - Specification for a 4-PAA Sequence of Eighth-grade Writing Assessments27
Table 3 - Specification for a PAA Focused on Reviewing Literary Works
Table 4 - Specification for a PAA Focused on Writing a Persuasive Memorandum31
Table 5 - Graphic Organizer Design for a Scaffolding Critical-reading Question34
Table 6 - Specification for a PAA Focused on Writing an Expository Report35
Table 7 - Specification for a PAA Focused on Writing a Persuasive Essay
Table 8 - Adjusted R2 for Linear Regression Models Predicting Human Scores on the
Fall 2007 Pilots41

1. Background

This paper presents the results of research on the summative writing assessment being designed as part of the ETS research initiative, Cognitively Based Assessments of, as, and for Learning (CBAL). It builds upon the accomplishments of research on writing assessment conducted in 2006–2007, work summarized in Deane et al. (2007, 2008). Among the central goals of this research are the following principles (see Appendix F for the full set of CBAL guiding principles).

Accountability tests, formative assessments, and professional support will be derived from the same conceptual base. That base will be built upon cognitive research, state standards, and curricular considerations. As such, accountability testing, formative assessment, and professional development will reinforce one another because task formats and content will be more consistent with conceptions of what students should know and be able to do than in current assessment systems, which are typically built around multiple-choice questions.

Because of their integrative, foundational, and engaging nature, CBAL accountability tasks should be viewed by teachers and students as worthwhile learning experiences in and of themselves. Preparing for the test should have the effect of improving student learning, not just enhancing performance on the test. CBAL assessments should help students view assessment positively, as a part of the learning process.

In addition, the design of the assessment is intended to conform to the principles of Evidence-Centered Design (Mislevy et al. 2003), and thus the test creation process has been structured so as to build a strong argument that the test collects the right evidence with which to assess the intended construct. In order to achieve these goals, we have sought to design a test that is, simultaneously, responsive to the literature on writing cognition and to the literature on best practices in language arts and writing instruction. To that end, we have worked with language arts teachers and other educational professionals not only to conduct initial pilots but to use their feedback to make the test design more effectively support student learning.

This report serves two purposes. On one hand, it provides a comprehensive overview of the CBAL Writing Assessment design. On the other hand, it documents progress during 2008, and will therefore outline where thinking has evolved as a result of an intensive process of

of piloting materials, working with teachers, scoring, and reviewing student responses. The structure naturally falls into two major sections: The first is comprised of sections 1–4, which outline the competency model, evidence model, and our final test blueprints for the year's worth of tests developed by the end of 2008. With each blueprint we will discuss the results of the pilots for each test, including review of scoring, followed by an explanation of why we chose the design modifications reflected in the blueprints. After this section, a series of appendices document in detail the revised evidence model based upon the new competency model and the summative tests piloted thus far, including screen shots, rubrics, and benchmark papers.

1.1 Leading Ideas From the Literature Review

Perhaps the most important conclusion reached in the literature review (Deane et al., 2008) is the importance to be attached to critical thinking, both with respect to writing as an act of communication, and with respect to the content being communicated. This conclusion is amply supported by a variety of facts drawn from the literature on the nature of writing expertise and the difference between novice and skilled writers. In particular, skilled writers spend more time planning and revising their work than novice writers; they focus more of their effort and attention on managing the development of content, and concern themselves less with its formal, surface characteristics; and they employ a variety of self-regulatory strategies (Bereiter & Scardamalia, 1987; Galbraith, 1999; Graham, 1997; Graham & Harris, 2000; Kellogg, 1988, McCutchen, Francis, & Kerr 1997; McCutchen, 2000). Moreover, novice writers benefit from instruction on planning and revision strategies and from being helped to think about topic-relevant content (De La Paz, 2005; De La Paz & Graham, 1997a, 1997b, 2002; Graham & Perin, 2006; Hillocks, 1987; Kellogg, 1988; Quinlan, 2004).

Bereiter and Scardamalia (1987) characterize the difference between novice and skilled authors as the difference between a *knowledge-telling* approach and a *knowledge-transforming* approach to writing. In a knowledge-telling approach, the focus of the writer's effort is on the process of putting words on the page. Whatever ideas the author is able to mobilize are assumed to be good enough; writing takes place as a direct translation of those ideas into words, and as soon as the words are on the page, the writer is finished. In a knowledge-transforming approach, writing is a recursive process of knowledge-development and knowledge-expression. Planning is more than organizing existing ideas; it is an active

process of questioning, research, and rethinking. When text is produced, it is not viewed as the final product, but is subjected to systematic evaluation, revision, reworking in the light of a critical evaluation both of the content being communicated and its effectiveness in advancing the author's rhetorical goals.

Knowledge-transforming is by its nature a much more effortful and sophisticated process than knowledge-telling, and so it is not particularly surprising that novice writers default to a knowledge-telling approach. It is, however, useful to consider in greater detail why authors might fail to use a knowledge-transforming approach to writing, as these reasons themselves suggest instructional strategies. The literature suggests five categories of explanation:

- Interference effects (undeveloped or inefficient literacy skills)
- Lack of strategic writing skills
- Insufficient topic-specific knowledge
- Weak content reasoning and research skills
- Unformed or rudimentary rhetorical goals.

Interference effects. Writing processes compete in working memory. The high-level, strategic skills required for a knowledge-transforming approach to writing place heavy demands on memory and attention. In many novice writers, absence or more likely, inefficiency of lower-level skills such as oral fluency, transcription, and text decoding (reading) makes it impossible to free up the working memory capacity needed for strategic thought (Kellogg, 2001; Olive & Kellogg, 2002; Pearl, 1979; Piolat, Roussey, Olive, & Farioli, 1996; Torrance & Galbraith, 2005). The ability to produce text fluently and easily depends both upon oral fluency (Shanahan, 2006) and upon basic transcription abilities (Bourdin & Fayol, 1994, 2000), and thus can become slow and effortful if any of these component skills function inefficiently. Similarly, the ability to monitor and reflect upon one's own writing, which is critical to planning and revision, depends in large part upon aspects of reading skill, both decoding and higher verbal comprehension, and thus reading difficulties can compromise revision and planning (Hayes, 1996, 2004; Kaufer, Hayes, & Flower, 1986; McCutchen et al., 1997).

Lack of strategic writing skills. Even skilled writers can be limited by working memory capacity, so that they cannot handle all aspects of the writing task simultaneously. A

significant element in writing skill is the ability to intersperse planning, text production, and evaluation, sometimes switching back and forth rapidly among tasks, and other times, devoting significant blocks of time to a single activity (Matsuhashi, 1981; Schilperoord, 2002). Controlling writing processes so that the choice of activities is strategically appropriate and maximally efficient is itself a skill, one that takes time to acquire, and novice writers typically do not manage it well (cf. Coirier, Anderson, & Chanquoy, 1999, for an application of these ideas to persuasive writing).

Insufficient topic-specific knowledge. Knowledge of the topic about which one has to write is a critical determinant of writing success. All writing models presuppose a critical role for long-term memory in which the subject matter of writing must be retrieved, either in a top-down fashion (Bereiter & Scardamalia, 1987; Hayes & Flower, 1980) or in a more bottom-up manner (Galbraith, 1999, Galbraith & Torrance, 1999). Those who already possess the knowledge needed to write about a subject are at an advantage. Moreover, the kinds of critical thinking needed to pursue a knowledge-transforming approach to writing arguably require at least basic topic knowledge to support judgments of relevance and plausibility and to support reasoning about content. Thus it is not surprising that topic knowledge is a major predictor of writing quality (DeGroff, 1987; Langer, 1985; McCutchen, 1986).

Weak content reasoning and research skills. The essence of the knowledge-transforming approach is that writing is not viewed as a mere expressive act, but as part and parcel of a habit of critical thinking in which the act of writing serves as the occasion for, and the focus of, a complex form of problem-solving. While many of the problems the writer faces are rhetorical, having to do with audience and purpose, these goals typically require the author to develop ideas, to identify information needed (but not possessed), and to obtain that information, whether by observation, inference, argument, or through research. These skills are arguably among the most important skills needed for academic writing (cf. Hillocks, 1987, meta-analysis, which indicated the critical importance of inquiry strategies to improve student writing, and the related arguments in Hillocks, 1995).

Complicating the picture is the fact that the reasoning required successfully to complete a writing task varies with purpose, audience, and genre. In a narrative writing task, for instance, reasoning about actions and motives is likely to be relatively important, whereas an expository writing task is more likely to place an emphasis on such skills as definition,

generalization, and analogy, and a persuasive argumentation task would emphasize evidence and refutation. There is thus a wide collection of content reasoning skills needed for effective writing, and it is entirely possible that individuals will be strong on some of these skills and weak on others, though overall the evidence is that one cannot assume that novice writers, or even all adults, possess more than a fraction of the requisite content reasoning and research skills needed to support a knowledge-transforming approach to writing (Felton & Kuhn 2001, Kuhn, 1991; Kuhn, Katz, & Dean, 2004; Means & Voss, 1996; Perkins, 1985; Perkins, Allen, & Hafner, 1983).

Unformed or rudimentary rhetorical goals. All of the issues surveyed thus far depend upon the setting of appropriate rhetorical goals. The key weakness of the knowledge-telling approach to writing is that it effectively assumes a single goal for writing: the expression of existing knowledge, modified minimally to suit the task. A sophisticated writer must be aware that all writing is communication within a social context, in which the author must take the audience into account, collaborate with others, and, more generally, act within one or more communities of practice with well-defined expectations about the role writing fills within each community's sphere of action.

There is evidence that students benefit from instructional activity that clarifies the intended audience and makes the writer's obligations to that audience clearer (Cohen & Riel, 1989; Daiute, 1986; Daiute & Dalton, 1993; Yarrow & Topping, 2001). In particular, activities which make the act of writing more social and interactive, such as peer review, appear to have a strong beneficial impact on student writing (Graham & Perin, 2006).

Just as critically, each act of writing and each mode and genre of writing operate within a system of social norms and expectations. There is clear evidence that students benefit when writing instruction is structured to enable students to internalize these social norms (Flower, 1989; Kent, 2003; Kostouli, 2005), and this idea can usefully be extended by viewing the writing classroom as functioning best when it is explicitly designed to enculturate students to participate in academic and other writing communities of practice (Beaufort, 2000).

The key point is that skilled writers have the necessary knowledge and community connections to set appropriate discourse goals. Such goals will be salient for students only to the extent that the community and the audience is made real and present within their writing

experiences, and will be practicable only to the extent that they have acquired the cognitive capacities or skills needed to achieve them.

Managing the transition to knowledge-transforming strategies. The considerations adduced thus far imply, collectively, that the goal of writing instruction is to enable novice writers to surmount the barriers that prevent them from adopting a knowledge-transforming approach, and to provide them with what they need to learn how to set appropriate rhetorical goals, reason appropriately about content, and manage their writing activities efficiently, with minimal problems due to inefficiencies in underlying component skills. In effect, the purpose of writing instruction is to manage the transition from knowledge-telling to knowledgetransforming. The challenge, in turn, is how to structure tests in such a way as to encourage appropriate strategies—strategies that enable students to overcome weaknesses in component skills, that encourage appropriate focus of attention on a more complex writing process, that strengthen topical knowledge and topic-appropriate content reasoning, and that encourage students to acquire the skills needed to set appropriate rhetorical goals. These goals inform our design, since they imply (among other things) that the tasks on a writing test should require meaningful engagement with writing tasks that genuinely require critical thinking, but that they should also provide enough information about the topic to enable weaker students to find something to say about it.

1.2 Leading Ideas From the Pedagogical Literature

The pedagogical literature also provides useful insights. We begin with the assumption that the structure of a test, by modeling desirable traits, may help to encourage desirable teaching practices. If so, we can draw on best practices for ideas, structuring the test to reflect the behaviors we wish to model. Viewed in this light, the literature strongly supports an approach in which tasks are placed in meaningful contexts, sequenced in functionally related sequences, and clearly related to activities students can expect to have to deal with both in school and in real life.

Let us start with general overviews of teaching methods, such as the approach to formative assessment presented by Leahy, Lyon, Thompson, and Wiliam (2005). According to their framework, effective teachers

 clarify and share what students are supposed to learn and the standards by which students will be judged,

- engineer classroom discussions, tasks, and activities that provide clear evidence of student learning,
- provide feedback that focuses students on what they need to learn,
- encourage students to take an active role as learners, and
- promote a classroom culture in which students collaborate and actively support one another's learning.

Similar themes emerge from Langer's (2001) analysis of teaching practices that produce more effective performance in the language arts. According to Langer's analysis, the most effective language arts programs were those in which

- skills and knowledge are taught in multiple types of lessons
- teachers make connections across lessons, between courses, with life outside the classroom
- students learn strategies for doing the work
- students are expected to be generative thinkers
- classrooms foster collaborations that deepen student thinking
- tests are deconstructed to inform curriculum and instruction

The shared emphases include making standards explicit; encouraging meaningful connections among ideas; and relying on a situated, collaborative approach to learning.

Graham and Perin (2007) present a meta-analysis of best practices in writing instruction, which can readily be interpreted in terms of Langer's categories, along the following lines:

- Multiple types of lessons, teaching component skills through methods such as sentence combining, while taking advantage of word processing and other supporting technologies
- Integrated assessment, providing clear rubrics for quality writing and models of good writing

- Integrated instruction connecting writing to meaningful goals in classroom and beyond, through the use of prewriting, inquiry activities, and writing across the curriculum
- Explicit **strategy instruction** not only in general writing processes but in specific, critical writing tasks such as **summarization**
- Fostering collaborative writing by adopting a process writing approach and structuring classroom activities to make writing a natural, shared task for achieving a variety of instructional goals.

These considerations, in combination with those discussed in section 1.1, led to the following general guidelines (as expressed in the 2007 CBAL project report, Deane et al., 2007):

- Each test typically (but not always) focuses on one genre or mode of discourse and the critical thinking/writing strategies associated with that mode of discourse.
- Each test presents a multi-part "project" and is structured around a scenario or situation that provides a context and purpose for a series of related tasks.
- Usually two or three short tasks precede a longer, more integrated writing task; sometimes a short follow-up task comes after the long task. The short tasks serve two purposes: (a) to engage the student in the kinds of thinking necessary to accomplish the long task, thereby providing a type of scaffolding, and (b) to assess component skills. Most of the tasks are constructed-response, with selected-response tasks used judiciously and, it is hoped, imaginatively (e.g., to assess understanding of the principles involved in adapting texts to a particular kind of audience or to assess the ability to apply writing guidelines to a particular type of text).

The literature review and competency model clearly acknowledge the importance of background knowledge in writing proficiency. In order to avoid measuring such knowledge about a topic, instead of actual writing and thinking skills, many of the tests include resource documents upon which the students are required or encouraged to draw. This approach also permits assessment of research skills

crucial to academic writing. The argument for this structure is that it captures the emphasis on critical thinking identified in the literature review and provides a structure that translates readily into a more formative context. The scenario structure allows all tasks to be set in a context that makes their intended relations clear; the acceptance of source materials makes the tasks more meaningful and more like the sorts of tasks students might actually have to carry out in classrooms or in later life. At the end of 2007, however, the design of the CBAL Writing Assessment was still very much open to discussion. A first draft of the competency model and evidence model had been developed, and draft assessments had been created to instantiate them, but it was not yet possible to convert the project's leading ideas into a test blueprint.

1.3. Developments in 2008

Work in 2008 focused on a recursive process in which we designed tests, worked with teachers to make classroom use of the task structures embodied in those tests, and administered the materials thus developed, in small pilot studies. We used this recursive process to revise existing elements of the test design (the competency model and the evidence model) and to formulate a more detailed test specification based upon our experience with pilot tests. The rest of this report will focus successively on every element of the test design, providing a full description and motivation for each. The following major topics will be addressed:

- 1. Revisions to the competency model.
- 2. Development of a blueprint for a year's sequence focused on the 8th grade, along with general specifications for CBAL writing assessments.
- A detailed presentation of the tests incorporated in that year's blueprint (which
 includes tests piloted late in 2007, but analyzed in 2008). This presentation
 includes test design, test presentation, and results and analysis from the pilot
 administrations.
- 4. Revisions to the evidence model, taking into account analysis of the pilots and related work on automated scoring, detailed in a separate report. The evidence model revisions are presented in Appendix A, as they are quite lengthy.

2. Revisions to the Competency Model

The competency model for the CBAL Writing Assessment has to satisfy several constraints. First, it functions as the specification for the student model within an evidence-centered design approach to test development. That is, it specifies exactly what construct is being tested and therefore reported. Second, it reflects (and links back to) cognitive models and research. That is, the competency model is not a cognitive model, but the skills it measures are cognitively defined and can be related to specific parts of a cognitive model of writing skill. Third, it communicates instructional objectives to teachers, since it defines the skills students are expected to master. At the end of 2007, the CBAL Writing competency model had been extensively motivated on construct and cognitive grounds (Deane et al., 2007, 2008), but still needed to be grounded pedagogically. Thus, most of the revisions to the competency model to be outlined below are motivated either by conceptual issues that came to the fore when the competency model was presented to teachers during 2008, or by the need to adapt the model to make it clearer and a more useful guide for teachers.

2.1 The Competency Model at the End Of 2007

At the end of 2007, the competency model was represented, at the highest level of abstraction, by the diagram in Figure 1. While the essential content of the model was well-justified, several problems emerged that need to be addressed:

- Clarity of presentation. Many of the node labels proved to be difficult for teachers to grasp easily, as evidenced by the interpretations they placed on them when the competency model had to be presented in summary form.
- **Emphasis.** The fact that critical thinking skills were placed at the bottom of the model diagram was sometimes misinterpreted by teachers as putting too much emphasis on language skills, since those were presented first.
- **Handling of prerequisite skills.** Some of the skills in the competency model, e.g., the nodes *Speak/Understand*, *Read/Decode*, and *Inscribe (Handwriting/*

Keyboarding) have a fundamentally different status from the other nodes in the model. They represent true prerequisite skills—skills that must be mastered if the writing construct is to be measured meaningfully. Their inclusion under the *Use Language and Literacy Skills* node was somewhat confusing, since they do not represent skills one would actually design the writing test to measure.

- The role of Standard Written English. Another key issue involved the conceptualization of language and literacy skills. A developmental analysis of automated scoring features for CBAL writing (see Deane & Quinlan, 2010) indicated a three-construct model, along the following lines:
 - Adherence to Conventions: Presence or absence of grammar, usage,
 spelling, and mechanics errors
 - o *Communicating in a Formal, Written English Style:* control of relatively marked, written vocabulary; grammatical richness (syntactic variety), and appropriate use of marked stylistic features such as the passive
 - Fluency in Producing Structured Texts: ability to produce longer documents, with internal structure (paragraphs, a hierarchical outline) where each segment is well-developed

The evidence for these constructs suggested a reorganization that would line up the natural categories for automatic scoring more closely with the competency model. On this view, fluency in producing structured texts is part of the second layer of the competency model (*Use Strategies to Manage the Writing Process*). Adherence to conventions, and being able to communicate in a formal, written English style, belong to the first layer (*Use Language and Literacy Skills*). These considerations imply a reorganization of the competency model.

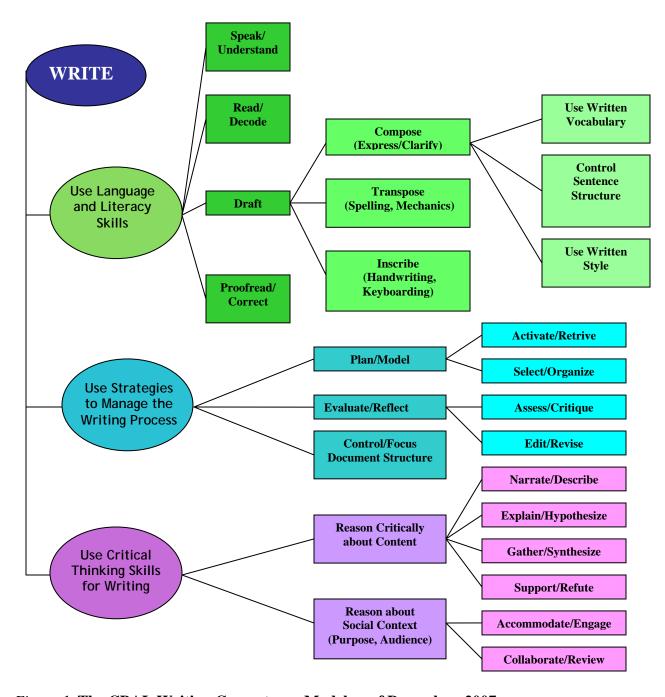


Figure 1. The CBAL Writing Competency Model as of December, 2007.

- that proved to be problematic is the way it handled genre categories like narrative, persuasive and expository writing. The competency model preserved a one-to-one mapping to genres in its critical thinking section. That is, narrative texts mapped to *Narrate/Describe*, expository texts primarily to *Explain/Hypothesize*, and persuasive texts to *Support/Refute*. But by design these nodes in the model were intended to capture the underlying critical thinking skills most critically deployed by each genre. They were not intended to represent the ability to write in each genre. This subtlety proved difficult to communicate either to teachers or to test developers, in part because genre represents a confluence across multiple levels of the competency model, subsuming a purpose for writing, a set of key critical thinking skills needed to achieve that purpose, a family of organizational plans, and a typical stylistic profile.
- The relationship between the critical thinking nodes and received categories such as Bloom's taxonomy. A number of teachers, when presented with the competency model, commented that it covered much of the same ground as Bloom's taxonomy of educational objectives (Bloom, 1956; Anderson, & Krathwohl, 2000). While it is important that the categories of the competency model retain their connection to modern cognitive theories, such as Kuhn (1991)'s theory of argumentation, it is also important that the categories should be transparent to teachers, allowing them to grasp exactly what is intended by each node of the model as rapidly as possible.
- The role of social reasoning. The competency model is intended to capture the very important role that social reasoning plays, both in the act of communication (where it figures in the understanding of social roles, purposes for writing, and modeling audience reactions) and in the content of text, particularly, but not exclusively, in narrative text, where literary concepts like narrator and protagonist presuppose a critical understanding of how people perceive and interact. In its 2007 form, certain key concepts, such as perspective, are not highlighted, and nothing highlights the strong connection between social reasoning about content and social reasoning about the communicative act itself.

• The role of revision. The competency model in its current version does not highlight revision as a key element in learning to write well. While it is arguably a complex composite involving evaluation and revision strategies, further review and feedback during 2008 indicated that the model needed to capture more accurately the role of revision in learning how to write.

2.2 The Revised Competency Model

These concerns, in combination, motivated a significant rewriting of the competency model designed to make it clearer, cleaner, and more obviously motivated. Figure 2 presents this revised model.

Explanation of the Writing Competency Model. The sections that follow present short explanations of each node in the model. Footnotes are included that explicate how each node links to the literature review and other supporting research.

- The Fundamental Construct. This model is concerned with the ability to produce and revise thoughtful, well-structured documents in formal written English. It primarily addresses that ability as a vehicle for critical thinking skills that enable writers to address their audience perceptively and engage thoughtfully with the subject matter they cover, thus achieving achieve well-defined rhetorical goals.²
- Use Critical Thinking Skills for Writing. Writing is a complex communicative process in which the writer addresses other people with a message. Effective writing requires that the writer think effectively and critically about the people addressed (the audience, and other relevant individuals or social groups) and about the content to be addressed.
- Reason Critically About Social Context (Perspective, Purpose, Audience). Students should not write in a vacuum; they should possess the social reasoning skills needed to support effective writing, which always occurs in a social context, always involves an audience whose perspective may differ from the authors, and whose content often addresses people and their interactions. They need to be able to understand their purpose in writing in the light of the full social dynamic involved in effective communication.

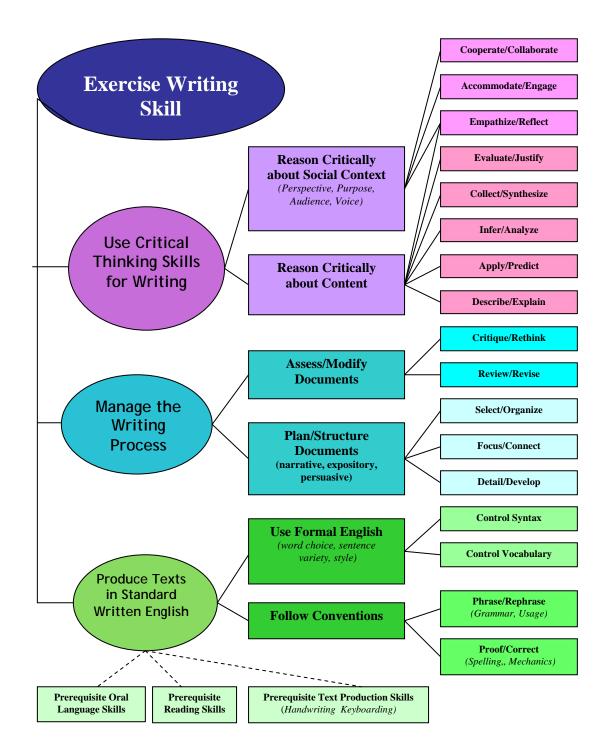


Figure 2. A revised competency model for CBAL Writing

- Cooperate/Collaborate. Writing does not occur in a vacuum, but in a social context in which people may sometimes function as authors, sometimes as readers, sometimes as editors, reviewers, or copyeditors, among others. Students should be able to function effectively in each of the major roles implicit in modern writing practices, and thus should be sensitive to and reasonably proficient in fulfilling the requirements of each role.³
- Accommodate/Engage. The communicative act is itself a transaction between the author and a (real or potential) audience, in which the author critically needs to predict and control the way an audience will respond. To the extent that authors are aware of, and responsive to, their audience, the texts they produce will be adjusted to accommodate the audience's perspective and to engage the audience's interest. Shifts in the audience or purpose of a text should be accompanied by an ability to adapt what is said to suit the new target. Ideally the writer communicates in a strong, effective personal voice in which content, style, and rhetorical strategies vividly work together to engage the audience and achieve the author's purpose.⁴
- Empathize/Reflect. The social interactions involved in the act of writing for a literate audience presuppose a critical self-understanding, in which the writer is consciously aware of the social dynamics that govern the act of writing. When students have this critical self-awareness, they are able to perceive multiple viewpoints, to understand human motivations, biases, and interactions at all levels of interaction, and therefore should be able to write reflective texts that demonstrate a deep awareness of (and responsiveness to) their own and other people's perspectives.⁵

The underlying complex of abilities can be applied to textual content and not just to the act of communication, which is why this node is cross-classified as an instance of content reasoning. The ability to mentally model social situations, and thus to empathize with social agents and comprehend their perspectives, motives, and interactions has a direct application in a variety of narrative forms of writing. A common and important application comes when these

abilities are applied (perhaps in somewhat specialized ways) to analyze literary and historical narrative.⁶

- Reason Critically About Content. Students should be able to perform a wide variety of writing tasks that require the application of critical thinking skills to the content of the document they are producing. Critical thinking about content is a fundamental part of writing skill and cannot be separated cleanly from the process of planning, producing, assessing, and modifying texts. However, particular types of writing exercise different combinations of critical thinking skill, and high-level writing requires a flexible ability to integrate critical thought with writing. The nodes in this part of the writing competency model are purposefully aligned (in descending order) with the six higher levels of thinking in Bloom's taxonomy of educational objectives, without relying on it exclusively (since of course the categories presented align closely with categories in the literature review.)
- e Evaluate/Justify. Students should be able to evaluate texts and ideas against appropriate criteria; they should be able to make the basis for that evaluation explicit; and they should be able to construct documents that explicitly argue for or against a specific evaluation or position. As a result, they should be able to evaluate proposals and similar text types using explicit guidelines or standards, support or refute a position by adducing appropriate reasons and evidence, or critique someone else's argument by systematically evaluating reasons and evidence. A common and important application comes when these abilities are deployed to produce persuasive essays, though the same cognitive skills are involved in a wide variety of text genres.⁷
- Collect/Synthesize. Students should be able to perform a variety of rhetorical moves that reflect capacity to gather, coordinate, and combine information intelligently and selectively, so as to create integrated models; and should also be able to create new formulations and performances that reflect an integrated understanding of their subject matter. Thus they should be able to research a subject, report their findings, generate novel interpretations, and create texts or

- performances embodying them. This ability includes the skills needed to assess the relevance and usefulness of sources and to combine information from multiple sources effectively and perhaps also creatively.⁸
- moves that reflect the capacity to work within a complex, hierarchically organized system of ideas. This includes the ability to construct definitions, to use an outline or some other hierarchical model to represent a complex set of ideas, and more generally to analyze a subject by breaking it down into successively smaller pieces, deducing some details from general principles, and inferring others where they are not already explicitly supplied. They should also be able to use those skills to analyze other people's texts.
- **Apply/Predict.** Students should be able to perform a variety of rhetorical moves that reflect the capacity to apply existing knowledge to new situations, including stating hypotheses about what is going on in a specific situation, making predictions that follow from a hypothesis, explicating one document in terms of ideas from another, or applying the ideas presented in a text to a specific situation or example. ¹⁰
- **Describe/Explain.** Students should be able to perform a variety of rhetorical moves that reflect the capacity to build a mental model and apply it to specific cases, including paraphrasing statements, summarizing the content of source materials, describing specific cases, classifying specific cases by assigning them to a category, generalizing over examples, building comparisons and analogies, illustrating general statements, and, in general, explaining a system of ideas in their own words.¹¹
- Manage the Writing Process. Students should have all the skills necessary to
 produce high-quality, detailed, richly structured documents. At high levels of
 performance, the writing process requires an ability to juggle many different
 types of activity, all of them critical to success.¹²

- Assess/Modify Documents. Students should control the skills and strategies
 necessary to evaluate (and improve) their own texts and to provide useful
 feedback to colleagues and peers.¹³
- Critique/Rethink. Students should be able to evaluate a document at the level
 of its rhetorical purpose and content, and provide feedback accordingly; they
 should also be able to modify their plans for a document based upon the results
 of such an assessment, which often requires a deep rethinking of rhetorical
 goals and content.¹⁴
- Review/Revise. Students should be able to evaluate a document formally, in terms of its overall plan and structure, and propose edits and revisions that restructure, elaborate, and clarify the way in which that document presents its content to the reader, without any fundamental changes to goals or content.¹⁵
- Plan/Structure Documents. Even in first draft, students should be able to produce longer documents with rich internal structure; e.g., they should produce detailed, unified, and cohesive individual paragraphs, with appropriate use of sentence structure and discourse markers to handle focus and emphasis; sequences of paragraphs that develop main and subordinate ideas; and texts with a clear overall structure that is well-suited to audience, genre, and rhetorical purpose. The range of document structures they should be able to produce should range across a variety of text types including narrative/literary, informational/expository, and persuasive/argumentative texts. 17
- Select/Organize. Students should control a variety of strategies for producing texts with rich internal organization, including template patterns (such as the five-paragraph essay), use of graphic organizers, and outlining, among others. Ideally students will be able to combine a variety of strategies to develop novel organizing patterns as needed to satisfy particular rhetorical demands. This capability entails the ability to narrow in on an appropriate topic or thesis for a piece of writing, and to select appropriate, relevant materials given the topic.¹⁸

- **Focus/Connect.** Students should control the techniques that make it possible to communicate connections between ideas and relative emphasis within a text. These include sentence-combining techniques, use of transition words, and a variety of other cohesive devices. They produce texts that can be read fluently, that do not force the reader to guess at the connection between ideas, and which correctly communicate to the readers what they should interpret as important, and what they should interpret secondary. ¹⁹
- **Detail/Develop.** Students should control a variety of strategies for fleshing out the details of a document and elaborating upon its ideas, including free-writing strategies to help them discover what they already know, as well as question-answering and annotation techniques, and use of graphic organizers, among others, and should therefore be able to retrieve details, flesh out ideas, and develop content effectively. They provide enough information in their texts to answer any questions most readers are likely to have.²⁰
- Produce Texts in Standard Written English. Students should be able to fluently write in a formal style using Standard Written English, avoiding non-standard dialects, excessive informality, slang, and other features inappropriate to a formal, Standard English style, while maintaining consistent adherence to written conventions, including those for spelling, usage, mechanics, and grammar. They should also be able flexibly to alter the degree of formality they employ depending on genre, task, audience, and setting.²¹
- Use Formal English. Students should have a full command of the syntactic and lexical resources of the language, and not be restricted to the registers characteristic of oral interactions in typically informal contexts. Their use of vocabulary should be lively and precise; their phrasing and expression, clear and forceful; their style, appropriate to the task and genre.²²
- Control Syntax. A high level of command of formal English will be reflected by syntactic variety, appropriate use of characteristically formal or academic syntactic patterns such as passives, and an ability to modify stylistic choices in context. We distinguish between the ability to employ a formal (and to switch

- to an informal) style, and those aspects of sentence control that indicate document structure and the relations and relative importance of ideas, which we assign to the Plan/Structure Documents node.²³
- Control Vocabulary. A high level of command of formal English will be reflected in an ability to use rarer, more specialized vocabulary when appropriate, without overuse in contexts where ordinary words will do just as well. At a high level of competency, choice of words will be precise with respect to meaning, effective and appropriate for genre and task, and appropriate given audience and purpose.²⁴
- Follow Conventions. Students should habitually follow the conventions governing written English, and should be able to copyedit documents to make them conform to the same.²⁵
- Phrase/Rephrase. Students should be able to produce sentences whose
 grammatical and lexical patterns reflect Standard English usage. When they
 write or are presented with sentences following nonstandard patterns, they
 should be able to rephrase them to conform to conventional English grammar
 and usage.²⁶
- Proof/Correct. Students should be able to read a document not conforming to standard conventions for transcribing texts into written English, and then make modifications to spelling and punctuation to bring them into accord with Standard English orthography.²⁷
- Prerequisite Oral Language Skills. For the writing construct to be assessed
 meaningfully, students must have achieved a minimum basic level of oral
 fluency in English, sufficient to support everyday communication and
 instruction in school.²⁸
- Prerequisite Reading Skills. For the writing construct to be assessed
 meaningfully, students must have achieved a baseline level of reading skill,
 sufficient to read and comprehend at or above a fourth grade level.²⁹

Prerequisite Text Production Skills. For the writing construct to be assessed meaningfully, students must have reasonably automated text production skills.
 Depending on the response mode, handwriting and/or keyboarding skills should have advanced to a level at which the mere act of setting words to paper does not require a student's undivided, conscious attention.³⁰

Distinctives. The competency model sketched above has, of necessity, much in common with existing writing standards and commonly applied approaches to assessing writing. It is possible, for instance, to map the 6-trait method for scoring writing (Culham, 2001; Spandel & Stiggins, 1990) onto it fairly directly, as the following table illustrates:

Table 1

How the Six-Trait Scoring Model Maps onto the CBAL Competency Model

Six-Trait Category	Closest CBAL Category	Differences
Ideas	Detail/Develop	The 6-trait Ideas category also includes elements of Focus/Connect.
Organization	Select/Organize	The 6-trait Organization category includes references to specific elements of standard templates (title, introduction, conclusion) left implicit in the CBAL category.
Voice	Think Critically about the Social Context	The 6-trait Voice category includes ideas like "taking a risk" that may depend upon the purpose for writing.
Word Choice	Control Vocabulary	The 6-trait Word Choice category emphasizes the idea that word choices are "lively" and "engaging," whereas the CBAL category focuses on mastery of formal written and not just oral English.
Sentence Fluency	Control Syntax	Some aspects of the 6-trait Sentence Fluency category (e.g., use of connectives) are also relevant to the Plan/Structure Document node in the CBAL model.
Conventions	Follow Conventions	Very little difference.

What this comparison underlines, however, is the decision to include critical thinking for writing in the CBAL Writing competency model. It is their inclusion that most distinguishes the CBAL Writing competency model from 6-trait scoring and from various other standards for writing. The essential point to note is that the hierarchy of critical thinking skills entails a hierarchy of writing tasks, ranging from simple descriptive pieces to complex persuasive and evaluative works. While texts vary along a variety of other dimensions—audience, purpose, topic, genre, among others—the literature review supports a view in which the critical-thinking demands of a writing task are an essential part of the construct.

Creating Test Specifications

At the end of 2007, we had developed rough outlines for a set of test specifications for CBAL Writing, but had not yet fleshed it out in full. Test specifications were accomplished iteratively, by creating Periodic Accountability Assessments (PAAs) for pilot administration, analyzing the results, getting feedback from teachers, and revising pilot designs before the next administration. By the end of 2008, we had created and piloted four PAAs in Maine and had developed an initial, full specification for both a proposed generic design for CBAL Writing assessments and a blueprint for a sequence of PAAs to be administered during the course of a single school year. The constraints embodied in the blueprints represent the interaction of several constraints on test design, including

- 1. maintaining high standards for validity, construct coverage, and reliability;
- 2. incorporating design elements that encourage appropriate instructional preparation for the competencies being tested; and
- covering a reasonably large range of tasks, including multiple genres (literary, informational, persuasive) and multiple types of writing requiring different combinations of critical thinking skills.

Section 3.1 presents our current thinking about CBAL Writing assessments in general, regardless of grade level. Section 3.2 presents our current thinking about the 8th grade assessments we actually developed, and provides a detailed blueprint for a year's worth of *Periodic Accountability Assessments*. This blueprint was arrived at only as a result of a year's worth of development and analysis, and so none of the assessments we have piloted exactly matches these specifications. In later sections we will review the PAAs that were

actually piloted, identify how they differ from the final blueprints, and lead up to their intended revisions in the light of usability study and pilot data.

3. 1 General Test Specification

Individual forms. Each test form is administered on the computer, requires approximately 90 minutes, and has the following characteristics:

- 1. Embodies a realistic scenario in which a series of related tasks unfold within an appropriate social context. The scenario is clearly established at the beginning of the PAA to give students a sense of what they will need to do, and why.
- 2. Contains a sustained writing task (30–45 minutes) that strongly exercises the ability to *use critical thinking skills for writing, plan and structure documents, use formal written English*, and *follow written conventions*. This task may require students to write an essay, memorandum, letter, proposal, newspaper article, or other document form that they may encounter outside of school. The specific form will be determined by the scenario. The writing needs to be formal enough, and directed to a mature enough audience, so as to require written rather than oral vocabulary and style. These documents will be scored for:
 - content (*Use Critical Thinking for Writing*, application of critical thinking skills to the subject),
 - demonstration of the ability to draft well-structured documents (*Plan/Structure Documents*),
 - general verbal maturity (*Master Formal English*, demonstration of gradeappropriate acquisition of formal written English), and
 - adherence to written conventions (*Follow Conventions*).
- 3. Contains a series of lead-in and/or follow-up tasks, each relatively short (5–20 minutes) that require the student to think about the content to be addressed and to engage fruitfully with the overall *critical-thinking* and *rhetorical* requirements implied in the scenario. These tasks should also satisfy the following general criteria:

- a. They introduce enough information, through reading materials or other sources, to enable students to write meaningfully about the subject.
- b. They require students to demonstrate critical thinking skills that are necessary to perform well in the scenario modeled by the test.
- c. They are either short writing or selected-response tasks that most students can reasonably be expected to have mastered by the target grade, but are prerequisite to successful performance on the longer writing tasks.
- d. Taken as a set, these tasks *scaffold*, and thus help model, what it means to perform well on the overall scenario and represent important stages of the thinking-and-writing process needed for successful performance. Ideally, the scenario should represent a task that would be difficult for many students at grade level to achieve without help, but which they can achieve if guided through the process step by step with appropriate scaffolding.
- e. The shorter tasks should contrast with the longer-writing task in important ways, exercising parts of the competency model not easily measured by an essay task alone. In particular:
 - At least one task should be a critical reading task without a written response, to help disentangle the ability to reason critically about content from general writing and drafting skills.
 - At least one task should require students to demonstrate the ability to assess and modify documents.
 - At least one task should allow students to write in a less formal style, addressing peers or younger students rather than elders, allowing them to demonstrate the ability to switch between a more formal and a more oral style, and more generally, an ability to adapt what they write to *purpose and audience (Think Critically about Social Context)*.
- 3. Presents grade-appropriate texts for students to read and think about. The purpose of these texts is not to assess reading skills but to give students content to consider—e.g., to summarize, to analyze, to synthesize, to evaluate—in preparation for writing.

The texts may be informative, persuasive, literary, research-based, or a part of any other genre relevant to the scenario and purpose for writing. The length of the texts must not exceed reasonable reading-time expectations for the target grade.

4. Supports thinking and writing activities with resources such as guidelines, writers' checklists, evaluation criteria, tips for getting started, or other reference materials to help students as they progress through the composing process.

Each year's sequence of periodic accountability assessments. The sequence of assessments given during any given year and grade level is selected to exercise a broad variety of critical reasoning skills set within an equally broad array of rhetorical situations. The focus and content of each periodic accountability assessment will be driven by critical thinking and rhetorical requirements, and not by surface form. In particular:

- 1. Each periodic accountability assessment will require the student to demonstrate control of a different type of critical thinking.
- 2. Each assessment will require students to demonstrate the ability to write in a particular genre or form for which that type of critical thinking is essential.
- 3. The distribution of critical thinking skills across forms will reflect reasonable gradelevel expectations about the type and range of critical thinking skills that students will be expected to demonstrate.
- 4. Each periodic accountability assessment should be self-contained. The order in which forms are administered should not matter, in order that test sequences can be adjusted to match curricular requirements.

We are currently considering alternative specific designs for the annual sequence of periodic accountability assessments, depending upon the acceptability of multiple PAAs during the school year. Our current thinking focuses on a four-PAA sequence that allows us to focus on each of the major types of critical thinking skill in the competency model in a separate PAA, but other models may be necessary depending on testing constraints.

3.2. Eighth Grade Test Specification

The annual sequence. The following set of draft specifications reflects our current thinking for an 8th grade PAA sequence. We currently expect to have a four-PAA sequence defined as follows:³¹

Table 2
Specification for a 4-PAA Sequence of Eighth-grade Writing Assessments

Proposed PAA sequence	Critical thinking skill(s) in focus	Typical source materials: For 8th grade, an upper limit of 1500 words per text, 6000 words per test	Genre/form in focus for culminating task	Task sequence
1	Describe/explain/infer/ analyze	Poems, short stories, student reviews	Expository/ review	Read and review literary works
2	Infer/analyze evaluate/ justify	Newspaper stories, journal articles, task guidelines, and sample memoranda	Persuasive / memorandum	Develop a recommendation that uses guidelines or standards to support a specific choice
3	Collect/ synthesize	Newspaper stories, journal articles, web pages and student notes	Expository/ report	Research a topic and report the results
4	Evaluate/justify	Newspaper documents, journal articles, editorials and student essays	Persuasive / essay	Develop a position on an issue, considering arguments pro and con

Between Fall 2007 and Fall 2008, we administered and scored early versions of the four PAAs sketched in Table 2. Each test was developed using preliminary thinking about appropriate test blueprints, as modified by results of the previous pilots. The sections that follow present the final blueprints, as determined after review of the pilots, and provide a sketch of the preliminary design, the pilot, and the results of scoring. Screenshots of the tests as administered, copies of actual rubrics used, and benchmark papers are included in the appendices.

Test specification #1: Expository/review. Table 3 presents the design initially developed for the first test in the intended sequence. Since this test was not piloted until Fall 2008, score results were not available when this report was prepared. See Appendix B1 for screenshots of the test as administered in the Fall 2008 pilot and appendices B2 and B3 for rubrics and benchmark papers.³²

Table 3
Specification for a PAA Focused on Reviewing Literary Works

Sequence on Test	Task Type	Task Details		Competencies in Focus	Timing
1	Literary Interpretation (lead-in)	Select theme and supporting details for a literary text	0 0	Describe/Explain Infer/Analyze Empathize/Reflect	10 minutes
2	Simulated Peer Feedback (lead-in)	Using explicit criteria, provide feedback on another student's review of a second literary text	0	Assess/Modify Documents Cooperate/Collaborate	20 minutes
3	Literary Review (culminating)	Using explicit criteria, review a third literary text and develop a thematic interpretation, including explanations of theme supported with details from the text	0 0 0 0 0	Describe/Explain Infer/Analyze Plan/Structure Documents Master Formal English Follow Written Conventions	45 minutes
4	Web Overview (follow-up)	Briefly summarize two of the three literary texts (the third summary is provided as a model) and write a brief overview to convince readers that these works are worth reading. Select a title and graphic appropriate to the content and rhetorical purpose of the three texts.	0 0	Describe/Explain (accurate one sentence summaries of stories) Evaluate/Justify (motivate choice of work for website) Accommodate/Engage (appropriate tone for audience, choice of title and graphic) Master Formal English (contrasting task not requiring a more formal style)	15 minutes

Discussion. The test was developed in consultation with teachers and teaching strategists in Portland, Maine, and administered in three middle schools in that school district. It was completed by two hundred and forty-four 8th-grade students, of whom two hundred and fifteen responded to the essay portion of the test. Given the high non-complete rate on the

essay task (12 percent), partly due to technology and platform issues, and also the fact that many students completed the test in less than half the allotted time, there may be issues with student motivation or with understanding of the tasks that should be examined in greater detail.³³

Based upon our initial examination of student responses and our review of the test materials in light of student responses, a number of issues were identified that will need to be addressed in future iterations of this design. In particular:

- 1. While this test form was intended in the specification to represent an expository review of a literary work – a fairly standard category including book reports and other responses to literature – all tasks focused on the problem of identifying the theme of a literary work. Identification of theme is an instance of the describe/explain node in the critical thinking part of the competency model, since it involves formulating a generalization to cover a specific set of examples. But any realistic expository review would not focus only on theme; other aspects of a work would be covered, e.g., plot, setting, characters, and it may be problematic that the test excluded all other literary elements from its purview.³⁴ Given the relative speed at which students completed the test, it might be appropriate to expand the critical reading portion of the test (Task 1 in the design) to assess whether students have a clear understanding of a variety of literary elements. Such an expansion would still primarily focus on the describe/explain node of the competency model, but it would provide much better information about whether students truly understand the literary concepts necessary to inform their other, writing-focused tasks. It appears very unlikely that a single selected-response item will provide enough information to reliably identify students who have trouble with forming clear mental models of the text, a receptive prerequisite to the expressive task of describing and explaining a work in an expository review.³⁵
- 2. An examination of student responses to Task 2 suggests that the vast majority of students took the bulleted list of guidelines provided with the task very literally, and proceeded by attempting to mechanically "check off" each item in the list. This

tendency was made more problematic by the fact that it was possible to do fairly well on this task as initially designed by responding in a surface way to very simple, easily identified guideline violations: in particular, by focusing on the lack of mention of title and author and one grammatical error. The task may have worked well in one of its purposes – scaffolding in students' minds the requirements of the essay task they would meet in task 3 – but in its current form it is problematic as a model for how students should respond in a peer review. As discussed above, the literature indicates that beginning authors tend to assess writing at a very particularistic, surface level, focusing on concrete items such as spelling or grammar, and that improvements in writing quality are strongly facilitated by a focus on content and critical thinking. Thus if this item is to measure the *Assess/Modify Documents* node accurately, in line with the evidence model for this node, the guidelines provided to students need to be revised, and the sample student response altered, to force more of a focus on content and critical thinking.

3. An examination of student responses to the third, essay task revealed some confusion among students about whether their task was primarily expository or primarily persuasive, with some students focusing their effort on explaining the theme, as intended, while others focused on the task of motivating the story's value for inclusion on the website. The phrasing of the guidelines provided allowed for this confusion. It would probably be best if they required students to explain multiple literary elements (summarizing plot; describing character and setting, explaining the theme) rather than focusing solely on the theme. If the persuasive element of the task were clearly focused – requiring students to explain plot, characters, setting, theme in such a way as to make it obvious why students would like to read the work, and made it clear that accuracy in exposition was the priority, there would be less risk of students focusing on the only part of the intended task.

Test Specification #2: Persuasive Memorandum. Table 4 presents the revised blueprint for a persuasive PAA administered in Fall 2007.³⁶

Table 4
Specification for a PAA focused on writing a persuasive memorandum

Sequence on Test	Task Type	Task Details	(Competencies in Focus	Timing
1	Rubric-based scoring (lead-in)	Assess how well a memo meets the rubric's criteria for a persuasive memo and select feedback based upon the rubric.	0	Assess/Modify Documents Cooperate/Collaborate	15 minutes
2	Explain Project with an Example (lead-in)	Write a paragraph explaining the concepts and task demands of an assignment to students who missed an earlier class. Use project Sample A to illustrate how to meet at least three of the guidelines for the project.	0 0	Describe/Explain Accommodate/Engage Master Formal English (contrasting task not requiring more formal style)	15 minutes
3	Compare Two Projects (lead-in)	Complete a graphic organizer comparing how well Sample B and Sample C meet the project guidelines.	0	Infer/Analyze Evaluate/Justify	15 minutes
4	Write a Persuasive Memorandum (culminating)	Consider two newly proposed projects. Write a memorandum to your teacher arguing in favor of choosing one project and against the other alternative	0 0 0	Infer/Analyze Evaluate/Justify Plan/Structure Documents Master Formal English Follow Written conventions	45 minutes

Discussion. The piloted test design differed from that shown in table 4 in several ways. The topic was *Choosing a Service-Learning Activity* and consisted of four tasks. In the first task, the 8th graders had to select three ways out of a list of six to improve the introduction of a persuasive letter. Ten minutes were allocated to this task. In the second task, for which 20 minutes were allocated, 8th graders had to evaluate the body of a persuasive

letter against a set of guidelines, and find one guideline that was met in the letter (Task 2a), and one that was not met (Task 2b). An explanation had to be provided as well. In the third task, the participants had to explain why an example followed the guidelines for service-learning. The allocated time was 15 minutes. The final task was the main task for which 45 minutes were provided. The task was to argue for choosing one service-learning activity and against choosing others. We will discuss the differences between this structure and that proposed in Table 4 below. See Appendix C1 for screenshots of the test as administered in the Fall 2007 pilot, referred to hereinafter as *Service Learning*.

This form was developed in consultation with teachers and teaching strategists in Portland, Maine, and administered in three middle schools in that school district. It was presented to 56 children, selected randomly from a sample of 112. While all students completed the test, parental consent was only obtained for 46 of the 56. There were, however, some administrative difficulties associated with the software that collected keystroke timing data, which caused keystroke logs only to be available for twenty-two of the forty-six available students. Analysis of the keystroke log data was conducted separately, as was analysis of student responses using e-rater and other NLP features; see Deane, Quinlan, and Kostin (forthcoming), Almond (forthcoming).

Scoring and exploratory psychometric analysis task were not completed until first quarter, 2008. Internal ETS test development staff scored all tasks. The generalizability of the results is limited due to the experimental nature of the tasks (it was the first pilot study conducted), selective sampling (teachers were handpicked and highly involved in the assessment), and the limited sample size.³⁷

Possible Revisions. Feedback from teachers in Maine indicated that they perceived this test as relatively hard, perhaps because the main essay task required students to do a three-way comparison presenting arguments for one and against both of the others. On the other hand, the distribution of scores was quite comparable between this test form and that for Test Specification #4, which was administered on the same day drawing students from the same classrooms.

Re-evaluation of the test design in light of the revised blueprint and the scoring results suggests that certain issues are of particular interest:

- Task 1 is a selected-response task requiring students to select ways to revise the introduction to a letter. The connection to the rest of the scenario is tangential, and the construct it tests ability to make judgments about recommended revisions is somewhat different than that tested in the other tasks, all of which essentially require drafting text in response to a set of stimulus passages. It is thus not clear to what extent the much lower correlation between Task 1 and the total test is a function of a difference in format (selected response vs. constructed response) and/or construct (revision vs. ability to plan, structure, and produce a document in standard written English).
- Task 2 is a revision task, but has many similarities to the other constructed- response writing tasks. It involves production of a text evaluating a response using explicit guidelines, even if in this case the guidelines comprise a rubric for assessing writing quality. The fact that it has a high correlation with the other constructed-response tasks, and not with Task 1, could be interpreted either as evidence that Task 1 is very unreliable, or that Task 2's scoring is excessively dominated by general writing considerations rather than by whatever information it provides about student skill at assessing writing.

This information raises conceptual issues. Ideally, the test design should provide richer information about students' assessment/revision strategies, and these two items do not appear to be doing the job. The final blueprint reflects a decision to try a different strategy. Instead of an isolated selected-response task (with little connection to the larger rubric), and a constructed response document critique, we intend to try out a complex selected-response task in which students apply a trait scoring rubric and then select feedback. The feedback would include the student clicking and dragging specific parts of the sample response as examples to support the feedback they have chosen.

This kind of complex selected-response task, in which students record an evaluation, select feedback based upon it, and link that feedback to specific points in the text, functions in effect as a relatively large but highly scaffolded sequence of decisions, which should provide a richer picture of students' ability to assess writing (at least where explicit guidelines are provide for students to use as the basis for evaluation) without conflating it with text production.³⁸ Another consequence of this

- change is that it would bring the test closer into line with the overall test blueprint, by explicitly including a critical reading task.
- Because Tasks 1 and 2 are essentially concerned with the Assess/Modify Documents node in the competency model, the first two tasks focused on the task of convincing the principal to let the students do a service learning project. The guidelines for evaluating one service learning project against another are not introduced until Task 3, where they do not really fit into a plausible scenario. We therefore propose modifying Task 3 so that the rhetorical situation is explaining the guidelines for a good service learning project to a classmate who had been out sick. This provides not only a plausible scenario but also a relatively informal short constructed response (writing) task, creating the contrast between a formal and an informal occasion for writing mentioned in the test specifications.

Since the original Task 3 focused on the a simple description of how the guidelines applied to an exemplary service learning proposal, there is something of a jump from Task 3 to Task 4, which critically requires high-level evaluation and analysis. Moreover, Task 4 requires students to make a three-way comparison, which may also increase the difficulty of the essay task. We therefore propose to add an additional critical reading task—a comparison between two proposals, evaluating how well either proposal meets guidelines for a good service learning project. This would be a nested selected response task along the following lines, modeled after the kinds of graphic organizers used to for structuring comparison/contrast tasks:

Table 5
Graphic organizer design for a scaffolding critical-reading question

	Project #1	Project #2
Guideline the project satisfies BEST	<select guideline=""></select>	<select guideline=""></select>
	<pre><select a="" guideline="" it="" reason="" satisfies="" this="" well="" why=""></select></pre>	<pre><select a="" guideline="" it="" reason="" satisfies="" this="" well="" why=""></select></pre>
Guideline the project satisfies WORST	<select guideline=""></select>	<select guideline=""></select>
	<pre><select a="" guideline="" it="" poorly="" reason="" satisfies="" this="" why=""></select></pre>	<select a="" it<br="" reason="" why="">satisfies this guideline poorly></select>

These changes underlie the specification presented in Table 4, which brings the design for this P.A.A. in line with the general specifications presented in section 3.1. We expect to revise and pilot this version of the test in 2009.

Test Specification #3: Expository/report. Table 6 presents the outline of a specification for a PAA focused on writing an expository report drawing on information from multiple sources. The originally piloted version of this test was developed in consultation with teachers and teaching strategists in Portland, ME, and piloted in three middle schools from that school district in Spring 2008. It was presented in conjunction with a set of formative materials focused on the same set of competencies but using a different topic. Scoring was conducted by teachers recruited from local New Jersey and Pennsylvania schools, under the supervision of ETS test developers. The generalizability of the results is fairly limited due to the experimental nature of the tasks (it was the first pilot study conducted of this test design), selective sampling (teachers were handpicked and highly involved in the assessment), and the small sample size. In sum, the results of this study should not be interpreted other than that they provide a first impression for how the piloted form functioned in a specific school district (Portland, ME) among 8th graders with highly involved teachers.

Table 6
Specification for a PAA focused on writing an expository report

Sequence on Test	Task Type	Task Details		Competencies in Focus	Timing
1	Select Sources (lead-in)	Examine short descriptions of possible sources in relation to essential	0	Collect/Synthesize	10 mins.
		questions for the report. Select the three			
		<i>most useful</i> sources and the essential questions they help answer. Select the			
		three <i>least useful</i> and indicate the standard(s) of usefulness they violate.			
2	Summarize a	Read a source article and write a	0	Describe/Explain	20 mins.
	source	summary, focusing on the essential	0	Master Formal English	
	(lead-in)	questions for the report.	0	Follow Written Conventions	
3	Report	Examine additional source articles and	0	Collect/Synthesize	45 mins.
	information	notes from other students and then	0	Plan/Structure Documents	
	drawn from	write a report explaining the issue,	0	Master Formal English	
	multiple sources	drawing on multiple sources as needed.	0	Follow Written Conventions	
	(culminating)				
4	Rewrite an	Read a flyer for the report and rewrite	0	Assess/Modify Documents	15 mins.
	announcement	the flyer to make it clearer and more	0	Accommodate/Engage	
	(follow-up)	engaging while keeping the critical	0	Master Formal English	
		content the same.		(contrasting task requiring a less	
				formal, more engaging style)	

Possible revisions. Review of the relatively small sample of student responses (133) suggested a range of possible revisions. ⁴⁰ We note the following in particular:

- Open-ended responses to the second part of Task 1 suggest that students were not using the guiding questions when they selected answers as *most useful*, but were, instead, relying on relatively weak heuristics; in some cases, they may have been doing little more than looking for a word in common between the question and the website they selected as *useful*. This result suggests that students at this level may need more scaffolding, or that instruction needs to focus more strongly on teaching students to make use of guiding or focusing questions while doing research. We therefore propose replacing the constructed-response part of Task 1 with a selected response, so that students will have to select which of the guiding questions a source will help to answer. Since the task design allows students to change their answer to the first level of selected response during the task, this two-level selected response task may help direct student thinking toward the intended types of evaluative thinking and prevent them short-circuiting their response on the basis of simple heuristics.
- Task 2, a summarization task, allowed students to cut and paste from the source, and contained instructions warning about use of quotation marks in that case. The results suggest that students took these instructions as permission to cut and paste freely, with some responses containing very little of the students' own words, even if their selections correctly captured the gist of the summarized text, and their own words correctly bridged the logical relations between their quotes. In order for the summarization task to work as desired, we propose disabling the cut-and-paste feature and changing the instructions to emphasize the need for students to summarize the text in their own words.
- Task 3, the full essay task, was given a social context that encouraged a relatively informal style, since students were instructed that they were preparing an article for *Teen Talk*, a publication aimed at people their own age. Given the general specifications we have adopted (section 3.1.1), this choice of audience is now problematic, since the long essay task should measure the ability to write formal English, and students writing to a peer audience may adopt a more informal style. We therefore plan to revise the social situation. Perhaps *Teen Talk* can be described as a newsletter produced by a journalism class, with the final product being distributed to parents and teachers as well

as fellow students, which should be enough to make the requirement that students write in formal English clear. ⁴¹ Also, the specification for Task 3 made it unclear whether the purpose of the essay was expository or persuasive, by adding a partially persuasive purpose. Given the overall design for the P.A.A. sequence, the instructions should be revised to make the task in this test clearly expository.

• One of the graphics for Task 4 (an alarm clock picture) was intended as a distractor, but is arguably a valid key. We should replace this graphic with one that will be more clearly a wrong answer.

With these proposed modifications, we obtain the specification given in Table 6.

Test Specification #4: Persuasive essay. Table 7 presents the revised test specification for a P.A.A. focused on writing a persuasive essay. The originally version of this test was developed in consultation with teachers and teaching strategists in Portland, ME, and piloted in three middle schools from that school district. It was piloted in Fall 2007 in conjunction with the other persuasive test described in the section Test Specification #2: Persuasive Memorandum (p.31). Scoring was performed by ETS test developers. The generalizability of the results is limited due to the experimental nature of the tasks (it was the first pilot study conducted), selective sampling (teachers were handpicked and highly involved in the assessment), and the limited sample size. 42

Table 7
Specification for a PAA focused on writing a persuasive essay

Sequence on Test	Task Type	Task Details	Competencies in Focus	Timing
1	Identify Pros and Cons (lead-in)	Read four texts and use them to fill out a graphic organizer giving the pros and cons on the issue.	Infer/AnalyzeEvaluate/Justify	10 mins.
2	Select Critical Questions (lead-in)	Given an assertion and a list of questions about that assertion, select the three questions that most need to be answered to support it.	Detail/DevelopEvaluate/Justify	20 mins.
3	Critical Peer Response (lead-in)	Write a note to the author of an argument pointing out information from sources that could be used to challenge key assertions.	 Accommodate/Engage Cooperate/Collaborate Critique/Rethink Evaluate/Justify Master Formal English (less formal document contrast) 	45 mins.
4	Persuasive Essay (culminating)	Take a position on an issue and develop a persuasive essay supporting one's own position and refuting the opposite view.	 Evaluate/Justify Plan/Structure Documents Master Formal English Follow Written Conventions 	15 mins.

The topic of the test was "Should Children's Advertising Be Banned?" The test as actually piloted consisted of four tasks. In the first task, the 8th graders had to list two arguments against advertising to children. For each argument, they were instructed to write a paragraph explaining why they agreed or disagreed with the argument. The time allocated to this task was 20 minutes. In the second task, two arguments for advertising to children were to be given. As in the first task, they were instructed to write a paragraph explaining why they agreed or disagreed with the argument, and the time allocated to this task was 20 minutes. Forty minutes were allocated to the third task, in which the children were instructed to write an essay expressing their view with respect to advertising to children. The third task was the main task. The first two tasks can be considered as scaffolding items for Task 3. Finally, Task 4 had a selected-response format. The participants had 10 minutes to select five changes out of a list of 10 to revise their essay into an essay that would be proper for a different audience (4th grade rather than the local newspaper).

The three constructed-response tasks of each form were scored analytically on three strands: Language and literacy skills for writing (Strand I), writing process management skills (Strand II), and critical thinking skills for writing (Strand III). A five- point scale (0–4) was used for each of the strands. Each response was rated by one to four different raters. The same Strand I and Strand II rubrics were used as on Test Specification #1, though the Strand III rubrics were different given task differences.

Possible revisions. An examination of student responses to this test, and feedback from teachers, suggest a number of problems with the original test design. Students' answers to the third, essay, task tend to reuse material from their answers to the first two tasks (dealing with pro and con arguments drawn from sources). In consequence, it is difficult to view the three tasks as truly independent from one another. In addition, this design conflates critical reading tasks (e.g., recognizing when an argument is being made pro and con) with responding to the argument thus recognized. In effect, the original design jumped straight from recognizing arguments to framing a full persuasive essay, without much scaffolding in between.

We are therefore considering a significantly different structure, with the following changes:

- Replace the original Tasks 1 and 2 by a single selected-response task in which students read the four articles and select sentences in those texts that present arguments for and against banning advertising to children, perhaps also selecting sentences providing counters to each argument if available. Thus modified, this is a pure critical reading task, per the test design.
- Add a task focused specifically on appropriate strategies for elaborating an
 argument. In particular, we suggest a selected-response task in which they identify
 which three from a larger set of critical questions most needs to be answered to
 develop an argument.
- Eliminate the existing selected response task, which has little direct connection to argument, and add instead a peer revision task focused on the dialectic between author and audience. In particular, students would be given a sample student paper that takes a straw man position with serious problems both in argumentation and in use of assertions easily challenged from the reading materials provided with the text. Their task would be to write a note to the author of that paper, pointing out to them where they make easily challenged assertions, noting specific points from the sources where their statements can be challenged, and if possible, suggesting points from the sources that could be used to strengthen the argument.

With these changes we obtain the design presented in Table 7.

4. General Discussion and Conclusions

This report has outlined work on the CBAL Writing summative assessment during 2008, focusing on revisions to the competency model, development of a set of general test specifications, and the development and piloting of four P.A.A. designs to cover 8th grade writing in particular. In this section we will consider certain implications of the pilot work, particularly with regard to test design and scoring; outline plans for future work, and consider connections and conclusions of this work in a larger context.

4.1 The Test Design

It is appropriate at this point to consider the relation between the test designs piloted and, ultimately, proposed, and other measures. One point in particular merits close

consideration, one that was explicit in the partially-specified design in 2007 and highlighted in the design and specifications proposed above: skillful writing requires a certain level of reading skill. While the test design does not evaluate prerequisite reading skills, it explicitly includes critical reading skills (and associated writing strategies) at the core of the competency model. Essentially, the critical reading tasks specified in the design define receptive prerequisites to successfully completing the (critical-thinking dependent) expressive tasks of writing.

This is one of the fundamental ways that this design varies from more usual writing assessment designs. In essence, it reflects the position that successful writing is closely integrated with general literacy and thinking skills. This position has certain implications. It would be entirely reasonable to think in terms of an integrated reading/writing test—a test of general literacy—of which the writing test forms a component part. It might not be feasible (for a variety of logistical and test design issues) to create a single, integrated summative test in which reading and writing are tested in conjunction with the development of critical thinking skills, but the logic of the competency model for writing (and for reading, see O'Reilly et al., forthcoming) is that these are fundamentally integrated skills, and that a test designed to assess progress in developing writing fundamentally requires the assessment of associated skills.

One implication of this position is that the critical reading skills on the writing test need to be assessed reliably enough to provide a baseline measure of student progress on achieving at least receptive competency in using critical thinking skills for writing. We are not convinced, given our pilot results so far, that we are measuring this skill reliably enough. In a joint reading/writing test, the reading test might provide sufficient independent measurement. But in the stand-alone design we are considering, it is at least necessary to provide richer versions of the critical reading tasks, along the lines proposed for a revised version of each test.

4.2. Scoring and Dimensionality

For the two pilot administrations we have scoring results for, there is a significant difference in the quality of scoring. On the Fall 2007, administration, scoring showed high reliability and high levels of inter-rater agreement, both of which achieved much lower levels in the Spring, 2008, administration (except on Strand I, the basic language skills). The major

difference between the two scoring sessions was that the Spring, 2008 scoring session was done by teachers drawn from the local (New Jersey/Pennsylvania) area rather than by ETS test development staff. This difference in results suggests that we should continue to score responses in-house until the tasks and scoring standards are more established and the project is ready to prepare for operational scoring.

In addition, the results of the psychometric analysis strongly suggests that the Strand I, II and III writing scores loaded on a single dimension, raising certain questions about the trait scoring implied by the competency model.

However, the results of the automated scoring study (Deane, Quinlan, & Kostin, forthcoming), which examined the Fall 2007 data, strongly suggest the possibility that only the critical thinking tasks may need to be scored by human graders. Strands I and II in the Fall 2007 scoring—for example, for the nodes *Produce Texts in Standard Written English* and *Use Strategies to Manage the Writing Process*—could be predicted from e-rater features. Linear regression models had adjusted R² as shown in table 18, though the size of the sample was too small to permit too much to be concluded from these results:

Table 8

Adjusted R^2 for linear regression models predicting human scores on the Fall 2007 pilots

	Service learning	Advertising to children
Strand I (Produce Texts in Standard Written English)	.83	.70
Strand II (Organization/Development)	.77	.85

In both Strand I models, the best two predictors were the e-rater Mechanics and Grammar features; in both Strand II models, the best two predictors were the e-rater Organization and Development features. These results suggest that even though the strand scores were strongly correlated, the differences between them were construct relevant.

In future work, we expect to conduct much larger-scale studies, with revised scoring models, to determine whether automated scoring models can be validated on an operational scale. The tentative conclusion suggested by the results to date is that it may be possible to

provide automated scoring for three of the second-level nodes in the competency model: Plan/Structure Documents, Produce Texts in Standard Written English, and Follow Conventions. To the extent that we use selected response tasks (even complex selected-response tasks) for Assess/Modify Documents and for critical reading tasks, only a relatively small part of the total scoring (critical thinking for textual responses) would require human scoring.

4.3. Use of Behavior Features

One of the innovations explored in this year's work was the exploration of behavioral features derived from keystroke logging of student responses. There were significant operational barriers encountered during administration due to interaction between the keystroke logging software in the test client and various bandwidth and/or security issues involving the target school district's internet gateway. However, as Deane, Quinlan, and Kostin (forthcoming) discovered, these features did look promising. Certain keystroke features, such as pause length between words, added prediction to a linear regression predicting Strand I scores from e-rater features. Certain others, such as total time spent backspacing, added prediction to a linear regression predicting Strand II scores from e-rater features. In each case, the total sample size was far too small to allow generalization, but was suggestive and consistent with construct predictions.

Here also the results suggest possible directions, though they hardly establish any firm conclusions. It appears possible that behavioral features will be useful predictors, in which case various issues need to be addressed, such as whether these features should be used primarily for diagnostic purposes, since they measure the thought patterns of the student rather than the quality of the final product.

5. Conclusions and Future Directions

This paper has described work in 2008 on the summative assessment for CBAL writing. During that year, we completed a year's worth of pilot Periodic Accountability Assessments, scored them, and analyzed the results. We completely revised the competency and evidence models upon which they are based, developed a blueprint or test specification to guide the design

more precisely, and proposed revisions to the existing set of four Periodic Accountability Assessments to bring them into line with this specification.

All of the work focused on the eighth grade, and as such, had a character of its own, which would be modified considerably if we considered writing assessment over the entire spectrum from fourth to twelfth grade. We expect future work to examine developmental issues and pedagogical sequences in depth, to continue to develop automated scoring for those parts of the competency model where it is appropriate, and to continue to refine and revise each assessment design.

References

- Almond, R., Deane, P., Quinlan, T., Wagner, M., & Sydorenko, T. (in press). *A**Preliminary Analysis of Keystroke Log Data from a Timed Writing Task. Princeton,

 *NJ: ETS.
- Anderson, L.W., & Krathwohl, D. R. (2000). A taxonomy for learning, teaching, and assessing: A revision of Bloom's taxonomy of educational objectives. New York, NY: Longman.
- Beaufort, A. (2000). Learning the trade: A social apprenticeship model for gaining writing expertise. *Written Communication* 17(2), 185–223.
- Bereiter, C., & Scardamalia, M. (1987). *The psychology of written composition*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Bloom, B. S. (1956). *Taxonomy of educational objectives, Handbook 1: Cognitive domain.*New York, NY: Longmans Green.
- Bourdin, B., & Fayol, M. (1994). Is written language production more difficult than oral language production? A working memory approach. *International Journal of Psychology*, 29(5), 591–620.
- Bourdin, B., & Fayol, M. (2000). Is graphic activity cognitively costly? A developmental approach. *Reading and Writing*, 13(3–4), 183–196.
- Cohen, M., & Riel, M. (1989). The effect of distant audiences on students' writing. *American Educational Research Journal*, 26(2), 143–159.
- Coirier, P., Andriessen, J. E. B., & Chanquoy, L. (1999). From planning to translating: The specificity of argumentative writing. In P. Coirier & J. Andriessen (Eds.),

- Foundations of argumentative text processing (pp. 1–28). Amsterdam, the Netherlands: Amsterdam University Press.
- Culham, R. (2001). 6+1 traits of writing: The complete guide. New York, NY: Scholastic.
- Daiute, C. (1986). Do 1 and 1 Make 2? Patterns of influence by collaborative authors.

 Written Communication, 3(3), 382–408.
- Daiute, C., & Dalton, B. (1993). Collaboration between children learning to write: Can novices be masters? *Cognition and Instruction*, *10*(4), 281–333.
- Deane, P., Odendahl, N., Fowles, M., Quinlan, T., Baldwin, D., & Redman, M. (2007). *The CBAL writing assessment: A design and first steps toward implementation*.

 Unpublished manuscript.
- Deane, P., Odendahl, N., Quinlan, T., Fowles, M., Welsh, C., & Bivens-Tatum, J. (2008).

 Cognitive Models of Writing: Writing Proficiency as a Complex Integrated Skill

 (ETS Research Report No. RR-08-55). Princeton, NJ: ETS.
- Deane, P., & Quinlan, T. (2010). What automated analyses of corpora can tell us about students' writing skills. *Journal of Writing Research*, 2(2), 151–177.
- DeGroff, L. J. C. (1987). The influence of prior knowledge on writing, conferencing, and revising. *Elementary School Journal*, 88(2), 105–118.
- De La Paz, S. (2005). Teaching historical reasoning and argumentative writing in culturally and academically diverse middle school classrooms. *Journal of Educational Psychology*, 97, 139–158.
- De La Paz, S., & Graham, S. (1997a). Effects of dictation and advanced planning instruction on the composing of students with writing and learning problems.

 *Journal of Educational Psychology, 89(2), 203–222.**

- De La Paz, S., & Graham, S. (1997b). Strategy instruction in planning: effects on the writing performance and behavior of students with learning difficulties. *Exceptional Children*, 63(2), 167–181.
- De La Paz, S., & Graham, S. (2002). Explicitly teaching strategies, skills and knowledge: Writing instruction in middle school classrooms. *Journal of Educational Psychology*, 94(4), 687–698.
- Felton, M., & Kuhn, D. (2001). The development of argumentative discourse skill. *Discourse Processes*, 32(2–3), 152–153.
- Flower, L. (1989.) Cognition, context, and theory building. *College Composition and Communication*, 40(3), 282–311.
- Galbraith, D. (1999). Writing as a knowledge constituting process. In M. Torrance & D. Galbraith (Eds.), *Knowing what to write: Conceptual processes in text production* (pp. 139–160). Amsterdam, the Netherlands: Amsterdam University Press.
- Graham, S. (1997). Executive control in the revising of students with learning and writing difficulties. *Journal of Educational Psychology*, 89, 223–234.
- Graham, S., & Harris, K. R. (2000). The role of self-regulation and transcription skills in writing development. *Educational Psychologist* 35(1), 3–12.
- Graham, S., & Perin, D. (2007). Writing next: Effective strategies to improve writing of adolescents in middle and high schools—A report to the Carnegie Corporation of New York, Washington, DC: Alliance for Educational Progress.
- Hayes, J. R. (1996). A new framework for understanding cognition and affect in writing. In C. M. Levy & S. Ransdell (Eds.), *The science of writing: Theories, methods*,

- individual differences, and applications (pp. 1–27). Mahwah, NJ: Lawrence Erlbaum.
- Hayes, J. R. (2004). What triggers revision? In G. Rijlaarsdam (Series Ed.) & L. Allal & L. Chanquoy (Vol. Eds.), *Studies in writing: Vol. 13: Revision: Cognitive and Instructional Processes* (pp. 189–207). Boston, MA: Kluwer Academic.
- Hayes, J. R., & Flower, L. S. (1980). Identifying the organization of writing processes. In L. Gregg & E. R. Steinberg (Eds.), *Cognitive processes in writing* (pp. 3–30). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Hillocks, G, Jr. (1987). Synthesis of research on teaching writing. *Educational Leadership*, 44(8), 71.
- Hillocks, G., Jr. (1995). *Teaching writing as reflective practice*. New York, NY: Teachers College Press.
- Kaufer, D. S., Hayes, J. R., & Flower, L. S. (1986). Composing written sentences.

 *Research in the Teaching of English, 20, 121–140.
- Kellogg, W. H. (1970). Syntactic maturity in schoolchildren and adults. *Monographs of the Society for Research in Child Development*, 134, 35(1), 1–61.
- Kellogg, R. (1988). Attentional overload and writing performance: Effects of rough draft and outline strategies. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 14(2), 355–365.
- Kellogg, R. T. (2001). Competition for working memory among writing processes. *The American Journal of Psychology*, 114(2), 175–191
- Kent, T. (2003). Post-process theory: beyond the writing-process paradigm. In L. Z. Bloom, D. A. Daiker & E. M. White (Eds.), *Composition Studies in the New*

- Millennium: Rereading the Past, Rewriting the Future. Carbondale, IL: Southern Illinois University Press.
- Kostouli, T. (2005). Writing in context(s): Textual practices and learning processes in sociocultural settings. Heidelberg, Germany: Springer Verlag.
- Kuhn, D. (1991). The skills of argument. Cambridge, England: Cambridge University Press.
- Kuhn, D., Katz, J. B., & Dean, D., Jr. (2004). Developing reason. *Thinking & Reasoning*, 10(2), 197–219.
- Kuhn, D., Shaw, V., & Felton, M. (1997). Effects of dyadic interaction on argumentative reasoning. *Cognition and Instruction*, 15(3), 287–315.
- Langer, J. A. (1985). Children's sense of genre: A study of performance on parallel reading and writing. *Written Communication*, 2(2), 157–187.
- Langer, J. A. (2001). Beating the odds: teaching middle and high school students to read and write well. *American Education Research Journal*, 38(4), 837–880.
- Leahy, S., Lyon, C., Thompson, M., & Wiliam, D. (2005). Classroom assessment: Minute by minute, day by day. *Educational Leadership*, 63(3), 18–24.
- Matsuhashi, A. (1981). Pausing and planning: The tempo of written discourse production.

 *Research in the Teaching of English, 15(2), 113–134.
- McCutchen, D. (1986). Domain knowledge and linguistic knowledge in the development of writing ability. *Journal of Memory and Language*, 25(4), 431–444.
- McCutchen, D. (1996). A capacity theory of writing: Working memory in composition. *Educational Psychology Review*, 8(3), 299–325.
- McCutchen, D. (2000). Knowledge, processing, and working memory: Implications for a theory of Writing. *Educational Psychologist*, 35(1):13–23.

- McCutchen, D., Francis, M., & Kerr, S. (1997). Revising for meaning: Effects of knowledge and strategy. *Journal of Educational Psychology*, 89(4), 667–676.
- McCutchen, D. (2006). Cognitive factors in children's writing. In MacArthur, C.A.,

 Graham, S. and Fitzgerald, J., eds., *Handbook of Writing Research* (pp. 115–130).

 New York, NY: The Guilford Press.
- Means, M. L., & Voss, J. F. (1996). Who reasons well? Two studies of informal reasoning among children of different grade, ability, and knowledge levels. *Cognition & Instruction*, 14(2), 139.
- Mislevy, R. J., Steinberg, L. S., & Almond, R. G. (2003). On the structure of educational assessments. *Measurement: Interdisciplinary Research and Perspectives*, 1(1), 3–62.
- Olive, T., & Kellogg, R.T. (2002). Concurrent Activation of high- and low-level processes in written composition. *Memory and Cognition*, *30*(4), 594–600.
- O'Reilly, T., & Sheehan, K. M. (2009). Cognitively based assessment of, for and as learning: a framework for assessing reading competency (ETS Research Rep. RR-09-26). Princeton, NJ: ETS.
- O'Reilly, T., Sheehan, K., Fowles, M., & Hansen, E. (2008). *The CBAL Formative Assessment of Literacy Skill*. Unpublished manuscript.
- Pearl, S. (1979). The composing processes of unskilled college writers. *Research in the Teaching of English*, 13, 317–336.
- Perkins, D. N. (1985). Postprimary education has little impact on informal reasoning. *Journal of Educational Psychology*, 77(5), 562–571.

- Perkins, D. N., Allen, R., & Hafner, J. (1983). Difficulties in everyday reasoning. In W. Maxwell & J. Bruner (Eds.), *Thinking: The expanding frontier* (pp. 177–189). Philadelphia, PA: The Franklin Institute Press.
- Piolat, A., Roussey, J.-Y., Olive, T., & Farioli, F. (1996). Charge mentale et mobilisation des processus rédactionnels : examen de la procédure de Kellogg. *Psychologie Française*, 41(4), 339–354.
- Quinlan, T. (2004). Speech recognition technology and students with writing difficulties: improving fluency. *Journal of Educational Psychology*, 96(2), 337–346.
- Schilperoord, J. (2001). On the cognitive status of pauses in discourse production. In T.

 Olive & C. M. Levy (Eds.), *Contemporary tools and techniques for studying writing*(pp. 61–90). Amsterdam, the Netherlands: Kluwer.
- Shanahan, T. (2006). Relations among oral language, reading and writing development (pp. 171–183) in C. A. MacArthur, S. Graham, and J. Fitzgerald, (Eds.), *Handbook of Writing Research*. New York, NY: Guilford Press.
- Spandel, V., & Stiggins, R. J. (1990). *Creative writers: Linking assessment and writing instruction*. White Plains, NY: Longman.
- Torrance, M., & Galbraith, D. (2006). The processing demands of writing. In C.

 MacArthur, S. Graham & J. Fitzgerald (Eds.), *Handbook of Writing Research*. New

 York, NY: Guilford Press.
- Yarrow, F., & Topping, K. J. (2001). Collaborative writing: The effects of metacognitive prompting and structured peer interaction. *British Journal of Educational Psychology*, 71(2): 261–282.

Notes

- ¹ The development of *formative scenarios* that can be used in the classroom was undertaken under the aegis of a separate CBAL project, which involved both the creation of materials for formative use, and extensive interaction with teachers. While this work went on in a separate project, many of its results fed back into the summative project, and will be discussed where appropriate below.
- ² See Deane et al. (2008), pp. 22–24, for discussion of the role of critical thinking as part of the writing construct. See Deane et al. (2008), pp. 3–12 for discussion of the cognitive basis for skill in planning and producing structured documents.
- ³ See Deane et al. (2008), pp. 12–16 for discussion of writing competency viewed as enculturation in a social practice.
- ⁴ See Deane et al. (2008), pp. 17–19 for discussion of the role of the audience and the author's stance toward the audience.
- ⁵ See Deane et al. (2008), pp. 55–61 for a discussion of the role of social-evaluative skills, e.g., skills requiring sensitivity to a theory of mind and to social context, and how these interact with various genres of writing.
- ⁶ Deane et al. (2008), p. 56, provides a diagram that represents social-evaluative skills as applied to narrative in the more literary sense. The construct should be understood to also include a more generic view of social-evaluative skills, focusing on theory of mind, mental simulation of social interactions, and thus the ability to adopt a perspective and to mentally place oneself inside someone else's point of view.
- ⁷ See Deane et al. (2008), pp. 39–44, 53–54, 58–61, particularly the diagram on p. 61 for discussion of evaluative skills focused specifically on argumentation. Labeling the node evaluate/justify instead of support/refute highlights the role that these skills can play in a variety of tasks and modes of writing, rather than exclusively applying them to persuasive writing.
- ⁸ The complex of skills identified here as collect/synthesize, earlier termed gather/synthesize, were not clearly distinguished in the literature review, though see the discussion of prior knowledge and relevance in Deane et al. (2008), pp. 47–51. The importance of the cluster of skills identified here became evident in the course of reviewing state standards and working

- with teachers, for whom the development of research and synthesis skills is an important element from the middle grades onward.
- ⁹ See the discussion of exposition-related critical thinking skills in Deane et al. (2008), pp. 52–53, 55–59, which correspond roughly to what we here call infer/analyze, except that that discussion implicitly subsumes lower levels of critical thinking (application and explanation) into the same category. It is worth breaking these levels out as shown in this model, since the resulting categories map well onto existing educational practice and help to communicate more clearly and precisely what kinds of critical thinking students must master in order to become effective writers.
- ¹⁰ This description is based upon the judgment that many of the skills involved in the scientific method e.g., determining how a theory applies, and devising experiments that based upon the predictions of a theory are a specialized but also critical instance of the ability to apply ideas.
- ¹¹ Describe/Explain may be viewed as the expressive counterpart of the kinds of verbal skills fundamental to reading comprehension (so-called 'model building skills'; see O'Reilly et al. forthcoming).
- 12 See Deane et al. (2008), pp. 9–12 and the diagram on p. 15.
- ¹³ The importance of peer review and similar collaborative processes in writing—and of the evaluative skills that support such activities—are highlighted throughout the literature. See Deane et al., pp. 32–34, 38 for discussion.
- Note the conclusions of McCutchen (1997) that effective revision at higher skill levels requires development of critical reading skills. The 'Critique/Rethink' node is intended to include both these kinds of receptive skills, and the ability to use them to plan/replan one's writing at a conceptual level. See also McCutchen (2006: 123–124). See also O'Reilly & Sheehan (submitted) for connections to the CBAL Reading competency model.
- ¹⁵ Note that this kind of revision necessarily involves a task schema for revision that involves a problem-solving approach, even if the problems to be solved are identifiable at specific points in the text. This entails not only possession of appropriate task schemas, but also the supporting skills needed to differentiate one's intended message from what is actually

communicated by the text. See McCutchen (2006: 124–125) for a discussion of some of the issues involved.

- ¹⁶ The analysis presented in the companion report by Deane, Quinlan & Kostin (forthcoming) indicates that automated scoring features measuring development, sentence complexity, and organization load on a common factor, which we interpret as the ability fluently to create documents with complex internal structure, and thus as measuring the plan/structure document node. The same analysis suggested a breakdown of the remaining features measurable by automated NLP techniques into two further categories, to be discussed below.
- ¹⁷ See Deane et al. (2008), pp. 24–30 for a discussion of some of the cognitive factors that underlie variation across different text types.

- There is a specific construct decision made by this definition that is important to note. It identifies acquisition of oral skills in English as a prerequisite skill, and places the focus on the acquisition of the richness of vocabulary and syntax characteristic of writers skilled in the use of formal written English. Two implications follow. First, since ELL students may distinguish a different pattern of acquisition, it will be important to identify students' level of language acquisition, as the kind of writing assessment posited here will not be appropriate before students achieve reasonable oral fluency. Second, we are placing an emphasis on acquisition of the formal code; creative uses of language are not part of the construct at this level, but reflect critical thinking of the sort usually labeled as achievement of effective 'voice' on the part of the writer.
- ²² Deane, Quinlan, & Kostin (forthcoming) identify a 2nd order factor over NLP features that appears to measure the continuum between informal/casual/oral language on one hand, and formal/studied/written or academic language on the other. We identify the dimension measured by these features as measuring the ability to use formal English, whether in an oral or a written context.

¹⁸ Deane et al. (2008), pp. 3–7, 10, 24–28.

¹⁹ The phenomena covered by the Focus/Connect node of the competency model comprise the kinds of textual cues covered in Deane et al., (2008), pp. 21–22, 27–30.

²⁰ See Deane et al., p. 37; pp. 48–54.

- ²³ This node corresponds in part to the "Sentence Fluency" category of 6-trait scoring models (Culham, 2001) and also to the "Sentence Variety" feature of many writing evaluations. Note, however, that sentence length/sentence complexity measures appear from the analysis in Deane, Quinlan, & Kostin (forthcoming) to form part of the "Plan/Structure Document" competency.
- ²⁴ This node corresponds roughly to the "Word Choice" feature in 6-trait scoring (Culham, 2001) and to vocabulary features in e-rater.
- ²⁵ This node corresponds roughly to the "Conventions" feature in 6-trait scoring (Culham, 2001).
- ²⁶This node corresponds roughly to the "Grammar" and "Usage" features in e-Rater.
- ²⁷ This node corresponds roughly to the "Mechanics" feature in e-Rater.
- ²⁸ See Deane et al., p. 36.
- ²⁹ See Deane et al., pp. 8–9. Also O'Reilley and Sheehan. (2009).
- ³⁰ See Deane et al., p. 8.
- Note that the test focuses on expository and persuasive writing, not narrative. This is a consequence of the strong focus on learning expository and persuasive writing in 8th grade. Other grades would have very different distributions of tasks, as we expect to explore in future work. The selection of tasks in the current set was motivated in part by examination of middle school writing standards in Maine and several other states.
- ³² This test was administered in conjunction with a set of formative materials focused on the topic "robots and ethics." The formative materials were developed under the companion formative writing project, and are described in that project's report.
- ³³ See Appendix F for a description of the administration procedures.
- ³⁴ These literary elements are viewed essentially as part of the metalanguage of talking about narrative, and as such, are viewed in our model as content-oriented aspects of the empathize/reflect competency.
- ³⁵ Note that this first item is in fact a critical reading item, reflecting the necessary overlap in the design between reading and writing, especially with regard to critical thinking tasks. As noted above, measuring relevant critical reading skills is of necessity part of the design.

- ³⁶ This test was administered in conjunction with a companion set of formative materials focused on the topic of junk food, which were described in the 2007 formative project report.
- ³⁷ A psychometric analysis is available in an internal project report from Frank Rijmen. On this, as the other analyses reported below, the reliability analysis yielded a Cronbach's Alpha above 0.8.
- ³⁸ We ultimately envisage a task in which the student's sequence of choices leads to an automatically constructed feedback letter, one that students are allowed to revise by going back and redoing their choices until the final product matches their evaluation of the text. Scoring the resulting nested set of decisions presents technical challenges, but would represent a highly scaffolded version of the text evaluation task that separates the ability to assess another student's writing from their ability to draft text in general. This may be more appropriate for 8th grade students, who may have limited previous experience critiquing other students' writing, and may provide a kind of template from which instruction in peer assessment would naturally proceed.
- ³⁹ The topic was "e-waste" (the environmentally friendly disposal of obsolete electronic devices). The formative materials were developed as part of the companion, formative writing project, and are described in detail in that report.
- ⁴⁰ A psychometric analysis of this pilot was performed by Frank Rijmen, and the results are available in an internal project report. Here, as with the previous design, the reliability analysis had a Cronbach's Alpha above 0.8.
- ⁴¹ Another issue with task 3 is the large set of notes presenting predigested information. Cut and paste needs to be disallowed as in task 2, and it might be worth considering whether there is any way to make the task of reading the notes less burdensome, since students do not encounter most of the information presented in the notes until they undertake the full essay task. We have not proposed modifying the test to address this issue, primarily because the obvious ways to address it would either add extra tasks for which there is little time, or would introduce dependencies among tasks that would not be desirable.
- ⁴² A psychometric report on this test is available in an internal project report from Frank Rijmen. The sample size was very small (56 8th graders), though the reliability analysis indicated a Cronbach's Alpha above 0.8.

List of Appendices

	Page
A - Revising the Evidence Model	57
A1 - Evidence Model for the Competency Model Variable Reason Critically About Social	
Context (PERSPECTIVE, PURPOSE, AUDIENCE)	58
A2 - Evidence Model for the Competency Model Variable, Reason Critically About Conten	t61
A3 - Evidence Model for the Competency Model Variable Assess/Modify Documents	67
A4 - Evidence Model For the Competency Model Node Plan/Structure Documents	71
A5 - Evidence Model for the Competency Model Node Master Formal English	73
A6 - Evidence Model for the Competency Model Node Follow Written Conventions	75
B - Screenshots for Test Design #1, Expository Review of Literature (Administered Fall	
2008; Scoring Underway)	77
B2 - Rubrics for Test Design #1 ("Book Club" Test)	87
B3 - Benchmark Papers For The "Book Club" Test	96
C1- Screenshots for Test Design #2, Persuasive Memorandum (Service Learning	
Administered Fall 2007; Scored/Analyzed In 2008)	109
C2 - Frequency Tables for Test Design #2	114
C3 - Rubrics for the Service Learning Test	121
C4a - Benchmark Responses Illustrating Sentence-Level Skills	129
C4b - Benchmark Responses Illustrating Document-Level Skills	133
D1 - Screen Shots for Test Design #3, Expository Report (Teenagers & Sleep Administered	
Spring 2008; Scored Summer 2008)	146
D2 - Frequency Tables at the Strand Within Task Level for Test Design #3	154
D3 - Keys and Rubrics for the <i>Teenagers & Sleep</i> Test	159
D4 - Benchmark Responses for Teenagers & Sleep Test	165
E1 - Screen Shots for Test Design #4, Persuasive Essay (Advertising to Children	
Administered Fall 2007; Scored/Analyzed in 2008)	188
E2 - Frequency Tables at the Strand Within Task Level for Test Design #4	193
E3 - Rubrics for the Advertising to Children' Test	197
E4 - Benchmark Responses for Advertising to Children Test	201
F - Principles of the CBAL Initiative	212
G - Administration Procedures	213

Appendix A

Revising the Evidence Model

At the end of 2007, we had had very little experience scoring CBAL Writing PAAs. In 2008 we scored four different PAAs, developed rubrics and topic guides, identified benchmark papers, conducted automated scoring research, collected keystroke and activity log data from computerized administrations, and various other scoring and analysis activities. More detail about the scoring experience will be provided in detailed sections about each of the PAAs scored during 2008, but this experience led, in combination with the revision of the competency model, to a much more detailed specification of the evidence we expect to be relevant to the measurement of student performance. The sections that follow present this revised evidence model, altered to provide a specification of performance levels at the second tier of the competency model (*Reason Critically about Social Context; Reason Critically about Content; Assess/Modify Documents; Plan/Structure Documents; Master Formal English; Follow Conventions*). This appears to be the lowest level at which we can reasonably expect to provide feedback about student performance based upon our experiences to date.⁴³

In the form presented below, each part of the evidence model is presented as a 'pseudo-rubric' with four performance levels. These are designed to provide a concrete picture of how responses would provide evidence for performance levels, without being used directly as rubrics with which to score individual responses. The actual rubrics used in scoring will be provided in detail sections presenting results for each individual PAA. It should be read as our best summary of an evidence model, specifying what kinds of information about student performance will be taken as evidence for high or low performance in each of the 8 mid-level nodes of the competency model.

The evidence specified here may be evidence about features of the final product, features of student performance while producing that product, or may address other work products, as specified below. We include features of the text response that require human scoring; features of the text response that we believe can be automatically scored; and features of student behavioral logs, among others. The sections below summarize the current state of the evidence model, with later sections discussing motivations for change.

Appendix A1

Evidence Model for the Competency Model Variable Reason Critically About Social Context (PERSPECTIVE, PURPOSE, AUDIENCE)

This node of the competency model corresponds to an assessment of a student's overall handling of the rhetorical task set by the prompt (assuming an essay response). It includes the variables usually subsumed under the heading of *voice* and the social transaction between purpose and audience.

4 EXCELLENT -- Typical performance patterns at each level

Cooperate/Collaborate:

• Demonstrate a deep awareness of and effective adherence to the norms and standards that govern the social role the writer is currently fulfilling (e.g., author, coauthor, reviewer, editor, copyeditor, etc.)

Among other things, this awareness entails that the author

- o understands the process of writing as a negotiation between the author and other socially interested parties, rather than as a pure act of individual creation, and therefore is willing to submit what he or she writes to the judgment of other interested parties, who may play such roles as collaborators, reviewers, editors, publishers, and readers.
- o is able to make effective use of publicly available resources, such as dictionaries, manuals, and style guides as means by which to conform to the social expectations of other interested parties
- o makes effective use of critiques and responses from reviewers to determine how to edit and revise a document in line with the social expectations of his particular writing community
- o is willing to provide critiques and responses to others and is aware of the kinds of feedback most likely to be helpful

Accommodate/Engage:

- Demonstrate a clear understanding of the purpose of the task and mobilize effective strategies to accomplish that purpose
- Demonstrate a clear understanding of how to approach the audience given current task and social role, choosing a voice and tone that maximizes the chance of success
- By choice of language, illustrations, details, etc., demonstrate an awareness of the audience, selecting material that is likely to consistently engage the interest of readers

Empathize/Reflect:

- Demonstrate both awareness of and empathy with multiple points of view, and a clear sense of their importance and significance
- Be fair-minded, avoiding bias or stereotyping, and in general, demonstrates an ability to step outside one's own particular perspective and take an objective view

- Demonstrate a clear understanding of the purposes, strategies, and interactions/conflicts of all parties
 - 2 ADEQUATE -- An "Adequate" response may do any of the following things, while failing to achieve the requirements for an excellent rating:

Cooperate/Collaborate:

- Adhere reasonably well to the norms and standards that govern the social role the writer is currently fulfilling
 - o understands the process of writing as a negotiation between the author and other socially interested parties, rather than as a pure act of individual creation, and therefore is willing to submit what he or she writes to the judgment of other interested parties, who may play such roles as collaborators, reviewers, editors, publishers, and readers.
 - Makes appropriate use of publicly available resources, such as dictionaries, manuals, and style guides as means by which to conform to the social expectations of other interested parties
 - o is willing to use critiques and responses from reviewers to determine how to edit and revise a document in line with the social expectations of his particular writing community
 - o is willing to provide critiques and responses to others' texts when asked to do so, and is aware that such responses should target more than surface features.

Accommodate/Engage:

- Demonstrate an understanding of the purpose of the task and mobilize reasonable strategies to accomplish that purpose
- Demonstrate some understanding of the need to choose a voice and tone that will be effective with the target audience
- By choice of language, illustrations, details, etc., demonstrate an awareness of the audience, sometimes achieving effects likely to engage the interest of readers

Empathize/Reflect:

- Demonstrate awareness of multiple points of view & ability to take them into account
- Maintain a generally fair attitude, avoiding bias or stereotyping.
- Demonstrate an awareness of the purposes, strategies, and interactions/conflicts of all parties

2 LIMITED -- A "Limited" response displays <u>one or more</u> of the following problems, where the author:

Cooperate/Collaborate:

- Demonstrates awareness of—but sometimes violates--the norms and standards that govern the social role the writer is currently fulfilling
 - In particular a writer may
 - o be resistant to the idea of a multiple-stage, negotiated writing process and avoids it unless forced
 - o fail to use resources such as dictionaries, manuals, or style guides unless required to

do so

o provide minimally useful feedback when asked to do so, with feedback focused almost entirely on surface characteristics of the text

Accommodate/Engage

- Does not seem completely clear about the purpose of the task
- Mobilizes rhetorical strategies inappropriate to the task at hand
- Writes in a voice or tone clearly inappropriate to the audience
- Makes little sustained effort to use language, illustrations and details to engage the interest and emotional involvement of the audience

Empathize/Reflect

- Demonstrates some awareness of multiple points of view but shows difficulty consistently taking them into account
- Occasionally lapses into bias or stereotyping.
- Fails to fully and accurately represent the purposes, strategies, and interactions/conflicts of key parties

1 MINIMAL -- A "Minimal" response displays <u>one or more</u> of the following problems, where the author:

Cooperate/Collaborate:

- Responds in a manner completely inappropriate to the norms and standards that govern the social role the writer is currently fulfilling [Cooperate/Collaborate]

 As a result the author may:
 - o only produce texts in a one-shot manner, and makes no significant changes if asked to revise
 - o consistently fail to provide more than perfunctory feedback when asked to review another author's work

Accommodate/Engage:

- Misunderstands the purpose of the task
- Shows no consistent voice / has no clear perspective of his or her own
- Displays no effort to engage the interest of the audience

Empathize/Reflect:

- Displays a consistently inappropriate egocentric point of view
- Demonstrates a generally biased or stereotyped evaluation of other parties.
- Misrepresents or ignores purposes, strategies, and interactions/conflicts of key parties

Appendix A2

Evidence Model for the Competency Model Variable, Reason Critically About Content

This node of the competency was already described in fairly great detail in Deane et al. (2007), although the change in the structure of this part of the competency model required some modifications to the evidence model, both for presentation purposes and to accurately represent the restructuring of the model at the lowest level, so that it corresponds more closely with the pedagogic categories of Bloom's taxonomy.⁴⁴

Examination of the categories presented below will generally demonstrate that the evidence model for "support/refute" from Deane et al. (2007) maps onto the "evaluate/justify" node of the current model, such as "gather/synthesize" maps onto "collect/synthesize" and "explain/hypothesize" mostly maps onto describe/explain, though some parts of it correspond to the new nodes "infer/analyze" and to "apply/predict".

Note that the focus continues to be on the application of these skills to writing; while reasoning skills have many uses, it is their connection to the content of writing that makes them relevant in our model.

4 EXCELLENT -- An "Excellent" response will consistently do many of the following things, depending on task requirements:

Evaluate/Justify:

- Accurately identify key issues
- Take a clear stance (whether it is communicated indirectly or explicitly represented by a thesis statement)
- Accurately identify critical assumptions
- Consistently base its own arguments on reasonable, consistent, and valid assumptions
- Support its position effectively using valid logic and well-reasoned arguments This includes being able to
 - o precisely and clearly formulate claims
 - o elaborate effectively on these claims by providing supporting arguments subclaims)
 - o provide convincing evidence to support the truth of claims and subclaims
- Address and if necessary effectively rebut arguments against the position that it takes This includes being able to
 - o Evaluate whether claims are consistent with prior knowledge
 - o Recognize whether evidence supports or refutes an argument
 - o Generate effective counterarguments to a claim, by questioning the warrants of an argument, finding alternative explanations, or challenging supporting evidence
- Accurately evaluate the relative strength of arguments This includes being able to

ins includes being able to

- Accurately predict how plausible and appealing an argument will be to an audience
- o Accurately predict how credible evidence will seem to an audience
- o Anticipate counterarguments likely to occur to members of an audience

Collect/Synthesize:

- Draw upon a wide variety of relevant sources of information to support its position This may include:
 - o Sophisticated, reflective use of personal experience and observation
 - Use of effective strategies for collecting information from primary, secondary and tertiary sources
- Demonstrate an ability to select the most reliable, credible sources of information. This may include:
 - O Deploying sophisticated strategies for assessing whether a source is reliable, biased, and/or relevant
 - o Deploying effective research strategies to seek out relevant information
- Gather enough credible and relevant information to support its position effectively This may include:
 - o Effective use of the meta-language of research to describe a source's provenance, validity, and usefulness
 - o full control of the linguistic devices for quoting and paraphrasing and attributing source material, such that the writer consistently uses these devices appropriately and correctly
 - o full control of appropriate strategies for managing the process of gathering information and reporting it to others, including appropriate strategies for note-taking and record-keeping and formatting/presentation of formal research papers
- Effectively integrate information from multiple sources
 - O Using facts from multiple sources to create effective inferences about the reliability of any one source
 - O Using sophisticated inferential strategies to resolve inconsistencies and fill in gaps when combining information from multiple sources

Infer/Analyze:

- Distinguish clearly between fact and inference
- Demonstrate the ability to distinguish between probable and merely plausible inferences
- Demonstrate the ability to make insightful rather than merely superficial inferences.
- Accurately identify significant consequences and implications of its materials
- Demonstrate an in-depth, precisely accurate understanding of the scope and complexity of the material it addresses

Apply/Predict:

- Identify insightful hypotheses and propose specific, testable predictions based upon them
- Create effective and insightful applications of an existing model
- Reason accurately in applied contexts such as making inferences about cause and effect
- Create accurate simulations, invented narratives or other effective generative applications of a system of ideas

Describe/Explain

- Effectively illustrate general ideas with examples
- Formulate insightful comparisons and analogies
- Induce accurate and insightful generalizations over specific examples
- Describe specific situations or examples accurately and vividly
- Distinguish effectively between main and subordinate ideas, using this ability to summarize information accurately
- Restate or paraphrase information accurately in one's own words

3 ADEQUATE -- An "Adequate" response will do many of the following things depending on task requirements, while failing to achieve the requirements for an excellent rating:

Evaluate/Justify

- Take a stance on the issues
- Identify some but not all key issues
- Accurately identify some but not all critical assumptions
- Base its own arguments on plausible assumptions
- Support its position with logical arguments

This includes:

- Clearly formulating claims
- Elaborating on these claims by providing plausible supporting arguments (subclaims)
- Providing plausible evidence to support the truth of claims and subclaims
- Address some arguments against the position that it takes

This includes:

- evaluating whether claims are consistent with prior knowledge
- recognizing whether evidence supports or refutes an argument
- recognizing obviously invalid arguments and identifying clearly questionable evidence
- Evaluate the relative strength of arguments.

This includes:

- o Recognizing when an argument is particularly strong or weak
- o Identifying arguments with strong surface appeal for a particular audience
- Recognizing when an audience will find evidence immediately plausible or questionable

Collect/Synthesize:

- Draw upon relevant sources of information to support its position This may include:
 - o Making reasonable and appropriate use of personal experience and observation
 - Making reasonable and appropriate use of information from primary, secondary and tertiary sources
- Demonstrate an ability to select credible sources of information This may include:

- O Deploying reasonable strategies for assessing whether a source is reliable, biased, and/or relevant
- o Deploying appropriate research strategies to seek out relevant information
- Gather enough credible/relevant information to provide plausible support This may include:
 - O Some ability to keep track of which parts of a complex set of ideas are drawn from sources and which are provided by the author
 - o Some understanding of the social reasons for attributing sources
 - O Some ability to use the meta-language of research to describe a source's provenance, validity, and usefulness
 - O Ability to make appropriate use of the linguistic devices for quoting and paraphrasing and attributing source material
 - o Appropriate strategies for managing the process of gathering information and reporting it to others
- Integrate information from multiple sources
 - O Using facts from multiple sources to create plausible inferences about the reliability of any one source
 - O Using plausible inferential strategies to resolve inconsistencies and fill in gaps when combining information from multiple sources

Infer/Analyze

- Identify significant consequences and implications of the position it takes
- Usually distinguish fact and inference

Apply/Predict

- Identify hypotheses and propose predictions based upon them
- Create reasonable applications of an existing model
- Make reasonable inferences about cause and effect
- Create plausible simulations, invented narratives, or other generative applications of a system of ideas

Describe/Explain

- Illustrate general ideas with examples
- Induce generalizations over specific examples
- Create plausible comparisons and analogies
- Describe specific situations or examples with reasonable accuracy
- Distinguish between main and subordinate ideas, using this ability to summarize information reasonably accurately
- Restate or paraphrase information with reasonable accuracy in one's own words

2 LIMITED

A "Limited" response typically displays one or more of the following problems:

Evaluate/Justify:

- Take a vague, ambiguous or inconsistent stance on the issues
- Fail to distinguish key issues from red herrings
- Fail to identify and motivate critical assumptions
- Leave important claims unsupported and unelaborated
- Misrepresent or dismiss other positions without argument even when the task requires them to address alternate points of view
- Argue for its position using logical fallacies
- Accept questionable evidence uncritically
- Fail to address important counterarguments
- Present arguments without regard for appeal to the target audience
- Draw upon irrelevant, questionable, or inaccurate sources of information

Collect/Synthesize:

- Not provide enough information from multiple sources to build a strong case
- Rely primarily on a few easily accessed types of information, without making effective or complete use of the best sources of information for a task
- Place unreasonable trust or reliance in unreliable sources of information
- Fail to filter out irrelevant or useless information
- Place too much reliance on prior knowledge
- Fail to resolve inconsistencies or fill in gaps
- Fail to properly attribute source information

Infer/Analyze:

- Fail to consider important consequences or implications of the position it takes
- Take an overly simplistic approach to the subject

Apply/Predict:

- Fail to apply its position to obvious and relevant cases
- Fail to make appropriate predictions from a known model
- Make incorrect inferences about cause and effect

Describe/Explain:

- Often fail to illustrate its ideas effectively
- Often fail to provide necessary or important generalizations to indicate how specific examples should be interpreted
- Provide vague and imprecise descriptions
- Produce implausible simulations, invented narratives, or other generative applications of a system of ideas

1 MINIMAL

A "Minimal" response typically displays one or more of the following problems:

Evaluate/Justify:

- Fail to take a stance
- Refuse even to entertain the possibility of disagreement or rebuttal
- Present evidence against its position as if it supported it
- Consistently fail to provide reasons to support the position that it takes
- Consistently repeat claims instead of supporting them
- Consistently fail to provide supporting evidence to back up its arguments

Collect/Synthesize:

- Make no serious effort to synthesize information from a range of appropriate sources
- Treat all sources of information indiscriminately as equally valid and reliable
- Present isolated bits of information side by side in undigested form
- Consistently fail to attribute sources

Infer/Analyze

- Demonstrate a serious misunderstanding of its subject matter due to failure to draw appropriate inferences.
- Consistently draw inappropriate/inaccurate inferences based upon fallacious reasoning

Apply/Predict

- Fail to apply a model or make predictions from a model where the task requires it be applied
- Succumb to critical logical errors about cause and effect and other logical fallacies in the application of mental models to specific events

Describe/Explain

- Provide inaccurate or misleading illustrations and generalizations
- Seriously misrepresent the content of summarized or paraphrased materials
- Plagiarize

Appendix A3

Evidence Model for the Competency Model Variable *Assess/Modify Documents*

Revision did not play a major role in the competency/evidence model presented in Deane et al., 2007. While *revise/edit* was a node at the lowest level, the evidence model said very little about revision activities, and the emphasis of the competency model was on planning and evaluation. There were motivations for this approach. Revision is driven by recursively evaluating and (re)planning a document, which makes revision, cognitively, a more complex process built upon more fundamental skills (Deane et al., 2008, 3–7). However, the cognitive and pedagogical literature generally suggests that learning to revise is an essential part of learning to write at a high level (Deane et al. 2008, 9–11). The importance of acquiring strategies to guide revising also emerges strongly from a review of the pedagogical literature (Graham & Harris, 2000; Graham & Perin, 2006), and thus there is strong motivation to make revision more salient in the competency and evidence models.

However, by the very nature of the task, revision is wedded to evaluation. We can only expect revision to be successful if the writer is able to assess a text, determine what is right, what is wrong, and what needs to be changed, and then apply appropriate strategies to address issues identified by the assessment. The literature cited above suggests that students engage in meaningful revision – the sort associated with high-level performance – only gradually, in an environment that strongly encourages it and models appropriate strategies, and that most early attempts at revision are focused on surface form rather than dealing with content and purpose. We have therefore restructured the competency model to highlight more strongly the relation between evaluation and revision, while keeping those revision activities that respond to critical thinking and rhetorical knowledge separate from lower-level formal correction.

4 EXCELLENT -- An "Excellent" response consistently displays the following characteristics:

Critique/Rethink:

An "Excellent" response does many of the following things, depending on the precise nature of the task:

- Demonstrate an awareness of the rhetorical purpose of a document and propose or implement changes consistent with and effectively accomplishing that purpose In accomplishing that goal it may
 - o Identify questions or objections that may occur to readers and propose or

- implement effective strategies for addressing them.
- Evaluate the rhetorical strategy employed and propose or implement effective alternative strategies, even if they require major rewrites
- Demonstrate a strong grasp of the subject matter of the document and propose or implement changes consistent with effective reasoning about that subject
 - In accomplishing that goal, it may
 - o Identify weaknesses in the reasoning underlying a document and propose or implement effective strategies for addressing them
 - Evaluate the accuracy of content, and propose or implement changes/ rewordings that maximize its validity and defensibility.
- Evince a consistent sense of how the content hangs together and thus demonstrate an ability to identify and fully correct digressions, logical leaps, gaps needing to be filled in, and the like.
- Demonstrate an awareness of the effect of stylistic choices (tone, voice) on how effectively a document achieves its rhetorical goals, and propose or implement changes in manner of presentation that effectively support those goals.

Review/Revise:

- Demonstrate an ability to consider alternative methods of organizing and presenting document content, by selecting and implementing highly effective organizations given current rhetorical goals.
- Demonstrate an ability to modify stylistic and vocabulary choices in a text effectively and systematically, so as to suit genre and task requirements.
- Demonstrate an effective sense of priority and importance of material by being able to compress or expand document content accurately and appropriately as needed to satisfy rhetorical goals
- Demonstrate an ability to identify places in a text where the flow of ideas is not clear and then effectively rephrase and recombine sentences to maximize focus and coherence.
- Demonstrate an ability to rework a text effectively to improve clarity of phrasing, and precision of word choice.

Typically, students also spend significant amounts of time rereading and revising their texts.

3 ADEQUATE -- An "Adequate" response does many of the following things, depending on the precise nature of the task, while failing to achieve the levels of performance specified for an excellent response:

Assess/Critique:

- Demonstrate an awareness of the rhetorical purpose of a document and propose or implement changes consistent with that purpose
 In accomplishing that goal it may
 - o Identify questions or objections that may occur to readers and propose or implement strategies for addressing them.
- Demonstrate comprehension of the subject matter of the document and propose or implement changes consistent with a reasonable understanding of the subject

- In accomplishing that goal, it may
- o Identify weaknesses in the reasoning underlying a document and propose or implement strategies for addressing them
- Evaluate the accuracy of content, and propose or implement changes/ rewordings that improve validity and defensibility.

Review/Revise:

- Demonstrate an ability to identify and address digressions, logical leaps, gaps needing to be filled in, and the like.
- Demonstrate some awareness of the effect of stylistic choices (tone, voice) on how effectively a document achieves its rhetorical goals, and propose or implement changes in manner of presentation that support those goals.
- Demonstrate an ability to consider and implement alternative methods of organizing and presenting document content.
- Demonstrate an ability to modify stylistic and vocabulary choices in a text to suit genre and task requirements.
- Demonstrate an ability to compress or expand document content accurately and appropriately
- Demonstrate an ability rephrase and recombine sentences in a text so as to improve focus and coherence.
- Demonstrate an ability to identify places where clarity of phrasing, and precision of word choice can be improved, and to make some improvements in same.

Typically, some time is spent on rereading and revising texts.

2 LIMITED -- A "Limited" response displays one or more of the following problems:

Assess/Critique:

- **a)** While the response shows some awareness of the rhetorical purpose of the document, it fails to address important questions or objections that may occur to readers.
- **b**) Although the response reflects a basic understanding of the subject matter of the document, it does not consistently address inaccuracies, logical fallacies or other reasoning problems.
- c) Although the response reflects a clear understanding of the document's purpose and message, it does not consistently address digressions, logical leaps, gaps needing to be filled in, etc.
- **d**) Although it may reflect some awareness of the effect of stylistic choices (tone, voice) on how effectively a document achieves its rhetorical goals, it fails to propose or implement changes in manner of presentation necessary to support those goals.

Review/Revise:

- a) It may attempt an alternative method of organizing a document, but fail to implement it adequately
- **b**) It may attempt to modify stylistic and vocabulary choices, but fail to carry through consistently
- c) It may attempt to compress or expand document content in appropriate places, but with infelicities or inaccuracies of content as a result.
- **d**) It may correctly recognize problems in phrasing and word choice, yet fail to suggest or implement appropriate changes

Typically, little time is spent on rereading and revising texts.

1 MINIMAL -- A "Minimal" response displays <u>one or more</u> of the following problems:

Assess/Critique:

- a) It indicates a failure to understand the rhetorical purpose of the document
- **b)** It indicates an inaccurate understanding of the document's subject matter.
- c) It indicates a failure to grasp what is relevant and what is irrelevant given the document's focus and purpose.
- **d**) It indicates a failure to understand how stylistic choices will affect readers' responses to the document.

Review/Revise:

- a) It does not address organization issues when a document clearly requires reorganization, or it tries to force a document into an inappropriate form or template
- **b)** It does not address obvious problems with style and vocabulary
- c) It fails to introduce (or suggest introduction of) developing material where it is clearly necessary; alternatively, it fails to compress (or suggest compression of) inappropriately long and detailed content.
- **d**) It fails to address obvious problems in phrasing and word choice

Typically, almost no time is spent rereading and revising texts.

Appendix A4

Evidence Model For the Competency Model Node Plan/Structure Documents

The greatest changes in the competency and evidence model were driven by the work in automated scoring (see Deane, Quinlan & Kostin forthcoming). As outlined above, the factor analysis in that study indicated a 2nd order factor that could be interpreted quite naturally as measuring fluency in producing complex, structured documents. The features loading on this 2nd order factor involved measures of organization, development, and sentence length. This led naturally to a reconception of the *language use* dimension in terms of the dimensions that appear to be directly measured. The criteria given below represent the construct that results.

4 EXCELLENT -- Characteristics of a typical "Excellent" response at each level

Select/Organize

• Has an effective overall structure, with content organized logically throughout. There are clear paragraph breaks; each paragraph has a clear purpose and the overall structure and purpose of the document is easily grasped. The author consistently avoids tangents, and when elaborating on subsidiary ideas, clearly indicates how they fit into the overall document structure. This may or may not be achieved by explicit thesis and topic sentences, use of explicit transition words, placement of headings, and other explicit methods of indicating overall document structure, but the main idea of the document and of its component parts are clearly indicated and easily understood.

Focus/Connect

• Is well focused and coherent, with clear indications of the relative importance of ideas and easily grasped links and transitions between ideas at sentence and clause boundaries. The writer makes effective use of complex sentence structures to indicate relative importance and logical relationships, and structures both sentences and sentence sequences, avoiding ambiguity, so that the reader is easily able to follow the flow of ideas.

Detail/Develop

• Is well developed; main ideas are substantially supported with reasons, examples, facts, or other types of elaboration

3 ADEQUATE -- Characteristics of a typical "Adequate" response:

Select/Organize

• Has a clear overall structure, with content organized logically throughout most of the composition. Where paragraph breaks are missed, other textual cues make the shift in topic clear. There may be occasional tangents or apparently irrelevant material, but not so long or so

frequent as to detract from the essential unity of the composition.

Focus/Connect

• Is focused and coherent, with a generally clear indication of how one idea relates to the next at clause and sentence boundaries. There may be places where the connection between ideas is unclear and has to be inferred, but the gaps and jumps and ambiguities are not so large or so problematic as to obscure the general intent of the text. The author makes some use of complex sentence structures and other syntactic devices to integrate and focus ideas across clauses.

Detail/Develop

• Is adequately developed; main ideas are supported with reasons, examples, facts, or other types of elaboration.

2 LIMITED -- A "Limited" response typically displays one or more of the following problems:

Select/Organize

• Is poorly structured and organized. This may involve frequent tangents, lack of an overall outline (so that the discussion wanders randomly from one point to the next), and the like. There may be no clear overall thesis, and even if there is, it may be hard to identify the topics for individual paragraphs.

Focus/Connect

• Is poorly focused and weak in coherence. Conceptual transitions between among sentences and clauses may not be explicitly marked even where the reason for the transition is hard to infer. Seemingly unimportant ideas may be emphasized by the way sentences are phrased, and there may be significant ambiguities, so that it may take some work for the reader to follow the flow of ideas.

Detail/Develop

• Is insufficiently developed, with few reasons, examples, facts, or other support

1 - MINIMAL -- A "Minimal" response typically displays <u>one or more</u> of the following problems:

Select/Organize

• Lacks structure and organization. This may take the form of a short response without paragraph structure when the task requires an elaborate document, or it may take the form of a rambling, 'stream-of-consciousness text' at greater length that is unstructured and disorganized throughout. It may be difficult to identify the main idea or subsidiary topics with any certainty.

Focus/Connect

• Lacks focus and coherence throughout. Topic shifts are unpredictable and illogical. There is little clear logical flow in which one idea follows naturally from the next even at the sentence level, and there may be serious ambiguities in expression. It may consistently use short, choppy sentences in an additive way, with no formal indication of how ideas are related across clauses.

Detail/Develop

• Is undeveloped, lacking reasons, examples, facts, or other types of support

Appendix A5

Evidence Model for the Competency Model Node Master Formal English

Given other adjustments to the competency model, and the evidence from the automated scoring analysis, there is a definite shift at this level of the competency model toward assessing mastery of formal Standard Written English, with a distinction being drawn between that ability and the ability to correct errors in written conventions. The following evidence model represents that construct:

4 EXCELLENT -- Characteristics of a typical response (one or more paragraphs) at level

Control Syntax

- Effective choice of style and register to suit the task. In particular, an excellent response
 - Effectively uses (but does not over-use or misuse) grammatical constructions typically associated with written style, such as passives, logical connectives, and attributive adjectives
 - Where appropriate, effectively uses grammatical constructions typically associated with an oral or colloquial style, such as contractions.
- A wide variety of sentence types and grammatical constructions, effectively chosen for clear communication. In particular, an excellent response
 - o Varies sentence structures effectively to maintain clarity and interest
 - o Consistently avoids unnecessarily complex or confusing sentence patterns

Control Vocabulary

- A wide range of vocabulary, precise and well-chosen. In particular, it
 - o Uses all words accurately and idiomatically to achieve clarity and interest
 - o Uses topic-specific words (Tier III words) to communicate clearly and precisely
 - o Uses words typically used in written texts (Tier II Words) where appropriate
 - Effectively uses (but does not over-use or misuse) abstract language and nominalizations
 - Effectively uses more complex, Latinate vocabulary in a way that demonstrates mastery of a variety of word-building techniques
 - o Makes effective, vivid use of common, simple vocabulary (Tier I words)

3 ADEQUATE -- An "Adequate" response displays:

Control Syntax:

- A reasonable ability to adopt a style and register appropriate to the task. In particular, an adequate response
 - O Shows reasonable control of written style, making some use of typically written constructions such as passives, logical connectives, and attributive adjectives
 - o Shows reasonable control of spoken style, making use of typically

- oral/conversational patterns only when appropriate, and consistently avoiding inappropriate use of oral patterns in written contexts.
- Reasonably well-formed sentences, varied in length and structure for clear communication. In particular, an adequate response
 - Varies word choice and sentence structures enough to avoid awkwardness and redundancy
 - o Contains relatively few sentences that are simultaneously complex and confusing

Control Vocabulary:

- A range of vocabulary, with most words used appropriately. I.e., an adequate response
 - o Rarely misuses words. When errors in word choice appear, they only occur with infrequent, academically-oriented or morphologically complex vocabulary.
 - o Uses some appropriate topic-specific vocabulary (Tier III)
 - o Makes reasonable use of at least relatively common abstract, academic, Latinate words (Tier II) without lapsing into an awkward, obscure style.
 - o Shows reasonable ability to choose clear phrasing using ordinary vocabulary (Tier I words) without lapsing into an overly repetitive or oral style.

2 - LIMITED -- A "Limited" response displays one or more of the following problems:

Control Syntax:

- Occasional lapses into an inappropriately oral style, characterized by a lack of syntactic variety, a vocabulary consisting almost entirely of ordinary everyday words, an overuse of pronouns, and/or a subjective presentation focused on inappropriate expressions of personal opinion and reaction
- Occasional lapses into an awkward, unclear style, characterized by repetitive word choice, unclear references, uneven control of sentence structure, and/or ambiguity in expression

Control Vocabulary

• Inappropriate word choices, possibly reflecting an attempt to use vocabulary above the students' normal level of expressive mastery

1 - MINIMAL -- A "Minimal" response displays one or more of the following problems:

Control Syntax

- An excessively oral style, characterized by a lack of syntactic variety, a vocabulary consisting almost entirely of ordinary everyday words, an overuse of pronouns, and/or a tendency toward a subjective presentation in which expressions of personal opinion and reaction are excessively common.
- A persistently awkward, unclear style, characterized by repetitiveness, unclear references, uneven control of sentence structure, and/or ambiguity in expression

Control Vocabulary

 Frequent, inaccurate word choices likely to cause confusion, involving vocabulary critical for the task

Appendix A6

Evidence Model for the Competency Model Node Follow Written Conventions

And finally, the last major category in the competency model is a straightforward matter of conventional correctness.

4 EXCELLENT -- Characteristics of a typical response (one or more paragraphs) at each level

- Few, if any errors in grammar [Phrase/Rephrase]
- Few, if any errors in usage for common grammatical categories (articles, prepositions, etc.) [Phrase/Rephrase]
- Few, if any errors in spelling [Proof/Correct]
- Few, if any errors in capitalization and punctuation Capitalization and punctuation patterns are both standard and consistent [Proof/Correct]
- Few errors reflecting hasty, careless text production, such as missing or repeated or inverted words [Proof/Correct]

3 ADEQUATE -- An "Adequate" response typically displays:

- Only minor errors in grammar. Those grammar errors that do appear are the sort that are in fact entirely appropriate in informal or oral contexts for Standard English, but are normally edited out from formal, written documents. [Phrase/Rephrase]
- Usage of articles, prepositions, and other common grammatical categories is generally correct and where incorrect, seldom impedes ease of reading. [Phrase/Rephrase]
- Only minor errors in spelling. Most of the spelling errors that do appear are of the sorts that reflect typographical errors rather than a lack of knowledge of how to transpose words into standard written form. There may be some errors that reflect lack of spelling knowledge for rarer words, but they should still be reasonable spellings in terms of the underlying orthographic systems of the language, with little or no confusion of common homophones [Proof/Correct]
- Few or no errors in the capitalization and punctuation necessary to indicate basic clause structure (including few or no run-ons or comma splices). Other capitalization or punctuation errors may be present as long as they do not impede ease of reading. [Proof/Correct]
- Careless errors such as missing, repeated, inverted words are not so frequent as to impede ease of reading. [Proof/Correct]

2 LIMITED -- A "Limited" response typically displays <u>one or more</u> of the following problems:

• Numerous errors in grammar, which may occasionally impede ease of reading

[Phrase/Rephrase]

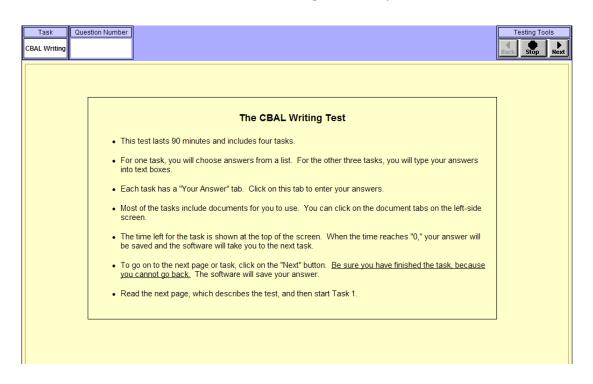
- Some errors in capitalization and punctuation that confuse basic clause structure, or else frequent minor errors in capitalization and punctuation, such as misuse of apostrophes and hyphens. [Phrase/Rephrase]
- Numerous spelling errors, reflecting incomplete mastery of English orthography. Misspellings tend to be plausible rather than confused (for instance, misspellings may often involve the wrong vowel spelling for an unstressed syllable). This may also involve frequent confusion of common homophones. [Proof/Correct]
- A significant number of careless errors, such as typos and instances where words are missing, repeated or inverted, sufficient to impede ease of reading. [Proof/Correct]
- Errors in the usage of articles, prepositions and other common grammatical categories, sufficient to impede ease of reading [Phrase/Rephrase]

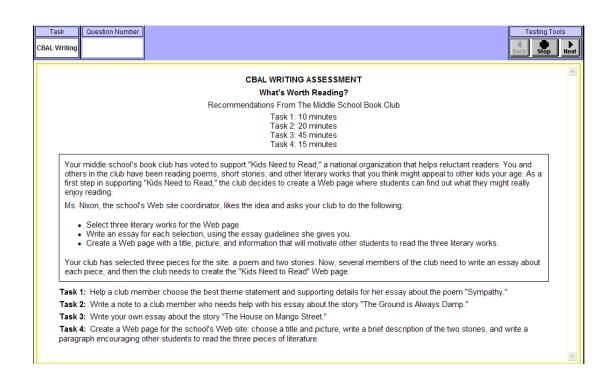
1 MINIMAL -- A "Minimal" response displays one or more of the following problems:

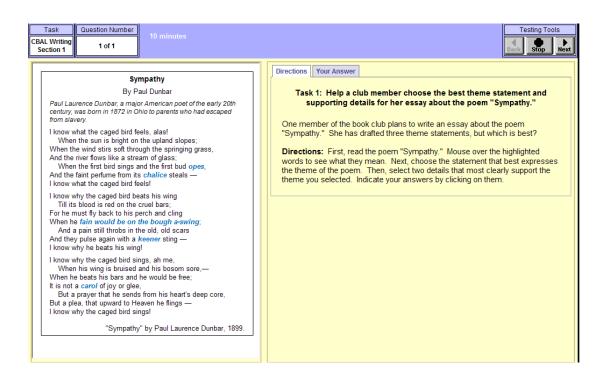
- a) Serious and pervasive ungrammaticality of the text produced [Phrase/Rephrase]
- **b)** Serious spelling errors reflecting minimal knowledge of standard English orthography, including pervasive confusion of common homophones [Proof/Correct]
- c) Consistent failure to follow fundamental conventions for punctuation and capitalization, such as a consistent failure to mark sentence boundaries with appropriate use of periods and capitals, or else pervasive minor errors in capitalization and punctuation [Proof/Correct]
- **d**) A pervasive pattern of text production errors, such as typos or instances where words are missing, repeated, inverted, or otherwise confused in ways that disrupt ease of reading. [Proof/Correct]
- e) Systematic errors in usage [Phrase/Rephrase]

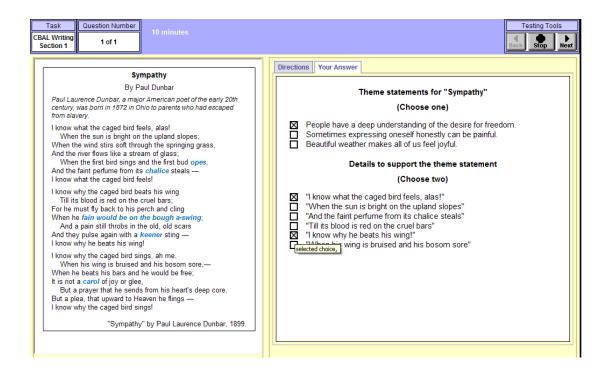
Appendix B

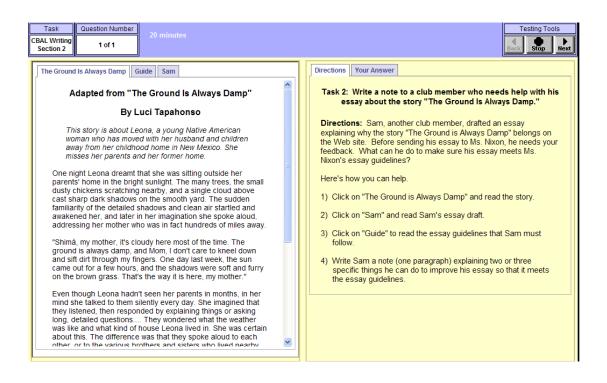
Screenshots for Test Design #1, Expository Review of Literature (Administered Fall 2008; Scoring Underway)

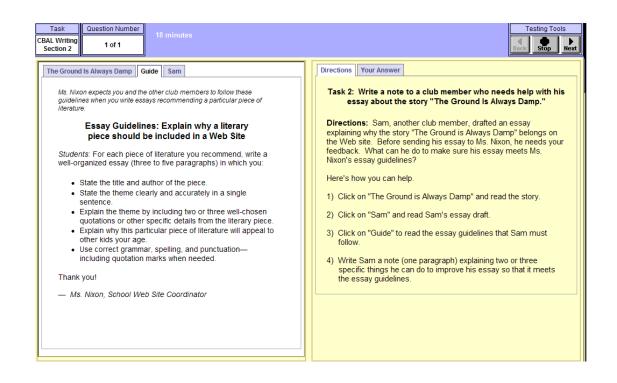


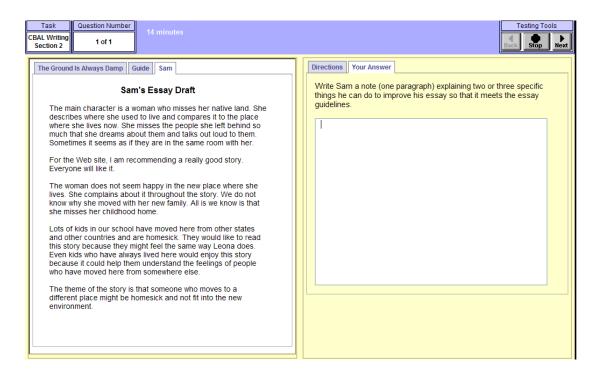


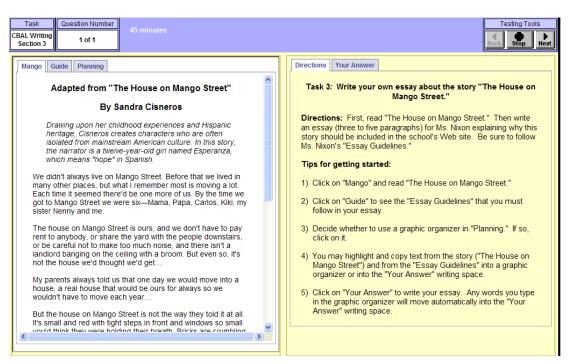


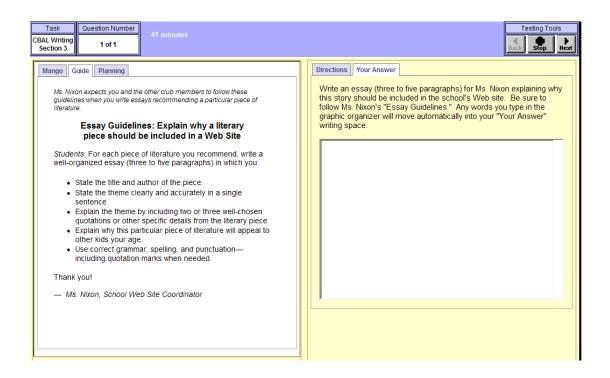


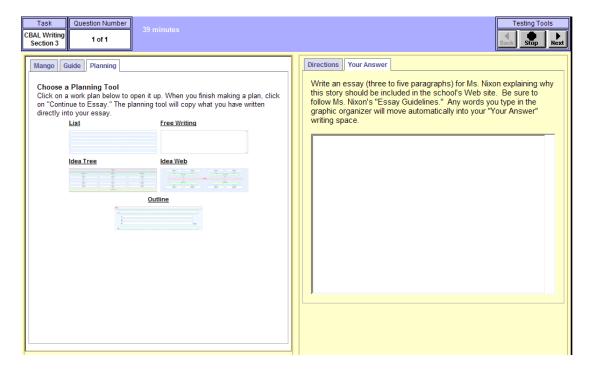




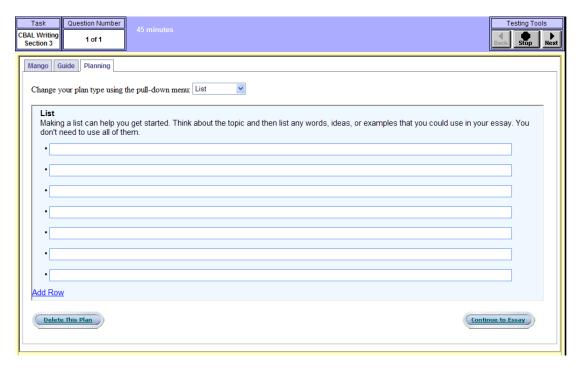




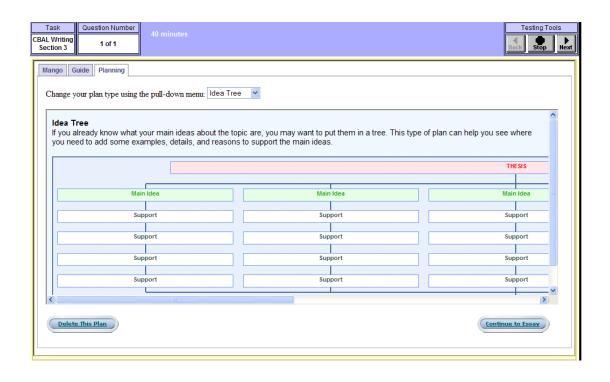


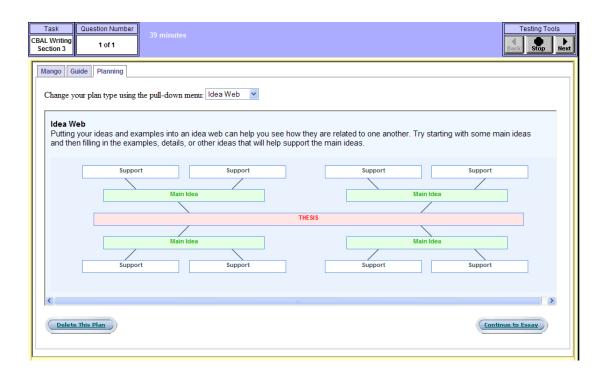


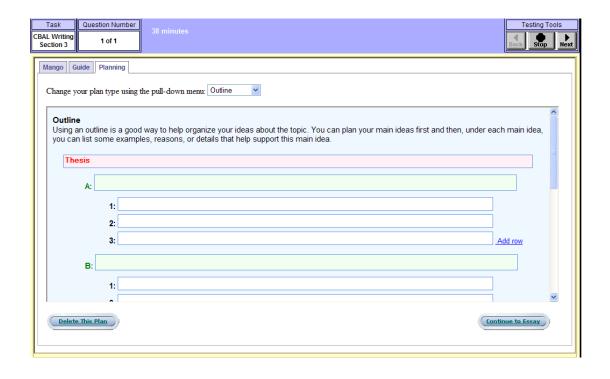
Note: The planning screen shown here and those that follow are the same on each test, but only shown for this one.

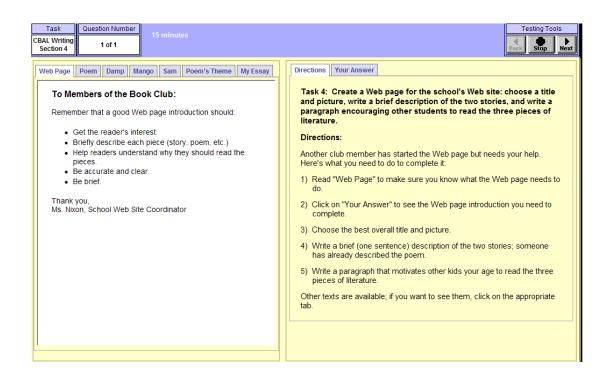


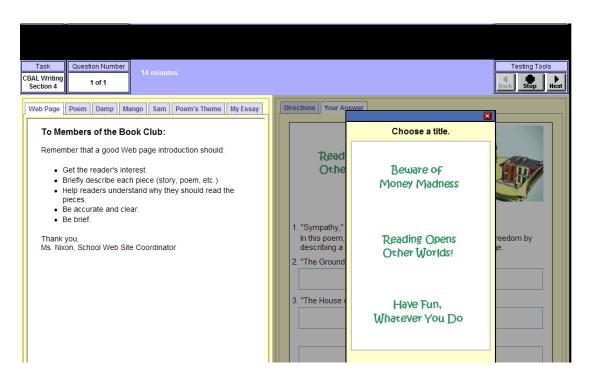




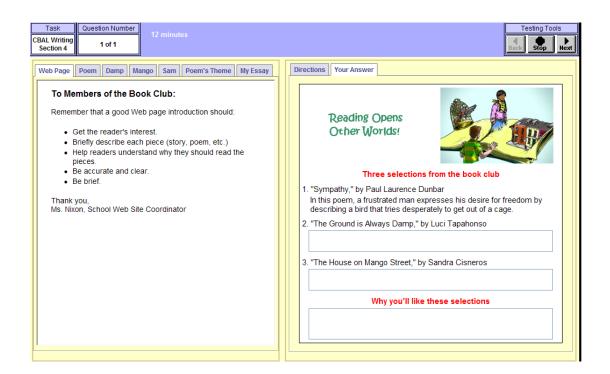












Appendix B2

Rubrics for Test Design #1 ('Book Club' Test)

'BOOK CLUB' TASK 1 – KEYS FOR Questions 1, 2, 3

- 1. Select best theme statement for poem: A, "People have a deep understanding..."
- 2. Select first supporting detail: A, "I know what..."
- 3. Select second supporting detail: E, "I know why...."

'BOOK CLUB' TASK 2 -- Write a note (one paragraph) to a peer explaining two or three specific things to do to improve the essay so that it meets the essay guidelines.

Task 2, Part A. Does the Response Address Another Student?

S	The response is clearly written to another student.
N-S	The response is not clearly written to another student – e.g., it may be written to the teacher, or the audience may be unclear.

Note: This is a yes/no decision made about all responses that are on task. It does not affect the score assigned on the main 4-point rubric for this task. Off-task responses may or may not be addressed to a student, but they are scored 0 (a) (off-task), as indicated in the 4-point rubric.

Task 2, Part B. How Well Does the Response Critique the Peer's Essay?

4 EXCELLENT --An "Excellent" response typically displays ALL of the following characteristics:

- a) Gives useful comments that are likely to improve the overall organization and at least one other problem in the peer's essay
- **b)** Has an appropriate and helpful tone

3 ADEQUATE -- An "Adequate" response typically displays ALL of the following characteristics:

- a) Gives comments that might improve (or would not weaken) at least two problems in the essay but may not address overall organization, or focuses only on the problem of organization
- **b)** Has an appropriate tone

2 LIMITED -- A "Limited" response typically displays ONE OR MORE of the following characteristics:

- a) Addresses at least one problem in the essay, but some of the advice may be confusing, irrelevant or misleading
- b) Has an inappropriate or unhelpful tone

1 MINIMAL -- A "Minimal" response typically displays ONE OR MORE of the following characteristics:

- **a**) Does not address a problem in the essay or offers fundamentally confused, inaccurate or counterproductive advice
- **b)** Has an extremely inappropriate or unhelpful tone

0 NO CREDIT -- A response receives "No Credit" if it does not allow the student's critical thinking to be judged in relation to the assigned task. For example, it might:

- a) Seriously misconstrue the task (e.g., by performing the peer's assignment rather than responding to the peer's essay)
- b) Not provide enough of the student's own writing to show understanding of the task
- c) Not be written in English
- **d**) Be off topic
- e) Be blank
- f) Consist only of random keystrokes

'BOOK CLUB' TASK 2: TOPIC NOTES

Part A: Does Response Address Student?

The response does not need to mention Sam by name, but the note should be "talking to" the person who wrote the essay. Students who write about the response (as if to a teacher) should receive a 1 on the Part A rubric.

Part B: How Well Does the Response Critique the Peer's Essay?

Overall *disorganization* of Sam's essay:

At the 4 level, responses address the *overall* disorganization of Sam's essay (as well as one other problem). Observations about the disorganization of Sam's essay may include that it lacks a clear introduction and conclusion and the paragraph transitions are choppy. The organization could be improved simply by moving the "For the Web site" and "The theme of the story..." sentences to the beginning.

Note: If a response simply identifies *part* of Sam's essay as out of order or inappropriate, the response may be credited for noting a problem, but not a problem with overall organization. (See first bullet under "Other problems" below.)

Other problems in Sam's essay include but are not limited to:

- Individual parts of the essay are out of order or contain too much, too little, or the wrong kind of information. For example, the summary of the story should not be provided in the first paragraph; the second paragraph is unconnected to the first and third paragraphs; the theme in the last paragraph should be near the story details that support it.
- The essay needs additional supporting details (e.g., details that clearly support the theme, as the theme statement at the end of Sam's essay is not clearly connected to details provided earlier in the essay).
- The title ("The Ground is Always Damp") and/or the author's name (Luci Tapahonso) is missing.
- The essay contains an error ("All is we know").
- The reference to "we" could be considered confusing.
- The essay may need quotations from the text. (While Sam does include details from the text, thus meeting the guideline requiring "two or three well-chosen quotations or other specific details," students may argue that Sam needs to include quotations as well.)

Note: Successful responses may include a minor misunderstanding of the scenario (e.g., instructing Sam to include the title and author so that other *students* will know what the book is, when Sam's audience is really Ms. Nixon) as long as the misunderstanding is not central to the responses and does not interfere with the overall helpfulness of the advice.

No Credit/Off-Task

Students who take on Sam's assignment as their own and give their impressions of "The Ground is Always Damp" should be scored as 0-OT, as they have not performed the task that was assigned.

"BOOK CLUB" TASK 3-- Explain why "The House on Mango Street" should be included on the school's Web site.

4 EXCELLENT -- An "Excellent" response typically:

- a) Gives a clear and insightful interpretation of the theme
- b) Provides a compelling explanation of why students should or might want to read the text
- c) Integrates at least two well-chosen supporting details from the text
- d) Communicates effectively for the audience

3 ADEQUATE -- An "Adequate" response typically:

- a) Gives a generally clear and reasonable interpretation of the theme
- b) Provides a credible explanation of why students should or might want to read the text
- c) Includes at least two relevant supporting details from the text
- d) Communicates appropriately for the audience

2 LIMITED -- A "Limited" response typically displays ONE OR MORE of the following characteristics:

- a) Gives a weak (unclear, incomplete, or somewhat inaccurate) interpretation of the theme
- **b**) Provides a weak (unclear, incomplete, or only somewhat credible) explanation of why students should or might want to read the text
- c) Includes little relevant or reasonable supporting material from the text
- d) At times, communicates inappropriately for the audience

1 MINIMAL -- A "Minimal" response typically displays ONE OR MORE of the following characteristics:

- a) Gives little or no interpretation of the theme, or substantially misinterprets the theme
- b) Offers no credible reason why students should or might want to read the text
- c) Includes little or no supporting material from the text
- d) Communicates extremely inappropriately for the audience

'BOOK CLUB' TASK 4

Keys to selected-response questions 1 and 2

Question 1 - Choose the best title for the Web Page: B ("Reading opens...")

Question 2 - Choose best picture for the Web Page: C (open book)

Rubric for Task 4, Part 1:

Write a brief description of the literary work for the Web Page Intro

SCORE Description of a typical response at this score level

Good 2

- a) Gives a completely relevant and accurate description of the literary work in the student's own words
- b) Gives a completely relevant and accurate description of the literary work but it has been copied from the texts provided

Weak 1

- a) Gives a fairly relevant and accurate description of the literary work in the student's own words, but may be somewhat misleading or confusing
- b) Gives a relevant and accurate description of the literary work that exceeds an appropriate length (1-3 lines) for the task
- c) Gives some relevant information about the literary work but responds to a different task (e.g., explains why students would want to read the work)

No Credit 0

Gives little, if any, relevant or accurate description of the literary work; is partly off-task; may copy irrelevant information from the texts provided

Note: While the prompt specifies one sentence, students may write more than one sentence as long as the response is reasonably brief and focused.

Critical Thinking Rubric for Task 4, Part 2: Write a paragraph encouraging other students to read the three pieces of literature.

4 EXCELLENT --An "Excellent" response typically displays ALL of the following characteristics:

- a) Is written in a highly engaging or motivating tone
- **b**) Gives clear, persuasive and text-based reasons why students might want to (or should) read the three literary works

3 ADEQUATE -- An "Adequate" response typically displays ALL of the following characteristics:

- a) Is written in an appropriate tone
- **b**) Gives generally clear, relevant and text-based reasons why students might want to (or should) read at least two of the literary works

2 LIMITED --A "Limited" response typically displays ONE OR MORE of the following characteristics:

- a) Is not written in an appropriate tone
- **b)** Gives some reasons why students might want to (or should) read at least one of the three literary works, but the reasons may be extremely limited, confusing or not clearly text-based

1 MINIMAL -- A "Minimal" response typically displays ONE OR MORE of the following characteristics:

- a) Has a confused sense of audience
- **b)** May not refer to aspects of the texts but does encourage students to read the literary works
- c) Refers to aspects of the texts but does not encourage students to read the literary works

0 NO CREDIT -- A response receives "No Credit" if it does not allow the student's critical thinking to be judged in relation to the task. For example, it might:

- Not provide enough of the student's own writing to show understanding of the task
- Not be written in English
- Be off topic
- Be blank
- Consist only of random keystrokes

"BOOK CLUB" TASK 4 – TOPIC NOTES

Write a paragraph encouraging other students to read the three pieces of literature.

PURPOSE AND AUDIENCE

The title and directions tell students that they are to write a paragraph "motivating" other students to read the three pieces. However, the Web Page Guidelines say that they are to help students "understand why they should read the pieces." To what extent does this inconsistency cause a problem in scoring?

Also, the heading for this part of the Web Page is "Why You'll like these selections." To what extent do students interpret this "you" to mean themselves? Should they be penalized for this, or can we assume that their own reasons for liking the pieces would be persuasive to other students?

CONTENT – AMOUNT OF INFORMATION

Is it reasonable to expect Since the Web Page Guidelines say to "be brief" and the three works are listed and briefly described above, we should not require students to provide an argument that refers specifically to all three literary works.

Control of Written Conventions Grammar, Usage, Mechanics, and Spelling⁴⁵

4 EXCELLENT -- Characteristics of a typical "Excellent" response

An "Excellent" response consistently adheres to the conventions of Standard Written English. Few, if any, errors are displayed in:

- a) Grammar
- b) Usage (e.g., use of articles, prepositions)
- c) Spelling
- d) Capitalization and punctuation
- e) Text production (e.g., missing words)

3 ADEQUATE -- An "Adequate" response generally adheres to the conventions of Standard Written English

It contains only minor errors, which seldom impede ease of reading. Minor errors are displayed in <u>one or more</u> of the following areas:

- a) Grammar
- **b**) Usage (e.g., use of articles, prepositions)
- c) Spelling
- d) Capitalization and punctuation
- e) Text production (missing words, etc.)

2 LIMITED -- A "Limited" response displays numerous errors in Standard Written English.

The errors may occasionally impede ease of reading. Numerous errors are displayed in <u>one or more</u> of the following areas:

- a) Grammar
- **b**) Usage (e.g., use of articles, prepositions)
- c) Spelling
- d) Capitalization and punctuation
- e) Text production (e.g., missing words)

1 MINIMAL -- A "Minimal" response displays fundamental lack of control in Standard Written English.

The errors often disrupt ease of reading. Serious and pervasive errors are displayed in <u>one or</u> more of the following areas:

- a) Grammar
- **b**) Usage (e.g., use of articles, prepositions)
- c) Spelling
- **d**) Capitalization and punctuation
- e) Text production (e.g., missing words)

0 NO CREDIT -- A response receives "No Credit" for <u>any one</u> of the following reasons:

- a) Not enough of the student's own writing for sentence-level characteristics to be judged
- **b)** Not written in English
- c) Off topic
- d) Blank
- e) Random keystrokes

MASTER FORMAL ENGLISH SYNTAX AND WORD CHOICE (VOCABULARY)

Characteristics of a typical response (one or more paragraphs) at each score level

4 EXCELLENT -- An "Excellent" response displays:

- Correct, well-formed sentences, varied in length and structure for effective communication
- A wide range of vocabulary, precise and well-chosen

3 ADEQUATE -- An "Adequate" response displays:

- Reasonably well-formed sentences, varied in length and structure for clear communication
- A range of vocabulary, with words used appropriately

2 LIMITED -- A "Limited" response displays one or more of the following problems:

- a) Little variety in sentence length and structure
- **b)** Poor word choice, possibly causing some confusion
- c) Uneven control of sentence structure/word order, which occasionally impedes coherence

1 MINIMAL -- A "Minimal" response displays one or more of the following problems:

- f) No variety in sentence length and structure
- g) Frequent misuse of words or extremely limited word choice
- h) Persistent lack of control over sentence structure, which frequently disrupts coherence

0 NO CREDIT -- A response receives "No Credit" for any one of the following reasons:

- a) Not enough of the student's own writing for sentence-level characteristics to be judged
- **b)** Not written in English
- c) Off topic
- d) Blank
- e) Random keystrokes

Appendix B3

Benchmark Papers For The 'Book Club' Test

"BOOK CLUB" TASK 2

Write a note giving Sam advice on how to improve his essay according to the essay guidelines

EXCELLENT RESPONSES - SCORE 4

Score of 4-S

Sam, your essay is very well written in some parts of it, but in other parts of it there are some flaws. To start with, you don't even have an introduction. Instead, you just start into the main character, the settin and the plot, there's no hook in it and it doesn't grab a reader's attention. If you fix that, the reader will be much more interested. Second, you have two random sentences about the website right after the first paragraph. That is in the wrong place, if you delete that it will lessen the chance of the reader getting confused. Lastly, you don't have enough about the story in your essay. You only have seven sentences about it, that's hardly a full escripti. If you get more information on the book and then talk about the website, you'll have a much better essay. I hope you fix it and hope your website stuff doesn't blow up in your face! Sincerely, The writer guy

Score of 4-S

Dear Sam, Your essay was fairly good but you need to state the title and author so other students will know what the book is so they can find it to read. Also you should have put the theme first and then explained the theme instead of explaining it and then stateing it. "For the Web site, I am recommending a really good story. Everyone will like it." You should have put his in your first paragraph at the escripti not by itself as a paragraph

ADEQUATE RESPONSES – SCORE 3

Score of 3-S

Sam, I read your essay and it was verry well written, but there are some requirerments that you left out. One thing you left out is the title, "The Ground Is Always Damp", and the author, "Luci Tapahonso". Another thing I notice that you left out are the specific lines and examples from the story. If I were you, I would state some things that Leona was saying to her parents like, "it's cloudy here most of the time." And, "there are no mountains here". I hope that I have helped you to improve your essay.

Score of 3-S

Sam, think about the guide then look at your essay, are there differences? Yes, one you didn't state the authors name or title. Why did you put the website recommendation in the middle of your essay? Sam, you did not use quotes to support your theme, let alone the fact that your theme partially meets the standard for a well 97escrip theme. A well 97escrip essay includes all the I, and you only included some of the I.

Score of 3- NS

I think that sam should have followed the Essay I more. He had made many errors. He should have stated the title and author of the piece. The theme of the I should be in the I of the Essay. Lastly his grammer needs to be fixed. One sentence says "All is we know is that she misses her childhood home.

LIMITED RESPONSES - SCORE 2

Score of 2-S (a)

One thing is to check your grammer. Grammer is important I if you have bad grammer your story will sound bad. You should tell what see will miss in her home, Like what is spiecal in he house that see charised. Support you theme statement, with quotes. Try thinking of a title for the story.

Score of 2-S (a)

Dear Sam, In your essay draft you need to delete "for the website, I am recommending a really good story, everyone will like it," and put "I think This is a good story, you would like the story if you like to read." But, don't say "for the website" if you are not going to give the website. Delete "she complains about it throughout the story." You need to add the Author and the title of the story. You need to add two or three well-chosen quotes. Instead of saying all is we know, instead say all that I know is that..... What age group is this story for? sincerely, Note: This response gives some misleading information about the Web site. We have inserted quotation marks for ease of reading.

MINIMAL RESPONSES - SCORE 1

Score of 1-S (a)

Maybe add a little more so I can understand what lenoa is and what happen next.

Score of 1-S (a)

Sam, try to descrip everything on the guide to the fullest, so the teacher has no doubts about your paper. Especially when she asks for a hidden theme, like relating the story to other kids. Also reread your writing because some parts didn't make any sense, but could have if you fixed them. If you want to get an even better grade, use more descriptive words. Those are my suggestions to you.

NO CREDIT RESPONSE - SCORE 0

Score of 0-(a)

"The Ground Is Always Damp" By Luci Tapahonso This story is about a girl named Leona who moved away from her home town in New Mexico. Kids my age would like this story because every kid goes throught this,like when they first start going to sleepovers, they get home sick.

'BOOK CLUB' TASK 3

Explain why "The House on Mango Street" should be included on the school's Web site.

EXCELLENT RESPONSE - SCORE 4

"The House on Mango Street" reflects the life and soul of a young girl who has had to cope with change all her life. Anyone who has had to deal with the bitterness of life will easily be able to relate with the narrator, Esperanza. However, many students who are not interested in short stories without much dialog may not find this story very interesting. This is why "The House on Mango Street" should be included on our Web Site, to prove those students wrong.

Life for those without much opportunity or resources, like Esperanza and her friends and family, can often be hard. However, if you have the will to press on, you will go far beyond your expectations. Once students our age complete this story and all of its short passages, they will be able to apply this message to their everyday life. When Esperanza describes her need for a house, a "real house," it is at this point that you start to realize you may have more in common with the narrator than you thought. When Esperanza takes us into the lives of her neighbors, it makes you stop to think about what kinds of stories are yet to be found in the people around you.

If this story is included in our Web Site, we can show students that this is a story worth reading. A story for all who long for a dream house with "running water" and "a great big yard with grass growing without a fence," but instead find themselves in a place very much like Mango Street.

ADEQUATE RESPONSE - SCORE 3

This very nice story called "The House on Mango Street" should be included in the school's literary web site. It is a short piece by Sandra Cisneros. It is about this family who is always moving and can't find a right house to live in so they move into the house on Mango Street. However, not everyone is happy with the move.

The family's ambition is to move into a big house with at least three washrooms and running water. They wanted a white house with trees around it and a huge yard. This was the way they dreamed their house to be. However, the house they moved into was not like that. There is hardly any yard and only one washroom. Everyone in the house has to share one bedroom. The bricks are falling apart and the front door is very rust and worn. The main character in the house wants a better house. It says, "I want a real house. One I could point to. The house on Mango Street isn't it."

"The House on Mango Street" appeals to other kids my age because they might be feeling the same way about their house too. They might not like their house and be too embarrassed to tell other kids where they lived. Kids may feel to shy to tell others how they feel. So reading about people with the same problems may make them feel better. They might realize that they are not the only one with the same problem. It may help them share more about how they are feeling. As you can see know, I strongly suggest to put this story on the school's literary web site. It can help other kids and appeal to them.

ADEQUATE RESPONSE - SCORE 3 - 2nd example

She wants an oasis. A place where she doesn't have to worry about being to loud, or having mean neighbors. A house with a big front yard and a laundry room that she doesn't have to share with other people. A house she won't be embarassed to point to when someone asks where she lives. She doesn't want to move almost every year. She wants a stable and peaceful life with few disapointments and a place to call her own. But don't we all.

The House on Mango Street was written by Sandra Cinsneros about a 12-year-old Hispanic girl and her family moving from apartment to apartment when all she wants is a home to call her own. She doesn't want to worry about being broken into by a random person off the streets because the front door isn't locked. Esperanza and her family lived in an apartment that got broken into, the person broke into the laundry room one night.

Two days after the break in while Esperanza was standing outside her teacher walked by. This is the conversation they had.

"Where do you live?" she asked.

"There, I said pointing up to the third floor." You live there?"

There. I had to look to where she pointed—the third floor, the paint peeling, wooden bars

Papa had nailed on the windows so we wouldn't fall out. You live *there*? The way she said it made me feel like nothing. *There*. I lived *there*. I nodded.

This conversation really got Esperanza thinking about how she wanted a house to point to and call her own. A place she can be proud of and not have to worry about what people think about her house and her family.

Kids my age would like this book because most of them could relate. Today there are more people coming from the United States from other countries. Also because everyone wants to be proud of what they have and not feel insignificant about the life style there living.

LIMITED RESPONSE - SCORE 2 b

The story i read was called "The house on Mango Street" by Sandra Cisneros The theme of this book is wanting a real house that she can be proud of.

For example: When her teacher comes and says, "You live there?" Esperanza felt so awful, she wanted a real home she could be proud of. Also how she explained her house on Mango Street. "It's small and red with tight steps in front and windows so small you'd think they were holding their breath. Bricks are crumbling in places, and the front door is so swollen you have to push hard to get in. There is no front yard...."

This book will apeal to other kids my age because I am sure there are many kids like Esperanza. Many even, at my school.

MINIMAL RESPONSE - SCORE 1 a, b

Many things in life change; Nothing always stays the same. For example some people move a lot or get a new family member; stuff like that. Sometimes it takes longer for change to happen.Like in the story: The House on Mango Street By Sandra Cisneros the main character; a 12 year old girl named Esperanza wants to change the place she lives in. In the next few sentences i will be explaining it more clearly to you.

We all have something we don't like in life. In the story Esperanza doesn't like the her house. Teacher saying "You live in there?" Pointing to her house. We all have dreams of either having a new house making the most friends becoming something in life. Sometimes those dreams set goals in our lives. Esperanza's parents want a different house a better life and it is. Slowly and slowly things in life change. They no longer live in that old house. They no longer have a landlord/lady.

It's hard for some people to get used to the change. For others it's for the better.

No matter how hard you try to keep things the same they never do!!!

'BOOK CLUB' TASK 4 - Question 3

Write a sentence describing "The Ground is Always Damp"

GOOD DESCRIPTIONS - SCORE 2

Score of 2 (a)

This story is about a woman leaving her native land to be with her husband and children but she so dearly misses her family and house that she has grown up in all her life.

Score of 2 (b)

This story is about Leona, a young Native American woman who has moved with her husband and children away from her childhood home in New Mexico. She misses her parents and her former home.

WEAK DESCRIPTIONS - SCORE 1

Score of 1 (a)

In this story, a women missing her family, starts to talk to herself

Score of 1(b)(c)

In this story and Native American woman talks about her feeling and where she lives with her family and she talks out loud thinking that her mama and papa and other siblings will hear her. I think you shold read this story because it might realate to you when you are homesick maybe you have the same feelings and do things the same as this Native American.

Score of 1 (c)

Ever feel homesick? Well read this story and you might be able to compare with this native american. This story is about a native american that feels homesick. She wants to go back home but decides not to.

'BOOK CLUB' TASK 4 - Question 4

Write a sentence describing "The House on Mango Street"

GOOD DESCRIPTIONS - SCORE 2

Score of 2 (a)

This story is about a little girl that has to live in appartment after apartment until her family can finally aford a house of their own. But the house is all run down, and not what she expected

Score of 2 (b)

In this story a 12 year old girl longs to have a house. One she can point to when people ask he where she lives.IN the end she get a house but it is not the one she want to have, it was not the way she had planned for.

WEAK DESCRIPTIONS - SCORE 1

Score of 1 (a)

The house on mango street is a nice, sad, and funny comedy about a girl named Esperanza, and she moves a lot in the story. I think anyone will enjoy this.

Score of 1 (b)(c)

Thats not what i expected. This story is about a young girl with a family of 5 including her. They hear about a house a buy it, the house is not what she expected. The girl just wanted a place to call home because shes moved every year of her living{12 years}. Will she learn to live with? Read this and find out

'BOOK CLUB' TASK 4 -- Question 5

Why you'll like these selections

EXCELLENT RESPONSES – SCORE 4 (CRITICAL-THINKING SKILLS)

Score of 4

These selections are wonderful, they all make you think again. One is about someone wanting freedom, one missing her homeland, and another who wants a home to call in really believe that this is her home. Anyone can take a part in the selections. We have all been there or felt like this people have. Reading them can give you the thought that you're not alone.

Score of 4

You will like these selections because they are very heartfelt and moving, so to speak. They make you think of what you have and how some people have nothing at all. If you like stories about the needy and hopeful, you will like these literature pieces. You may be thinking, "Ew!! I hate literature!" But they don't seem like literature at all. They sound like normal stories. They are great stories indeed, and I hope you will take time to read them.

ADEQUATE RESPONSES – SCORE 3

(CRITICAL-THINKING SKILLS)

Score of 3

You would like these storyies because they may relate to something in your life. The house on mango street could make you realize that your life isnt all that bad compared to Esperenza. Also it could be interesting because she moves alot, and maybe you do to. In the ground is always damp, the girl misses her old home. She misses everything about it including its smell. She is homesick and also misses her parents. Maybe your parents arent around or you miss your old house and friends. You would enjoy these stories because there about kids that you could possibly relate to.

Score of 3

These are quality stories that can relate to you. If any of you have ever moved and felt homesick, The Ground is Always Damp is a perfect story for you. If you have ever rented a house or apartment or even have a house that is not ideal, The House on Mango Street is the book for you.

LIMITED RESPONSES – SCORE 2 (CRITICAL-THINKING SKILLS)

Score of 2 (b)

10617

These stories all are different in their own ways. They each can be expeirences that YOU can relate to. They are a great way to spend some time when your board and long for a story that will trigger your imagination.

Score of 2 (b)

11048

Readers'll like these selections because all of the characters have... character. They are great stories and readers will hopefully be able to connect to the books in more than one way.

MINIMAL RESPONSES – SCORE 1 (CRITICAL-THINKING SKILLS)

Score of 1 (b)

04102

the reason why you will like these sections are because they are good stories or peoms.

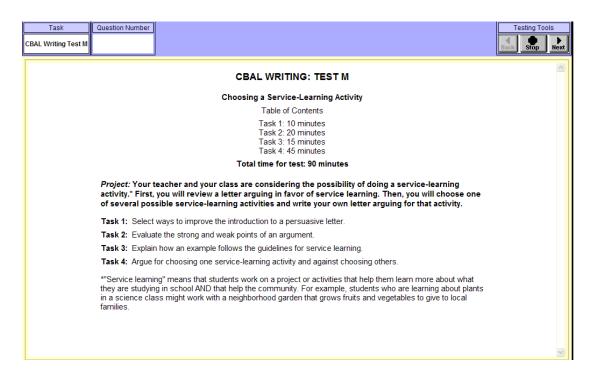
Score of 1 (b)

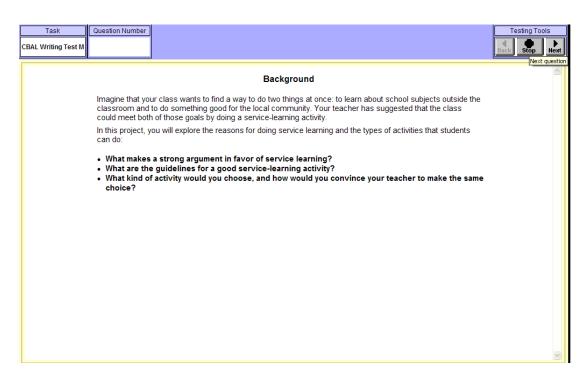
10908

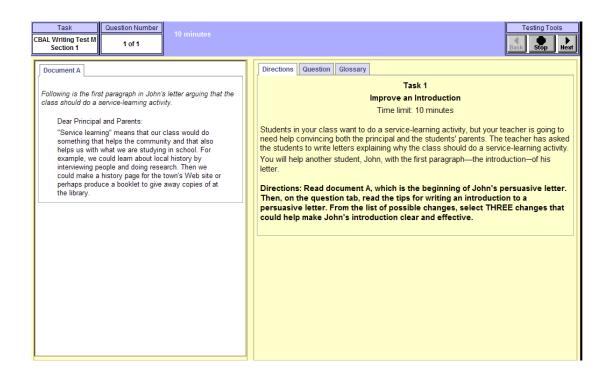
I would like other kids to pick these because they are interesting peices of writing and sometimes it is fun to read even if its boring.

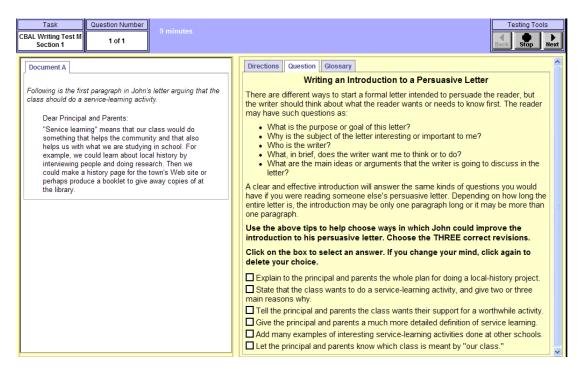
Appendix C1

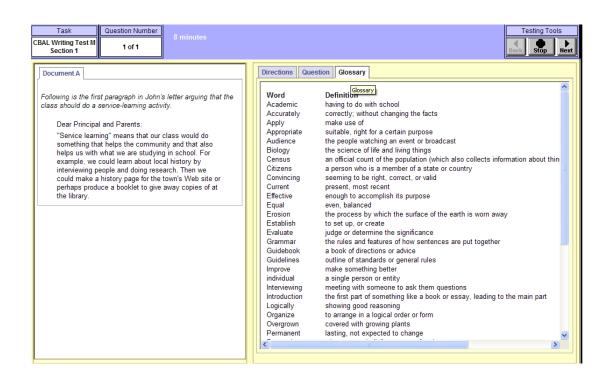
Screenshots for Test Design #2, Persuasive Memorandum (Service Learning Administered Fall 2007; Scored/Analyzed In 2008)

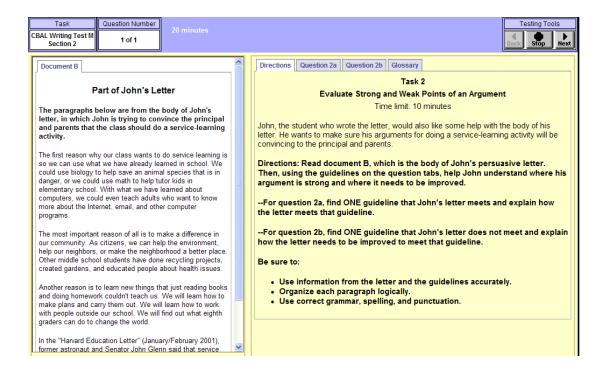


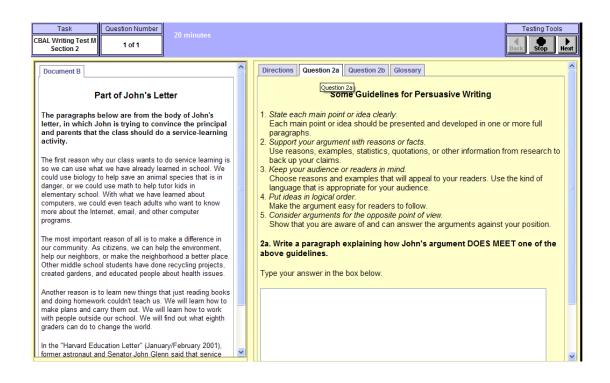


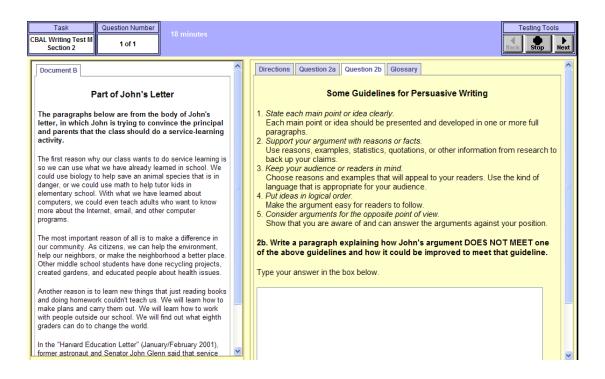


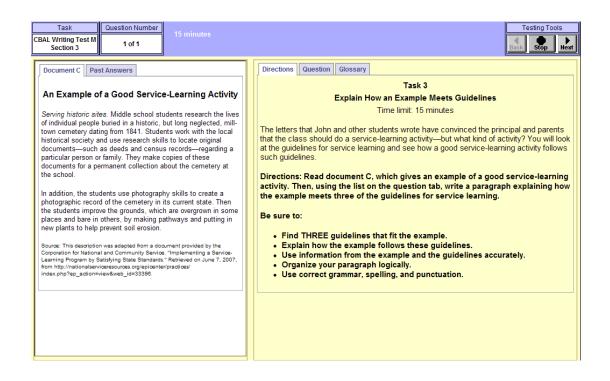


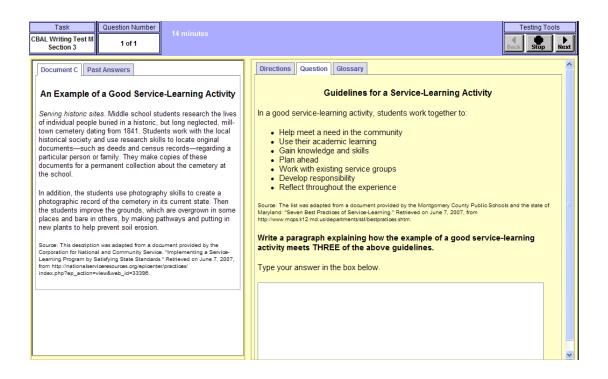












Appendix C2
Frequency Tables for Test Design #2

			Cumulative	
Task 1	Frequency	Percent	Frequency	Cumulative Percent
0	4	8.89	4	8.89
1	17	37.78	21	46.67
2	21	46.67	42	93.33
3	3	6.67	45	100.00

Task 2 Strand I	Frequency	Percent	Cumulative Frequency	Cumulative Percent
0.5	1	2.22	1	2.22
1	9	20.00	10	22.22
1.5	6	13.33	16	35.56
2	12	26.67	28	62.22
2.5	4	8.89	32	71.11
3	8	17.78	40	88.89
3.5	4	8.89	44	97.78
4	1	2.22	45	100.00

Task 2a Strand II	Frequency	Percent	Cumulative Frequency	Cumulative Percent
0	1	2.22	1	2.22
0.5	1	2.22	2	4.44
1	12	26.67	14	31.11
1.5	4	8.89	18	40.00
2	11	24.44	29	64.44
2.5	4	8.89	33	73.33
3	9	20.00	42	93.33
3.5	1	2.22	43	95.56
4	2	4.44	45	100.00

Task 2b Strand II	Frequency	Percent	Cumulative Frequency	Cumulative Percent
0	6	13.33	6	13.33
0.5	2	4.44	8	17.78
1	5	11.11	13	28.89
1.5	6	13.33	19	42.22
2	16	35.56	35	77.78
2.5	2	4.44	37	82.22
3	5	11.11	42	93.33
3.5	1	2.22	43	95.56
4	2	4.44	45	100.00

Task 2a Strand III	Frequency	Percent	Cumulative Frequency	Cumulative Percent
0	1	2.22	1	2.22
0.5	3	6.67	4	8.89
0.75	2	4.44	6	13.33
1	1	2.22	7	15.56
1.25	2	4.44	9	20.00
1.5	2	4.44	11	24.44
1.6666666667	1	2.22	12	26.67
1.75	4	8.89	16	35.56
2	8	17.78	24	53.33
2.25	4	8.89	28	62.22
2.5	5	11.11	33	73.33
2.75	5	11.11	38	84.44
3	3	6.67	41	91.11
3.25	2	4.44	43	95.56
4	2	4.44	45	100.00

Task 2b Strand III	Frequency	Percent	Cumulative Frequency	Cumulative Percent
0	6	13.33	6	13.33
0.5	2	4.44	8	17.78
0.75	1	2.22	9	20.00
1	1	2.22	10	22.22
1.25	2	4.44	12	26.67
1.33	1	2.22	13	28.89
1.5	3	6.67	16	35.56
1.75	5	11.11	21	46.67
2	4	8.89	25	55.56
2.25	2	4.44	27	60.00
2.5	5	11.11	32	71.11
2.75	3	6.67	35	77.78
3	4	8.89	39	86.67
3.25	4	8.89	43	95.56
3.5	1	2.22	44	97.78
4	1	2.22	45	100.00

Task 3 Strand I	Frequency	Percent	Cumulative Frequency	Cumulative Percent
0	2	4.44	2	4.44
0.5	3	6.67	5	11.11
1	5	11.11	10	22.22
1.5	5	11.11	15	33.33
2	11	24.44	26	57.78
2.5	9	20.00	35	77.78
3	6	13.33	41	91.11
3.5	3	6.67	44	97.78
4	1	2.22	45	100.00

Task 3 Strand II	Frequency	Percent	Cumulative Frequency	Cumulative Percent
0	4	8.89	4	8.89
0.5	1	2.22	5	11.11
1	6	13.33	11	24.44
1.5	4	8.89	15	33.33
2	13	28.89	28	62.22
2.5	4	8.89	32	71.11
3	8	17.78	40	88.89
4	5	11.11	45	100.00

Task 3 Strand III	Frequency	Percent	Cumulative Frequency	Cumulative Percent
0.5	1	2.22	1	2.22
0.67	1	2.22	2	4.44
1	12	26.67	14	31.11
1.33	2	4.44	16	35.56
1.5	5	11.11	21	46.67
2	4	8.89	25	55.56
2.5	8	17.78	33	73.33
3	7	15.56	40	88.89
3.33	1	2.22	41	91.11
3.5	1	2.22	42	93.33
4	3	6.67	45	100.00

Task 4 Strand I	Frequency	Percent	Cumulative Frequency	Cumulative Percent
0	2	4.44	2	4.44
1	5	11.11	7	15.56
1.5	6	13.33	13	28.89
2	7	15.56	20	44.44
2.5	12	26.67	32	71.11
3	4	8.89	36	80.00
3.5	5	11.11	41	91.11
4	4	8.89	45	100.00

Task 4 Strand II	Frequency	Percent	Cumulative Frequency	Cumulative Percent
0	1	2.22	1	2.22
1	6	13.33	7	15.56
1.5	5	11.11	12	26.67
2	9	20.00	21	46.67
2.5	7	15.56	28	62.22
3	7	15.56	35	77.78
3.5	8	17.78	43	95.56
4	2	4.44	45	100.00

Task 4 Strand III	Frequency	Percent	Cumulative Frequency	Cumulative Percent
0	1	2.22	1	2.22
0.5	1	2.22	2	4.44
1	8	17.78	10	22.22
1.5	2	4.44	12	26.67
2	14	31.11	26	57.78
2.5	5	11.11	31	68.89
3	9	20.00	40	88.89
3.5	2	4.44	42	93.33
4	3	6.67	45	100.00

Appendix C3

Rubrics for the Service Learning Test

Note: the rubrics for Strand I (Control of Sentence Structure, Grammar, Usage, Mechanics, Word Choice (Vocabulary), and Spelling) and Strand II (Organization and development) were constant across the pilots for the Service Learning, 'Children and Sleep and Advertising to Children assessments.

CONTROL OF SENTENCE STRUCTURE, GRAMMAR, USAGE, MECHANICS, WORD CHOICE (VOCABULARY), AND SPELLING

Characteristics of a typical response (one or more paragraphs) at each score level

4 EXCELLENT -- An "Excellent" response displays:

- Correct, well-formed sentences, varied in length and structure for effective communication
- A wide range of vocabulary, precise and well-chosen
- Few, if any, errors in grammar and mechanics, and spelling

3 ADEQUATE -- An "Adequate" response displays:

- Reasonably well-formed sentences, varied in length and structure for clear communication
- A range of vocabulary, with words used appropriately
- Only minor errors in grammar, mechanics, and spelling—not serious enough to impede ease of reading

2 LIMITED -- A "Limited" response displays one or more of the following problems:

- Poor word choice, possibly causing some confusion
- Little variety in sentence length and structure
- Numerous errors in grammar and mechanics, which may occasionally impede ease of reading
- Numerous spelling errors, which occasionally impede ease of reading
- Uneven control of sentence structure/word order, which occasionally impedes coherence

1 MINIMAL -- A "Minimal" response displays one or more of the following problems:

• Frequent misuse of words or extremely limited word choice

- No variety in sentence length and structure
- Serious and pervasive errors in grammar and mechanics, which frequently disrupt ease of reading
- Serious spelling errors, which frequently disrupt ease of reading Persistent lack of control over sentence structure, which frequently disrupts coherence

- Not enough of the student's own writing for sentence-level characteristics to be judged
- Not written in English
- Off topic
- Blank
- Random keystrokes

ORGANIZATION AND DEVELOPMENT

Characteristics of a typical response (one or more paragraphs) at each score level

4 EXCELLENT -- An "Excellent" response:

- Has an effective overall structure, with content organized logically throughout
- Is well focused and coherent, with a clear relationship between main and subordinate ideas and clear transitions and connections between ideas
- Is well developed; main ideas are substantially supported with reasons, examples, facts, or other types of elaboration

3 ADEQUATE -- An "Adequate" response:

- Has a clear overall structure, with content organized logically throughout most of the composition
- Is focused and coherent, with a generally clear relationship of main/subordinate ideas and appropriate transitions and connections between ideas
- Is adequately developed; main ideas are supported with reasons, examples, facts, or other types of elaboration

2 LIMITED -- A "Limited" response displays one or more of the following problems:

- Is poorly structured and organized
- Is poorly focused and weak in coherence
- Is insufficiently developed, with few reasons, examples, facts, or other types of support

1 MINIMAL -- A "Minimal" response displays one or more of the following problems:

- Is unstructured and disorganized throughout
- Lacks focus and coherence throughout
- Is undeveloped, lacking reasons, examples, facts, or other types of support

- Not enough of the student's own writing for document-level characteristics to be judged
- Not written in English
- Off topic
- Blank
- Random keystrokes

SERVICE LEARNING TASK 1 -- Improve an Introduction

Use the above tips to help choose ways in which John could improve the introduction to his persuasive letter. Choose the THREE correct revisions.

- 1. __ Explain to the principal and parents the whole plan for doing a local-history project.
- 2. _X_ State that the class wants to do a service-learning activity, and give two or three main reasons why.
- 3. _X_ Tell the principal and parents the class wants their support for a worthwhile activity.
- 4. __ Give the principal and parents a much more detailed definition of service learning.
- 5. __ Add many examples of interesting service-learning activities done at other schools.
- 6. _X_Let the principal and parents know which class is meant by "our class."

SERVICE LEARNING TEST M – TASK 2, PART A

Explain the Strong Points of an Argument

Task Directions: Find one guideline that John's argument meets and explain how his argument meets that guideline.

4 EXCELLENT -- An "Excellent" response:

- Identifies one of the guidelines
- Provides a well-reasoned, insightful explanation of how the argument meets the guideline
- Accurately uses specific information from the letter and guideline to provide a persuasive analysis

3 ADEQUATE -- An "Adequate" response:

- Identifies one of the guidelines
- Provides a reasonable explanation of how the argument meets the guideline
- Accurately uses relevant information from the letter and the guideline in the analysis

•

2 LIMITED -- A "Limited" response identifies one of the guidelines but displays one or more of the following problems:

- a) Gives a poor explanation of the argument's relationship to the guideline
- **b)** Refers to the argument or guideline inaccurately
- c) Is somewhat vague or confusing in its reasoning

1 MINIMAL -- A "Minimal" response displays one or more of the following problems:

- a) Does not identify one of the guidelines
- b) Does not accurately explain the argument's relationship to the guideline
- c) Seriously misrepresents the argument and/or the guideline
- d) Is often vague or confusing in its reasoning

- a) Not long enough for critical-thinking skills to be judged
- **b)** Not written in English
- c) Off topic
- d) Blank
- e) Only random key strokes

SERVICE LEARNING TEST M – TASK 2, PART B

Explain the Weak Points of an Argument

Task Directions: Find one guideline that John's argument fails to meet, and explain how to improve the argument to meet the guideline.

4 EXCELLENT -- An "Excellent" response:

- Identifies one of the guidelines
- Provides a well-reasoned, insightful explanation of how to improve the argument to meet the guideline
- Accurately uses specific information from the letter and guideline to provide a persuasive analysis

3 ADEQUATE -- An "Adequate" response:

- Identifies one of the guidelines
- Provides a reasonable explanation of how to improve the argument to meet the guideline
- Accurately uses relevant information from the letter and the guideline in the analysis

2 LIMITED -- A "Limited" response identifies one of the guidelines but displays <u>one or more</u> of the following problems:

- a) Gives a poor explanation of the relationship between the argument and the guideline
- **b)** Refers to the argument or guideline inaccurately
- c) Is somewhat vague or confusing in its reasoning

1 MINIMAL -- A "Minimal" response displays one or more of the following problems:

- a) Does not identify one of the guidelines
- b) Does not accurately explain the relationship between the guideline and the argument
- c) Seriously misrepresents the argument and/or the guideline
- **d**) Is often vague or confusing in its reasoning

- a) Not long enough for critical-thinking skills to be judged
- **b)** Not written in English
- c) Off topic
- d) Blank
- e) Only random key strokes

SERVICE LEARNING TEST M - TASK 3

Explain How an Example Meets Guidelines

Task Directions:

Find three guidelines that fit the example and explain how the example follows these guidelines

4 EXCELLENT -- An "Excellent" response:

- Identifies three guidelines that fit the example
- Gives a well-reasoned, insightful explanation of how the example meets each guideline
- Accurately uses specific information from the example and guideline to provide a persuasive analysis

3 ADEQUATE -- An "Adequate" response:

- Identifies three guidelines that fit the example
- Gives a reason
- able explanation of how the example meets each guideline
- Uses relevant information from the example and the guideline in the analysis

2 LIMITED -- A "Limited" response displays one or more of the following problems:

- a) Identifies only two guidelines, not three
- **b**) Gives a reasonable explanation of the relationship between the example and only two guidelines
- c) Uses relevant information but is somewhat vague or confusing, with little analysis

1 MINIMAL -- A "Minimal" response displays one or more of the following problems:

- a) Identifies only one guideline
- b) Gives a reasonable explanation of the relationship between the example and only one guideline
- c) Uses some relevant information but is often vague or confusing, or has no analysis

- a) Not long enough for critical-thinking skills to be judged
- **b)** Not written in English
- c) Off topic
- d) Blank
- e) Only random key strokes

Service Learning TEST M – TASK 4

Argue for Choosing One Service-Learning Activity and Against Choosing Two Others

4 EXCELLENT -- An "Excellent" response:

- Clearly identifies one of the activities for the class to do
- Argues in favor of the selected activity by using insightful, highly persuasive reasoning
- Argues *against* the other two activities by using insightful, highly persuasive reasoning
- Is accurate in its use of facts from the documents

3 ADEQUATE -- An "Adequate" response:

- Clearly identifies one of the activities for the class to do
- Argues in favor of the selected activity by using relevant reasons
- Argues *against* the other two activities by using relevant reasons
- Is generally accurate in its use of facts from the documents

2 LIMITED -- A "Limited" response displays one or more of the following problems:

- a) Does not explicitly identify one of the activities but implies which it should be
- **b)** Gives generally weak reasons to support the choice of activities
- c) Gives generally weak reasons to oppose the other two
- d) Contains inaccuracies in its use of facts
- e) Argues for one activity but neglects to oppose the other two

1 MINIMAL -- A "Minimal" response displays one or more of the following problems:

- a) Does not identify one of the activities for the class to do
- **b)** Gives few, if any, reasons to explain why the class should do one activity instead of the others
- c) Presents a seriously incoherent argument or fails to make any argument at all
- **d)** Is seriously inaccurate in its use of facts
- e) Is often vague or confusing in its reasoning

- a) Not long enough for critical-thinking skills to be judged
- **b)** Not written in English
- c) Off topic
- d) Blank
- e) Only random key strokes

Appendix C4a

Benchmark Responses Illustrating Sentence-Level Skills

SCORE 4 – EXCELLENT RESPONSE

(SENTENCE-LEVEL SKILLS)

TestM_S4Q1

Dear teacher.

In my opinion, activity A is the most reasonable and managable activity for the students to do. In performing songs with senior citizens, students would get to speak to the kinds of people they would not normally get to speak to. It would give them a new perspective on the world to hear senior citizens talk about what it was like when they were the students' age. Also, the students would get to learn new songs, and engage in the culture of music, which is a very good type of culture for students to be exposed to. Music creates emotions and feeds the spirit. It can teach empathy and it can change you all at once from being sad to happy of viseversa. Some of the senior citizens are highly skilled musicians, so the students would get to learn from them. Some of the students in the class do not sing or play an instrument, but that would be all right, because they would get to learn from the senior citizens. The senior citizens would also enjoy passing on their knowledge to young people. The issue of finding times when both groups can rehearse together is not an unbeatable one. Students could simply go up to the nursing home during school hours when the senior citizens are not busy.

I do not like the other two options as much. The main problem with activity B is that there is already a group picking up litter at the park. As long as the litter is being picked up, it doesn't

matter whose doing it. The students might not find very much trash to pick up, and they might

instead fool around. The other problem with option B is that the students aren't even learning about biology this year. So why would they do a project that is based on biology? People might think that students would enjoy taking water samples and writing up the results, but for

one year, that was all my class did, and let me tell you...it was duller than watching paint dry. Students would get bored with this option and probably resort to playing in the park.

As for activity C, it is a very good idea, and I would support it, except that it would not be accessable to very much of the student population. Only three groups of foreign students would be able to benefit from this option, and I do not think this is fair to the other new students. They might feel like they are being snubbed, or that they are not welcome in their new school. All students should have an equal chance at success in school. In conclusion, I think that activity A is the best activity for the students to do.

SCORE 3 – ADEQUATE

(SENTENCE-LEVEL SKILLS)

TestK_S3Q1

Befor reading these documents I thought that some ads were not good for children. After reading Documents A and B, I thought that ads were bad for kids, but after reading Documents C and D, I thought that ads were ok. Both are true, but it kind of depends on what the kids have for parents and friends. If a child has really good parents that talk to their kids and teach them many things, ads can be ok because they will know how the media is persuading them to buy things.

If a child has parents that don't really talk to them as much and don't explain things to them, then ads are not so good. Children that don't have parents to teach them about ads will be persuaded not only by the ad, but by their friends as well. If a child has friends that are buying or doing something, they will want to do the same because thats what they feel will make them be "cool".

Children that don't understand how the media is persuading them to buy things, will want something because they like the ads for it. If a child sees a funny ad for some junk food or a fastfood restaurant, they will want to choose the unhealthy foods over the healthy foods. In conclution, I think it really depends on the childs life style as to weather or not ads should be banned. If a child has a really good life with parents that explain things to them, it's ok. If a child has a life style that is not quite as good, and parents that don't talk to them about things as much, then certain ads should be banned.

SCORE 2 - LIMITED RESPONSE - 2c e

(SENTENCE-LEVEL SKILLS)

TestK_S3Q1

I do not agree with the U.S. government banning advertizing to children. I think that children are more influenced by their parents so if the parents tried to eduacate their children and exsplain how the advertizer try to persuade people and that thing arent always the same as they are on television I think that it would influence their children to make good choices about what the advertizment is telling them.

Researchers are alarmed that commercials persuade children to want the product so they nag their parents to buy it. I think that if an adult was not eduacated and saw a commercial that appealed to them and they didn't have money and they had people that took care of them and bought them the things they needed they would do the same. It's not just kids who are influence by advertizing I think that all age groups from children to elderly people without education on advertizing and the media could be persuaded by the media. The kids just need to be educated and it's the parents job to educate them about stuff like this, not the schools.

Snack food and candy are advertized showing fun things and famouse charactors like Spongebob Squarepants, like the cheetah for Cheetos, Tony the tiger the list could go on and I've never seen healthy food ads with mascots and fun ads. I think thats why maybe the snack food and candy ads are more apealing.

In europe a group made a survey for parents and 85% said that advertizing has little or no effect on their children. The other parent in the 25% should try to eduacate their children about advertizing and it persuasive messages, they could watch television with them and exsplain commercials to them like if a candy or snack commercial came on parents could exsplain where candy and snacks come in the world of food.

So I think that advertizement to children should not be banned but the parents should watch out for children and eduacate them on things the school dosn't since it's their job to raise their child and to them commen sence and how to deal with the world.

SCORE 1 - MINIMAL RESPONSE - 1 a b c d e

(SENTENCE-LEVEL SKILLS)

TestM_S4Q1

Well the read say help the pevent the soil and he what the kids in school show that we did good deed well the person who did it is right so we all do good deed well the person say we chose the class room that we what so the reads say about so the class Responsibility to help Academic. The person say what us to Apply the different class for the lessons each class that is good or bad Proposing th chose is right. For that activity with specific reasons?

Appendix C4b

Benchmark Responses Illustrating Document-Level Skills

SCORE 4 – EXCELLENT RESPONSE⁴⁶ (DOCUMENT-LEVEL SKILLS)

M S4Q 1

Dear Teacher.

The upcoming decison on which activity we should do for our service learning poiject is one that many are still considering. After considering them all, activity C should be our activity.

First off, activity C welcomes new students coming from other countries. It would be really hard to come from an other country and not know their language, or know how everything their works. With these guides, new kids will know will get to know everything better in a language that they understand. Kids who don't know english will get better grades because they will know how everything works and where classes are and everything. Kids would learn about other languages too.

Next, the students who already know english would learn more about other languages and culture and have a more open mind. Their would be less racism in our school, and kids would be more inviting and welcoming to different kids. There would also be more interest about learning different languages after learning alot about translators and what they do. Plus, all the kids may discover other interests about different cultures and explore that more broadening most of the many kids horizons.

Lastly, not only this year of eighth graders will learn from this. Many people think that this is bad that the eighth graders will only use them for a little amount of time, but all the following classes can reuse these guides and have all these benefits as well.

Overall, the introduction to our society for strangers is an important event, and changes what these people will see of our community. It is important to make them see it as a good place just like what it is, and best way for them to understand this is through the language that they understand. It will welcome the kids, open the minds of the kids from our community, and help generations of kids to come. Activity C should be our activity.

SCORE 3 – ADEQUATE RESPONSE

(DOCUMENT-LEVEL SKILLS)

M_S4Q1

Activity B should be chosen as tha service learning activity. It would benefiet the students that most out of all the activity because it has the best positive reasons and the negative reasons aren't really all that bad.

One of the reasons why it should be chosen is that it would help the community by picking up the liter by the creek. If the creek was clean there would be more people wanting to walk by it to take ajog or wllc there dogs. It would also just be a nice improvement to the landscape of the city. The student would benefiet from this by getting out of the classroom and getting some fresh air.

Another reason why this would be the best chioce is that the students could learn from it. They sould be able to study nature by taking water samples, observing plants, and writing up the results. Even though the students are learning about wheather and earth science they could always take a little sidestep to biology easily enough.

Activities A and C should not be chosen because there isn't enogh good benefeits for the students. Activity A woud be a bad chioce because not all the students know how to play a musical instrument and it would be hard to find time for them and the senior citizent to rehearse together. Activity C should not be chosen because by the time that the students fineish there guidbooks, most of the student would ahead be in high school. Also they couldn't translate all the guidbooks the the different languages they will need.

So in conclusion, looking at how task B benefiets the student academicly and helps there community it should be chosen over both Activities A and C which have far to many negative pionts to be a goos chioce for the students service learning project.

SCORE 2 – LIMITED RESPONSE – 2 b

(DOCUMENT-LEVEL SKILLS)

M-S4Q1

Adopt a creek. I would love to go to a local park and test the water, to see if it is okay for the animals to drink, swim, and kids to play in. Helping the local park is developing resonsibility. Helping to pick up trash, and testing the water to know the pH, and if it has germs, dead animals, or deadly water to not touch. We can also gain knowlege, beacsue if the creek has something that contains bad germs in it we can do something to help it. Gaining our Responsibility is a good thing because we can test the water, and then record it. Bring it back to the school researching how to keep it clean.

Whats bad for our community is saying there are already other people picking up litter in the park. Thats is really bad to say because anyone who is more willing to help, pick up the litter is helping. You can pick up litter even though people are picking it up too. Helping the community is good if you do it yourself too. It's not good for other people to pick up trash by themselves, helping is a good resonsibility, because its helping others to be clean and stay healthy.

 $SCORE\ 1 - MINIMAL - 1\ a\ b\ c$

(DOCUMENT-LEVEL SKILLS)

M_S4Q1

In this letter i thing the best thing to do for the community is to pick up trash to make the community and safe place and also clean. Nonething is better to do for your community beside picking up trash and not geting paid. We can also stop pulliton because it lead to global worming.

service learning TASK 2 – Evaluate a peer's essay

SCORE 4 – EXCELLENT RESPONSE (CRITICAL-THINKING SKILLS)

Part A – Evaluate the essay's strong points

Dear John,

Your argument meets the guidelines in several ways. I think you did a good job of stating your idea clearly. Each paragraph was clear as to what its topic was. You also did well supporting your arguments with reasons. After every statement, you had supporting reasons, which always makes persuasive writing better. For example, when you said "As citizens, we can help the environment, help our neighbors, or make the neighborhood a better place," you were supporting the previous statement: "The most important reason of all is to make a difference in our community." You kept your audience in mind because you spoke in a sophisticated way, but not so complex that you might have lost certain parents. You kept in mind that parents would be reading this by discussing how much their children could learn from the experience you are advocating for. You put your ideas in logical order by suggesting different ideas in each paragraph, not jumbling them all up together.

This also made the argument easy for readers to follow. You showed reasonable points in a good order. Overall, I think you did very well.

Part B – Evaluate the essay's weak points

Dear John,

Your argument is very well written, but there are still a few things you might want to take into account.

You might want to state each point a little bit more clearly. One way you could do this would be to state the main argument of each paragraph as the first sentence. In the last paragraph of your letter, you did not do this. Instead, you talked about the "Harvard Education Letter." I suggest you switch the topic sentence of your last paragraph with one of the other sentences in that same paragraph. You did a good job of supporting your arguments with reasons and facts, but a few more could not hurt. For example, in the third paragraph, why don't you give some statistic as to how community work helps middle schoolers. Otherwise, that whole paragraph is just based on your opinion, which does *not* make for a very good argument.

You should keep your readers in mind a little bit more by speaking more specifically to the principal. I feel like, through most of the letter, you were speaking to the parents, but do not forget your other audience.

One way you could keep your ideas in order would be to start off with a bunch of facts to grab people's *attention*. That way they will be more likely to listen to the opinion parts of your argument. You did not at all discuss arguments for the opposite point of view. Do not forget that some people might be skeptical. They might think that community service will not be educational for kids, so make sure to set them strait.

SCORE 4 – EXCELLENT RESPONSE – 2nd EXAMPLE (CRITICAL-THINKING SKILLS)

Part A – Evaluate the essay's strong points

Johns letter does state each point clearly. Johns first paragraph is all about how they could use what they have learned to help the community. He says they could use what they know about biology to tutor elementary school students, or use what they know about computers to teach adults about email. His second paragraph is about how important it is to do service learning because they will make a difference in the community. He gives examples of things other schools have done that really made a difference. In his third paragraph he explains that they will learn stuff that they just could not learn in school. Then he gives examples of what those things are. His third paragraph is a good way to close. He used a quote from a pretty respected person saying they thought it was a good idea to do service learning. Each paragraph was very clearly a different idea that the reader can easily pull out. John does a good job of stating all his points clearly.

Part B – Evaluate the essay's weak points

John does not support his information with facts. He does not show that it is proven that you will use things you learned in school on service learning. He does not show proof that what they do will make a difference in the community. He says what other schools did but he does not say if it was successful. He also does not show facts that they really will learn things they could not learn in a book or homework. Also his quote is just a quote it is not a fact So John did not do a good job of using facts to support his arguments.

SCORE 3 – ADEQUATE RESPONSE

(CRITICAL-THINKING SKILLS)

Part A – Evaluate the essay's strong points

John supports his argument with reasons or facts. He has a direct *quote from a* former astronaut and a senator supporting his idea. He also includes examples of other schools doing a service learning project successfully, and connected both of these facts to what he and his class wants to do to prove his argument.

Part B – Evaluate the essay's weak points

John needs to consider the arguments for the opposite point of view. He should include what the other side thinks and then tell why what they think is wrong. For example, maybe the opposing side says that it cost too much money to do these things. But can you put a price on learing, and you could even raise money for your school and the city *for* these projects. He needs to show why the opposing view is wrong.

SCORE 2 – LIMITED RESPONSE – 2 b

(CRITICAL-THINKING SKILLS)

Part A – Evaluate the essay's strong points

John has good support for his arguernents. He clearly gives the reasons why this project could be .usefull. John supports each of the reasons for his arguements with alot of detail. John has done well in supporting his arguement.

Part B – Evaluate the essay's weak points

John's arguments are not constructed for the more adult vocabulary. They are also written informally and persuasive writing is formal. John's arguments have a limited amount of more advanced vocabulary. He writes as though he where talking to his younger brother. Add more advanced wording, and make it a little more complex.

SCORE 1 – MINIMAL RESPONSE – 1 b/ 1 b c d

(CRITICAL-THINKING SKILLS)

Part A – Evaluate the essay's strong points

pick number 2 because He does makes a strong argument in favor of service learning. The kind of activity would I choose, and how would he convince your teacher to make the same choice.

Part B – Evaluate the essay's weak points

pick 5 because I don't think teachers would ask students parents for learing something.

service learning TASK 3 – Explain how an example meets guidelines

SCORE 4 – EXCELLENT RESPONSE (CRITICAL-THINKING SKILLS)

The example of a good service-learmong activity meets the guidelines of using their academic learning skills, gaining knowlege and skills, and devolping responsibility. It meets using their academic learning skills because the students have to use research skills and most middle schoolers know how to research documents. It meets gaining knoledge and skills because their learning about a particular or family, and their learning how to work with a camera and learning photography skills which most middle schoolers are not experts at. It meets devolping responsibility becasue they have to be careful with the cameras and cannont drop them or be rough with them. Those are the guidelines a example of a good sevicelearning activity meet or at least the ones I thought they meet.

SCORE 3 – ADEQUATE RESPONSE (CRITICAL-THINKING SKILLS)

The example of a good service-learning activeity meets many of the guidlines. First, it helps a need in the community because they say they are going to make paths so that people can come visit. They also want to make it so the cemetery is not so overgrown and they want to plant new plants. Secondly, they are gaining knowledge and skills by learning about the people and families who are buried there, it will help them with learning history. Thirdly, they develope resposibility because they are resposible for making the cemetery look nice and getting the records and histories of the people who are buried there for anybody who wants to know.

SCORE 2 - LIMITED REPONSE - 2 c

(CRITICAL-THINKING SKILLS)

The example of good service-learning activity meets using thier academic learning because they are still learning something and the project is not worthless. Another one is planning ahead. They are planning what they are going to do and how they are going to do the project. The last one is working with exsisting service groups. It says in the document that the source was provided by the Corporation for National and Community Service.

SCORE 1 - MINIMAL RESPONSE - 1 a c

(CRITICAL-THINKING SKILLS)

it has knowledge adn skills reasons why the should to the activity, it has facts and reasons they wanna do this kind of activity.

service learning TASK 4 – Argue for one activity and against two others

SCORE 4 – EXCELLENT RESPONSE

(CRITICAL-THINKING SKILLS)

Dear teacher,

In my opinion, activity A is the most reasonable and managable activity for the students to do. In performing songs with senior citizens, students would get to speak to the kinds of people they would not normally get to speak to. It would give them a new perspective on the world to hear senior citizens talk about what it was like when they were the students' age. Also, the students would get to learn new songs, and engage in the culture of music, which is a very good type of culture for students to be exposed to. Music creates emotions and feeds the spirit. It can teach empathy and it can change you all at once from being sad to happy of viseversa. Some of the senior citizens are highly skilled musicians, so the students would get to learn from them. Some of the students in the class do not sing or play an instrument, but that would be all right, because they would get to learn from the senior citizens. The senior citizens would also enjoy passing on their knowledge to young people. The issue of finding times when both groups can rehearse together is not an unbeatable one. Students could simply go up to the nursing home during school hours when the senior citizens are not busy.

I do not like the other two options as much. The main problem with activity B is that there is already a group picking up litter at the park. As long as the litter is being picked up, it doesn't matter whose doing it. The students might not find very much trash to pick up, and they might instead fool around. The other problem with option B is that the students aren't even learning about biology this year. So why would they do a project that is based on biology? People might think that students would enjoy taking water samples and writing up the results, but for one year, that was all my class did, and let me tell you...it was duller than watching paint dry. Students would get bored with this option and probably resort to playing in the park.

As for activity C, it is a very good idea, and I would support it, except that it would not be accessable to very much of the student population. Only three groups of foreign students would be able to benefit from this option, and I do not think this is fair to the other new students. They might feel like they are being snubbed, or that they are not welcome in their new school. All students should have an equal chance at success in school.

In conclusion, I think that activity A is the best activity for the students to do.

SCORE 3 – ADEQUATE RESPONSE

(CRITICAL-THINKING SKILLS)

The best activity in my point of view is activity B. Acitvity B is the best choice for a few reasons. This acitvity is the best choice, because keeping the environment clean is very important for the future, also they will learn a lot by dong something that is also helping out the community.

Activity B will be a better service-learning project, because there is a varity of things that the students can do, such as tkae water samples, pick up litter, observe animals and plants, and writing results. That means that if a student does not really enjoy watching the plants and animals there are other options they can choose from, overall it will make the majority of the studnets happy.

Activity A is not a great thing to do for a service-learning project, because a lot of students do not enjoy playing musical instruments, nevermind know how to, and most teenagers are to scared and shy to get up on the stage and sing. This project does not really help the students learn any vauleble information, the students do not need to know about history of a certain song.

Activty C is also a poor choice for a service-learning project, because it is not fair if only certain kids get to have the guide book, just because they speak that language. Also most kids will jsut skim it and throw it away, they will not apy to much attention to it. The guidebook is also not a important thing to have, mnay students hay moved to other states and not had a guidebook and they turned out fine, there usually will be one person who will help out that new student and make them feel welcome.

Overall activty B will be the best choice for a service-learning project.

SCORE 2 – LIMITED RESPONSE – 2 e

(CRITICAL-THINKING SKILLS)

Middle school students should learn new things and do service learning activity. The reason the class wants to do this activity is because the students gets to learn new things. That just reading books and doing homeworks couldn't teach us anything. We will learn how to built a boat and share with other schools and they might will learn something new from us too. The most important reason of all is to make a difference in our community as citizens. In other schools they learn new things every day and they do know things that we don't and thats why we should learn something new. In addition the students can built their own small baot and share facts with class. Then students will improve that they can learn things that they should learn. Not all students can built a nice perfect boat, but they can try their best to built it. Students needs lots and lots of skills to go to high school and college. This year students will have to learn new things, work and other suffs. They might need that skill for rest of their lives or probebly not at all. But they have to learn new educated things that will be important to their lives like Math, Science, and History. These things are important to their live some of them not all of them. Their are other people in another places or country that don't even know how to count so thats why we need these skills and learn new things.

SCORE 1 – MINIMAL RESPONSE – 1 a (CRITICAL-THINKING SKILLS)

The slow murmering of the creeck is peaceful and alot of people enjoy water noices. adopting the creek would be such an awsome idea because if we adopted a creek we would learn abbout many animals, earth, weather, and other things.

Animals are very important to us to eat or to be friend and if we adopt the creek then we can clean and filter the water so it would be more safe for animals to drink also if there is plastic the animals could be stengled to death witch is bad. Not good, Bad.

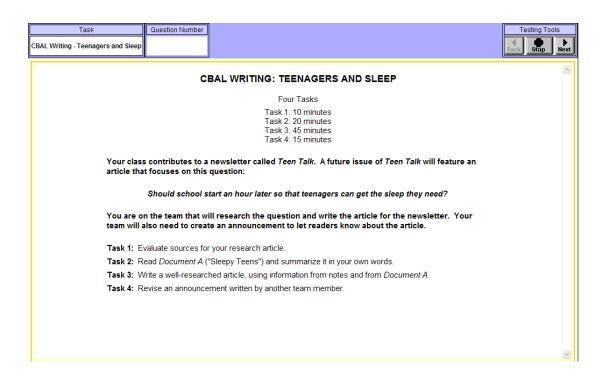
the earth is in the middle of globals wanning and each year about 72 gallons of carbon dioxide is dumpd in the air and thats only by one car.... if we cleaned up the creek we could be helping the air we breath and what everything breaths.

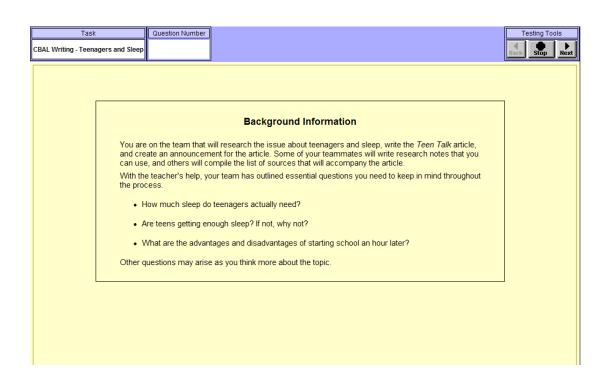
lastley, its very good to know about animals, earth, weather, and other things beacuse if we didnt know anything at all how would i be typing on this apple computer so if we cleaned up the creek and studied the surrouinding i have an idea the more and more people will becume more environmently friendly witch is good, not bad, good.

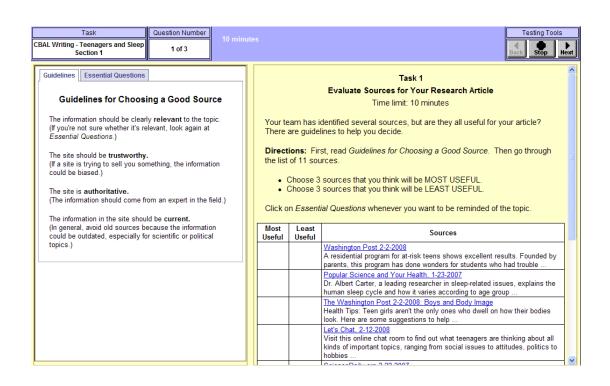
in conclutuion we sould adopt the creeck so w can learn about animals, eart, weather, and other thi that could affect the environment.

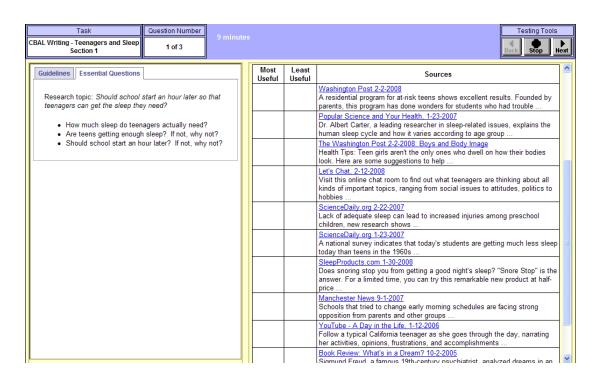
Appendix D1

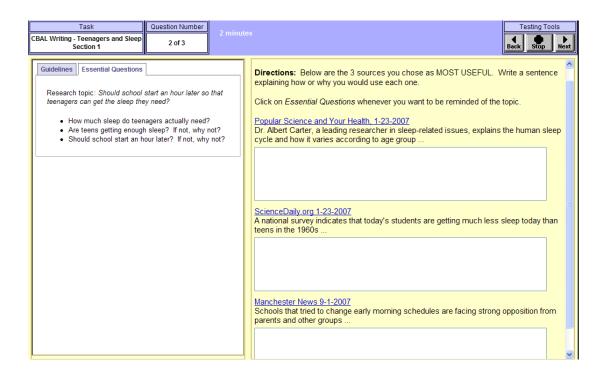
Screen Shots for Test Design #3, Expository Report (*Teenagers & Sleep* Administered Spring 2008; Scored Summer 2008)

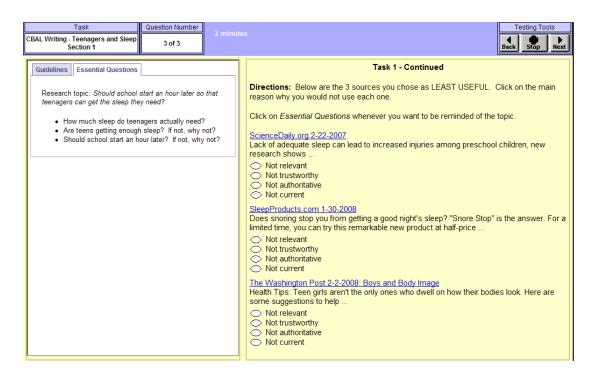


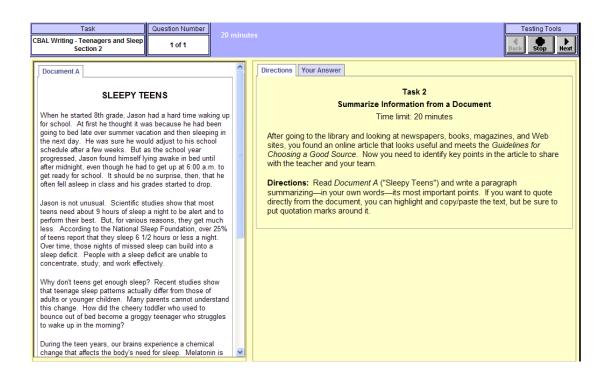


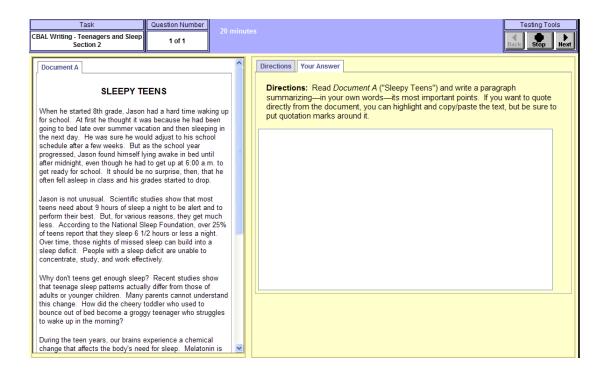


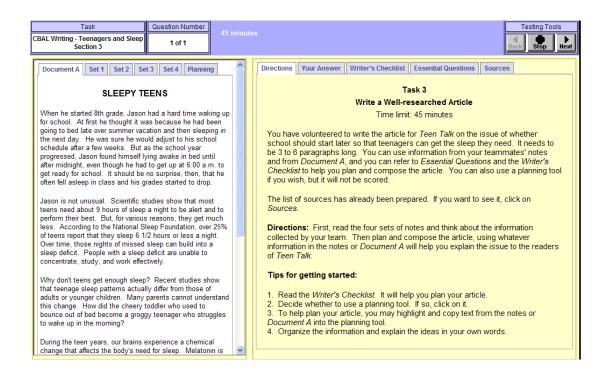


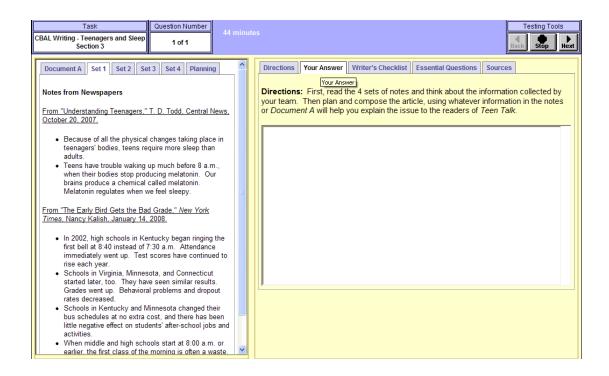


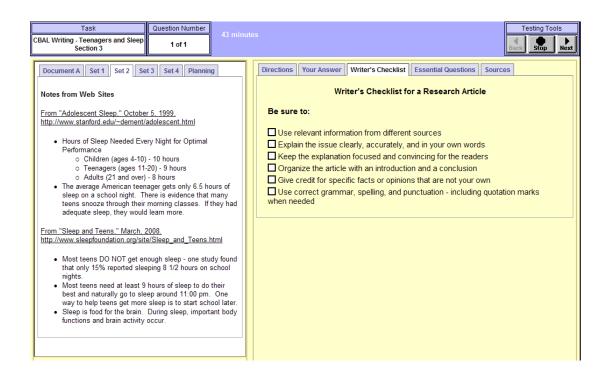


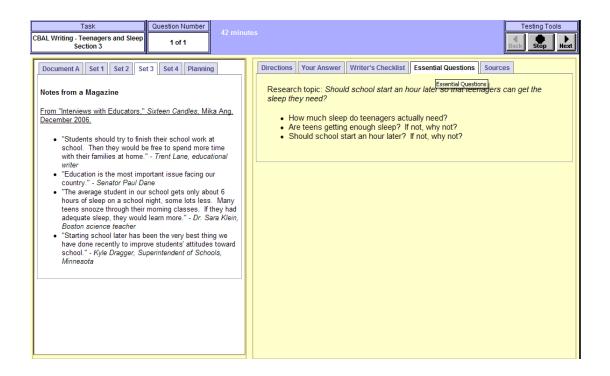


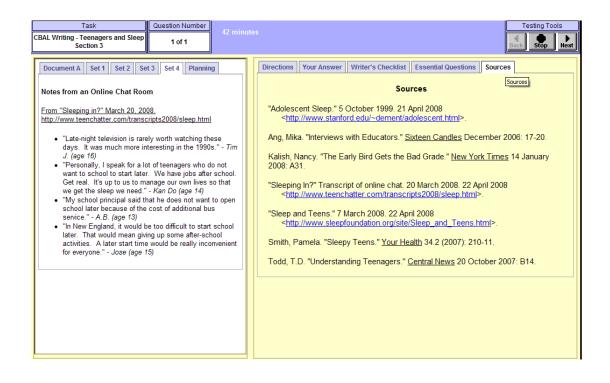


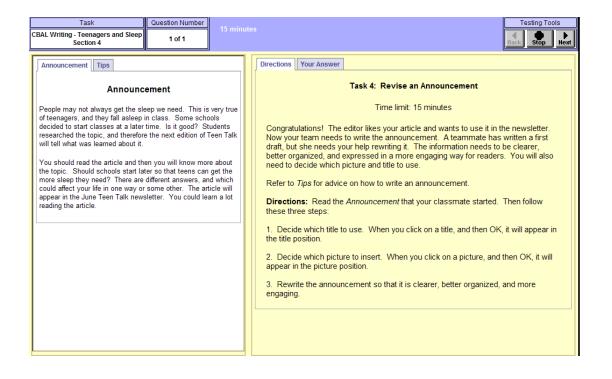


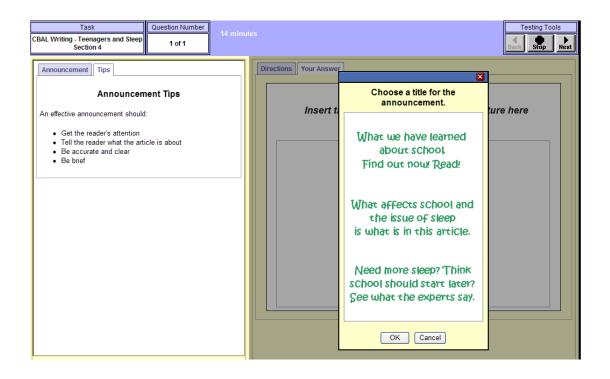


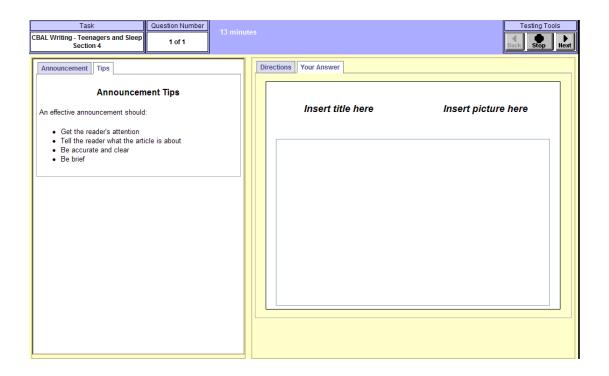












Appendix D2

Frequency Tables at the Strand Within Task Level for Test Design #3

Sleep_Task1a	Frequency	Percent		Cumulative Percent
0	8	6.02	8	6.02
1	43	32.33	51	38.35
2	68	51.13	119	89.47
3	14	10.53	133	100.00

Sleep_Task1c	Frequency	Percent	Cumulative Frequency	Cumulative Percent
0	31	23.31	31	23.31
1	36	27.07	67	50.38
2	39	29.32	106	79.70
3	27	20.30	133	100.00

Sleep_Task2_Strand1	Frequency	Percent	Cumulative Frequency	Cumulative Percent
0	7	5.26	7	5.26
0.5	4	3.01	11	8.27
1	15	11.28	26	19.55
1.5	13	9.77	39	29.32
2	23	17.29	62	46.62
2.5	19	14.29	81	60.90
3	30	22.56	111	83.46
3.5	20	15.04	131	98.50
4	2	1.50	133	100.00

Sleep_Task2_Strand2	Frequency	Percent	Cumulative Frequency	Cumulative Percent
0	7	5.26	7	5.26
0.5	9	6.77	16	12.03
1	13	9.77	29	21.80
1.5	20	15.04	49	36.84
2	28	21.05	77	57.89
2.5	16	12.03	93	69.92
3	23	17.29	116	87.22
3.5	15	11.28	131	98.50
4	2	1.50	133	100.00

Sleep_Task2_Strand3	Frequency	Percent	Cumulative Frequency	Cumulative Percent
0	21	15.79	21	15.79
0.5	1	0.75	22	16.54
1	28	21.05	50	37.59
1.5	11	8.27	61	45.86
2	27	20.30	88	66.17
2.5	16	12.03	104	78.20
3	16	12.03	120	90.23
3.5	10	7.52	130	97.74
4	3	2.26	133	100.00

Sleep_Task3_Strand1	Frequency	Percent	Cumulative Frequency	Cumulative Percent
0	25	18.80	25	18.80
0.5	3	2.26	28	21.05
1	5	3.76	33	24.81
1.5	9	6.77	42	31.58
2	20	15.04	62	46.62
2.5	16	12.03	78	58.65
3	20	15.04	98	73.68
3.5	22	16.54	120	90.23
4	13	9.77	133	100.00

Sleep_Task3_Strand2	Frequency	Percent	Cumulative Frequency	Cumulative Percent
0	26	19.55	26	19.55
0.5	3	2.26	29	21.80
1	18	13.53	47	35.34
1.5	15	11.28	62	46.62
2	10	7.52	72	54.14
2.5	20	15.04	92	69.17
3	19	14.29	111	83.46
3.5	12	9.02	123	92.48
4	10	7.52	133	100.00

Sleep_Task3_Strand3	Frequency	Percent	Cumulative Frequency	Cumulative Percent
0	30	22.56	30	22.56
0.5	7	5.26	37	27.82
1	27	20.30	64	48.12
1.5	21	15.79	85	63.91
2	14	10.53	99	74.44
2.5	17	12.78	116	87.22
3	11	8.27	127	95.49
3.5	4	3.01	131	98.50
4	2	1.50	133	100.00

Sleep_Task4_Strand1	Frequency	Percent	Cumulative Frequency	Cumulative Percent
0	21	15.79	21	15.79
0.5	3	2.26	24	18.05
1	4	3.01	28	21.05
1.5	12	9.02	40	30.08
2	11	8.27	51	38.35
2.5	19	14.29	70	52.63
3	30	22.56	100	75.19
3.5	25	18.80	125	93.98
4	8	6.02	133	100.00

Sleep_Task4_Strand2	Frequency	Percent	Cumulative Frequency	Cumulative Percent
0	19	14.29	19	14.29
0.5	5	3.76	24	18.05
1	11	8.27	35	26.32
1.5	13	9.77	48	36.09
2	13	9.77	61	45.86
2.5	25	18.80	86	64.66
3	31	23.31	117	87.97
3.5	13	9.77	130	97.74
4	3	2.26	133	100.00

Sleep_Task4_Strand3	Frequency	Percent	Cumulative Frequency	Cumulative Percent
0	32	24.06	32	24.06
0.5	2	1.50	34	25.56
1	20	15.04	54	40.60
1.5	8	6.02	62	46.62
2	8	6.02	70	52.63
2.5	27	20.30	97	72.93
3	24	18.05	121	90.98
3.5	10	7.52	131	98.50
4	2	1.50	133	100.00

Appendix D3

Keys and Rubrics for the Teenagers & Sleep Test

TASK 1

Question 1: Choose three sources that you think will be MOST USEFUL

B – Popular Science and Your Health

F—ScienceDaily.org

H – Manchester News

Question 2: Choose three sources that you think will be LEAST USEFUL

Any options other than B, F, H

Note: Question 3 responses were reviewed but not scored.

TEENAGERS & SLEEP - TASK 2

Write a summary of an article

Task Directions: Read the article "Sleepy Teens" and write a one-paragraph summary of its main points. (See *Topic Notes*.)

4 EXCELLENT -- An "Excellent" response:

- Succinctly summarizes the main points of the article and connects them clearly and accurately
- Includes no irrelevant, tangential, or trivial information

3 ADEQUATE -- An "Adequate" response:

- Summarizes all or most of the main points of the article and connects them in a generally clear and accurate way
- Includes little or no irrelevant, tangential, or trivial information

2 LIMITED -- A "Limited" response displays one or more of the following problems:

- a) Omits essential information from the article
- **b)** Misrepresents important points in the article
- c) Is at times unclear
- **d**) Combines summary with a significant amount of irrelevant or tangential information or minor details

1 MINIMAL -- A "Minimal" response is problematic in <u>one or more</u> of the following ways:

- a) Omits much of the essential information in the article
- **b)** Seriously misrepresents important points in the article
- c) Is generally unclear
- **d)** Fails to summarize the article and/or emphasizes irrelevant or tangential information or minor details

0 NO CREDIT -- A response receives "No Credit" for any one of the following reasons:

- f) Not enough of the student's own writing for critical-thinking skills to be judged
- g) Not written in English
- h) Off topic
- i) Blank
- j) Only random key strokes

TEENAGERS & SLEEP TOPIC NOTES -- TASK 2

Write a summary of an article

Sleepy Teens does not begin by stating a thesis statement or even a topic sentence; instead, it draws the reader into the topic by telling a little story about Jason struggling to wake up for school. The second paragraph transitions from Jason's experience to details about sleep research. The message is clear and the main points are easy enough to summarize, but they are not stated explicitly. Students have to tie them together and use their own words to express them.

Main points and major supporting details

According to Dr. Pamela Smith:

According to research studies, teens need about 9 hours of sleep per night, but over 25% of teens report that they get only $6\frac{1}{2}$ hours or less.

During the teenage years, many people do not get enough sleep because their bodies naturally produce melatonin on a later schedule from kids and adults.

Early school-starting times contribute to the problem.

Teens often struggle to be awake for their early classes.

"People with a sleep deficit are unable to concentrate, study, and work effectively."

<u>Information that should probably not be included</u>

- Details about Jason -- although some students might note that Dr. Smith began her article by giving an example of a teen who had trouble waking up for school
- Description of the "cheery toddler" bouncing out of bed

Paraphrasing/quoting Dr. Smith

Task directions tell students to use their own words and use quotation marks for direct quotations. Readers will need to be familiar with the text so that they can tell whether or not the student has used quotation marks correctly.

Author's rhetorical approach

Students may describe the way the author introduced and developed the information, but even a brilliant analysis of the rhetorical structure will not compensate for missing content or any misrepresentation of main points.

Personal information/opinions

Instead of summarizing the article, some students may merely give their own opinion about the issue. These responses receive a low score; they are off-task but not off-topic.

TEENAGERS & SLEEP - TASK 3

Write a well-researched article

Task Demands: Synthesizing information from multiple sources provided, write an article for a teen newsletter on the topic: Should school start later so that teenagers can get the sleep they need?

4 EXCELLENT – A typical "Excellent" response:

- Explains the issue clearly, completely, and accurately
- Insightfully synthesizes a range of important information and supporting details from multiple sources
- Communicates effectively for the audience

3 ADEQUATE – A typical "Adequate" response:

- Explains the issue in a generally clear, complete, and accurate way
- Synthesizes a range of important information and supporting details from multiple sources
- Communicates appropriately for the audience

2 LIMITED -- A "Limited" response displays one or more of the following problems:

- **a)** Includes important information from multiple sources but does not adequately synthesize it (lacks connections/explanations)
- **b**) Omits some important source information
- c) Misrepresents important points
- d) Includes a significant amount of unimportant information or minor details
- e) At times, communicates inappropriately for the audience

1 MINIMAL -- A "Minimal" response displays one or more of the following problems:

- a) Fails to synthesize the information
- **b)** Omits much of the important source information
- c) Seriously misrepresents important points
- d) Emphasizes unimportant information or minor details
- e) Communicates inappropriately for the audience

0 NO CREDIT -- A response receives "No Credit" for any one of the following reasons:

- a) Not enough of the student's own writing for critical-thinking skills to be judged
- **b)** Not written in English
- c) Off topic
- d) Blank
- e) Only random keystrokes

TEENAGERS & SLEEP – TASK 4

Rewrite an announcement

4 EXCELLENT -- An "Excellent" response:

- Engages the reader immediately and keeps the reader's attention
- Creates an announcement that is complete, accurate, and succinct
- Includes no irrelevant or inaccurate information

3 ADEQUATE -- An "Adequate" response:

- Addresses the audience adequately
- Creates an announcement that is generally complete and accurate
- Includes little, if any, irrelevant or inaccurate information

2 LIMITED -- A "Limited" response displays one or more of the following characteristics:

- a) Sometimes communicates inappropriately for the audience
- b) Omits some important information necessary to announce the article
- c) Includes some irrelevant or inaccurate information

1 MINIMAL -- A "Minimal" response displays <u>one or more</u> of the following characteristics:

- a) Does not communicate appropriately for the audience
- **b)** Does not announce the article
- c) Includes a substantial amount of irrelevant or inaccurate information

0 NO CREDIT -- A response receives "No Credit" for any one of the following reasons:

- a) Not enough of the student's own writing for critical-thinking skills to be judged
- **b)** Not written in English
- c) Off topic
- d) Blank
- e) Only random key strokes

TEENAGERS & SLEEP TOPIC NOTE

Rewrite an announcement

A "complete" announcement includes the following information:

• The next edition of *Teen Talk* will feature an article on the issue of whether schools should start later so that teenagers can get the sleep they need.

"Effective communication" for a teenage audience will:

- Have a lively tone
- Present information in a way that motivates students to want to read the article

Some students may make very few changes in the original text, but the bulleted characteristics of clarity and tone will still apply.

Appendix D4

Benchmark Responses for Teenagers & Sleep Test

TASK 4, Write a well-researched article47

SCORE 4 – EXCELLENT RESPONSE (SENTENCE-LEVEL SKILLS)

Teens need more sleep, it's a fact. Only a small percent actually get the 9 hours required for best performance. Later schools start times would give teens a better chance to learn. They miss most of the learning in their first period class because they can't stay awake. I'm sure all the teens out there reading this would agree. Studies in schools who have started opening later show that attendence is way up, grades are better, and many students are healthier.

SCORE 4 – EXCELLENT RESPONSE – 2nd example

(SENTENCE-LEVEL SKILLS)

Does an earlier start time for schools equal a negative impact on grade performance, as well as hurt test scores and attendance records? Research has shown that there is a legitimate link. Surveys, studies and records are showing that by pushing back the starting time for schools-only by an hourcan see an increase in student performance. Of those schools that have already set back the starting time, results are quite amazing. Research is also showing that during the teenage years, sleep cycles differ from those of adults. In all, it it seems that good things could come from pushing back school starting times just an hour.

Of everything that is being said over the topic, why school starting times should or shouldn't change, research is showing why it might be such a good idea after all to go ahead with the push. As children grow into teens, their bodies-and minds-change signifigantly. One of the largest changes deals with the production of a chemical called "Melatonin." This chemical is linked to sleep, as it's produced later in the night by teens, making it more difficult to fall asleep. By pushing back school starting times back after around 8:00am, when teen bodies stop the production of Melatonin, it can give ample time for teens to prepare to perform their best accademicly.

And those schools which have already pushed back starting times, namely those in the states of Kentucky, Virginia, Minnesota and Conneticuit, have seen wonderful results. High schools in Kentucky whom have pushed back start times-from 7:30 to 8:40-have seen a jump in attendence records. Test scores have also been on the rise. In the latter three of the mentioned states which followed Kentucky's example, there have been similar results, as well as increasing grade averages, and a dip in behavioral issues and dropouts. The results from these schools are model examples as to

why the starting times for schools should be set back.

Because of the sleep deficit, the first period If school is effectively wasted, as students find it a challenge to combat tiredness and stay awake. Teens would doze off and miss important sections of class discussions, ultimately hurting the teacher when the class is not knolagable of an important topic and when certain components must be re-taught.

In all, it seems to be quite bennificial to schools to bump back starting times just an hour. Research has shown that teens need that extra time to wake up and be more prepared for school, as well as preform well. Schools whom have already set back starting times are prime examples of what benefits can come of it. And also, when teens are more rested, they can focus more on their schoolwork and less on trying to stay awake.

SCORE 3 – ADEQUATE RESPONSE

(SENTENCE-LEVEL SKILLS)

I think that schools should start at the same time that they always have. If we change the time that school starts do you think that students will be able to get up? I think that if school starts later then they would want to stay out later and go to bed later than before. It really wouldn't help any thing. People from ages 11-20 years old need and average 9 hours of sleep to be functioning fully. Because of the physical changes that your bodies are going through we need more sleep than adults. One study found that only 15% of teens are getting 8 1/2 hours of sleep a night. The average teen gets only 6.5 hours of sleep on a school night. Many teens snooze through their morning classes. If they had adequate sleep, they would learn more. Also because most teens go to sleep around 11pm that gives them only 7 hours of sleep if they get up at 6 in the morning. "Personally, I speak for a lot of teenagers who do not want to school to start later. We have jobs after school. Get real. It's up to us to manage our own lives so that we get the sleep we need." says Kan Do age 14. Another reason i believe that school shouldn not start later is because of extra coricular activities, and many teenagers have jobs. Starting school later would restrict the free time they had even more. Students are already busy studying and just being active teens and spending time with friends. "In New England, it would be too difficult to start school later. That would mean giving up some after-school activities. A later start time would be really inconvenient for everyone." says Jose (age 15). These are the reasons that i believe school should not start later.

SCORE 3 – ADEQUATE RESPONSE – 2nd example

(SENTENCE-LEVEL SKILLS)

Should schools start earlier to increase teens metal health?

Waking up at 7:30 am isn't easy for us teens thats why 28% of middle schoolers and high schoolers ofter waste first class sleeping in, or falling asleep in class."According to the National Sleep Foundation, over 25% of teens report that they sleep 6 1/2 hours or less a night. Over time, those nights of missed sleep can build into a sleep deficit. People with a sleep deficit are unable to concentrate, study, and work effectively. "

Also, In 2002, high schools in Kentucky opened up at 8:40 instead of 7:30 a.m. Attendance immediately went up. And thats because more children and teens got to sleep in, which increased the attendance. "Schools in Virginia, Minnesota, and Connecticut started later, too. They have seen similar results. Grades went up. Behavioral problems and dropout rates decreased."

What teens are facing today is puberty. The big P word changes everything and messes most kids up, because of all those physical and mental changes taking place, teenagers' bodies require more sleep than adults.

Sleep is like food for the brain. When teens sleep, important body functions and brain activity occur. Most teens DO NOT get enough sleep. Only 15% of students sleep 8 1/2 hours on school nights. And "Most teens need at least 9 hours of sleep to do their best and naturally go to sleep around 11:00 pm. One way to help teens get more sleep is to start school later."

SCORE 2 – LIMITED RESPONSE

(SENTENCE-LEVEL SKILLS)

According to the National Sleep Foundation, over 25% of teens report that they sleep 6 1/2 hours or less a night.

On a recent study found, only 15% of teens get 8 1/2 hours of sleep on school nights. If teens got adequate sleep each night, they would be well rested and be more active in school. Most teens need about 9 hours of sleep to be ready and alert. If the brain doesnt get sleep, your body wont function the right way and youll do a lot of things differently.

There are a lot of physical changes and one teenagers body. This causes them to have a hard time falling asleep at night and waking up in the morning hours. Many teens stop producing melatonin

which they need to feel sleepy. Teens need more sleep than adults because they are'nt on a regular sleeping schedule yet like most adults.

Parents say that students should try and finish their work at school. They say this so the students can spend more time with their families. It would be a lot better for the students too so they can go to bed earlier. Many teens fall asleep through the first morning class so whats the point of having it?

In 2002 high schools in Kentucky began to start school later. Automatically there was'nt as many dropouts and there were more students showing up. Each students test scores went up by a lot sense they started to get the right amount of sleep they needed. If this happened in a lot of states around the U.S., more schools would be getting these same results.

A lot of people think that their kids dont get enough sleep from the summer sleep schedule. Going to bed late and waking up late is what pretty much all teens do. They think they cant get to sleep during the school year because of how they spent their summer sleeping schedule. But in my opinion, that isn't true. I think most kids are concerned about how their going to get a project done or they just have a lot of stuff on their mind. I know i've had experiences like that before. Therefor i think schools should make the start time later so students can be more awake and more alert.

SCORE 2 – LIMITED RESPONSE – 2nd example

(SENTENCE-LEVEL SKILLS)

Teens nead about nine in a half hours. This is neaded so that the teens can payatenction in school and not fall asleap in class.

Teens are not getting enuf sleep. Teens and going to bed later because of what there body tells them to so that is around eleven oclock at night. So are teens getting enuf sleep? No they are not becuse of that reson so that makes less awer of there actoins.

Should schools start later? No they shouldn't because that would make kids be less active in the morning and make the amarican population more fater then we all ready are.

SCORE 1 – MINIMAL RESPONSE

(SENTENCE-LEVEL SKILLS)

I agree that school should start one hour later because it will give us all time to wake up and start our day because I am one of those kids who sleep in school and it sucks so yes i agree.

SCORE 1 – MINIMAL RESPONSE – 2nd example

(SENTENCE-LEVEL SKILLS)

I think that teens need more sleep so school school start later if we get more sleep we will do beter work and pay more atesion in class.

"Teenagers & Sleep" TASK 4, Write a well-researched article

SCORE 4 – EXCELLENT RESPONSE (DOCUMENT-LEVEL SKILLS)

I believe that starting school later is a good thing. It will increase teens grades, improve their work ethic, and teens would be getting the sleep they need.

How much sleep do you get at night?

Most teens say they get only 6 to 6 1/2 hours of sleep, when they are supposed to be getting 9 hours of sleep. "According to the National Sleep Foundation, over 25% of teens report that they sleep 6 1/2 hours or less a night." "Most teens DO NOT get enough sleep - one study found that only 15% reported sleeping 8 1/2 hours on school nights."

How much sleep should you get?

If you are between the ages of 11 and 20, then in order to preform your best, you should be getting at least 9 hours of sleep at night. Most teens don't get nearly enough sleep therefore they do not do so well in school. They don't have the energy for school, because they go to sleep later and then have to wake up early for school. You should be getting 9 hours of sleep but that can be hard because at night "our brains produce a chemical called melatonin." This regulates when you fall asleep and wake up. For teens it doesn't start producing until later at night, and doesn't stop producing until later in the morning.

How can we get the sleep we need without falling asleep in class?

Some people have suggested starting school later in the morning. Then we could wake up at a time where teens can get a full 9 hours of sleep. Because of the personal changes teens are going through at this time in their life, they naturally need more sleep, because they don't fall asleep until later at night. Then again people think it is our responsibility to get the sleep we need, but also we can not control the changes going on. We can try to force our selves to sleep but then we won't be getting a good nights sleep, and that is important too.

Should we start school later?

"In 2002, high schools in Kentucky began ringing the first bell at 8:40 instead of 7:30 a.m. Attendance immediately went up. Test scores have continued to rise each year." Is this the answer to our problems?

Many people agree that starting school later would allow us to get the sleep we need without missing school. We would be more attentive to the lesson, and so we would learn more. Teens are waking up as early as 6 in the morning, obviously not getting 9 hours of sleep. If we started school later I think many, many teens would be getting the sleep they need in order to their best. I think it

would bring tests scores and our grades up a lot, because we would be listening and paying attention. School would be worth it.

I think that by reading this paper, you should be convinced that starting school later would definalty resolve the problem of the loss of sleep for teens. It would increase the attendance and scores of the students, teens would be getting 9 hours of sleep therefore performing their best. I think you, and many more teens would agree.

SCORE 4 – EXCELLENT RESPONSE – 2nd example (DOCUMENT-LEVEL SKILLS)

Does an earlier start time for schools equal a negative impact on grade performance, as well as hurt test scores and attendance records? Research has shown that there is a legitimate link. Surveys, studies and records are showing that by pushing back the starting time for schools-only by an hour-can see an increase in student performance. Of those schools that have already set back the starting time, results are quite amazing. Research is also showing that during the teenage years, sleep cycles differ from those of adults. In all, it it seems that good things could come from pushing back school starting times just an hour.

Of everything that is being said over the topic, why school starting times should or shouldn't change, research is showing why it might be such a good idea after all to go ahead with the push. As children grow into teens, their bodies-and minds-change signifigantly. One of the largest changes deals with the production of a chemical called "Melatonin." This chemical is linked to sleep, as it's produced later in the night by teens, making it more difficult to fall asleep. By pushing back school starting times back after around 8:00am, when teen bodies stop the production of Melatonin, it can give ample time for teens to prepare to perform their best accademicly.

And those schools which have already pushed back starting times, namely those in the states of Kentucky, Virginia, Minnesota and Conneticuit, have seen wonderful results. High schools in Kentucky whom have pushed back start times-from 7:30 to 8:40-have seen a jump in attendence records. Test scores have also been on the rise. In the latter three of the mentioned states which followed Kentucky's example, there have been

similar results, as well as increasing grade averages, and a dip in behavioral issues and dropouts. The results from these schools are model examples as to why the starting times for schools should be set back.

Because of the sleep deficit, the first period lf school is effectively wasted, as students find it a challenge to combat tiredness and stay awake. Teens would doze off and miss important sections of class discussions, ultimately hurting the teacher when the class is not knolagable of an important topic and when certain components must be re-taught.

In all, it seems to be quite bennificial to schools to bump back starting times just an hour. Research has shown that teens need that extra time to wake up and be more prepared for school, as well as preform well. Schools whom have already set back starting times are prime examples of what benifits can come of it. And also, when teens are more rested, they can focus more on thier schoolwork and less on trying to stay awake.

SCORE 3 – ADEQUATE RESPONSE

(DOCUMENT-LEVEL SKILLS)

There is a crisis of epidemic scale that involes our teens and sleep. This crisis is our teens are not getting nearly enough sleep as they should get. This is extremely damaging to health, education, and daily life. The right fix is to start shoool an hour later then normal.

Starting school an hour later then normal will instanly fix the education problem in many ways. The first being attendance will rise sharply. This will help everyone get a good education and also benifit the school for spearheading a national issue. The seconed wayh it will help is most states will have state wide testing in the morning. With teens only getting 6.5 out of 9 hours they need these can be extremely damaging to test scores. With bad test scores schools will not be able to get the funding they desvere due to No child left behind. The last way it will benifit the education of students is that 28% of students who do show up fall alsepp in their first class.

Health is a major factor with everyone. With the rising obestity rate in America the extra hour in the morning is vital to the health of teens. More and more teens everyday are skipping breakfast to get to school in time. Breakfast is important because it gives you the fuel you need to get to lunch and it starts the metobolisism. With more kids eating healthy breakfasts obesity rate would steadily go down. Schools would also not have to buy as much food for lunch because the students will have already eaten during hte morning; lowering food costs and allowing more budget for other important areas of concern.

Dialy life is as important teens as their own life. Weaither its tecting or after school sports teens need the afternoon for things.

SCORE 3 – ADEQUATE RESPONSE – 2nd example (DOCUMENT-LEVEL SKILLS)

Many people were talking about starting school an hour later for teens so that they can get more sleep during the night.

I think that we should keep the starting time for school for a few reasons. For the teens starting school early in the day is a good thing because they will get use to waking up early. If they get use to waking up early everyday that will be good for them later in life. If we moved the starting time of school most teens will just go to bed an hour later than they do now. If we moved the starting time back it wouldn't help the average teen's sleep habats.

Teens need a good social life, whether its hanging out with friends, after school activities or playing sports. If we were to keep them in school for another hour they wouldn't have enough time for these things.

If we started school later in the day it would not be a good thing for the parents. First, they would have to bring their kids to school an extra hour later which might cause them to be late for work. It also would effect how they sleep, seeing as they need 8 hours of sleep, they would need to go to bed while their kids were still awake because the teen will just stay up later.

I think that their are too many changes that would need to be made just so a kid could get an extra hour of sleep. It would be so much easier for the kid to go to bed a little earlier.

SCORE 2 – LIMITED RESPONSE (DOCUMENT-LEVEL SKILLS)

Teens should get nine hours of sleep because if they don't they won't do well in school, sports...ext. Their grades may start to drop as well as them falling asleep in class or being late to school because of sleeping in to late so they end up on missing one or two of their classes. Studies have shown that teens that sleep for nine hours or more do better in school and their grades are going up instead of dropping, their sports, if they are apart of a sport team, they have a better chance of doing great instead of doing terrible.

SCORE 2 – LIMITED RESPONSE – 2nd example (DOCUMENT-LEVEL SKILLS)

If school started later, than teens would get more sleep, but if they have many things going on later, such as homework, sports, a job, or parties, than what time would all of these thigns get done? "Personally, I speak for a lot of teenagers who do not want to school to start later. We have jobs after school. Get real. It's up to us to manage our own lives so that we get the sleep we need." - Kan Do (age 14).

If kids really wanted sleep, than they would have to manage it on their own time, like Kan Do says.

An example of this is if a student gets out of school and 3:30 instead of 2:00 p.m or even 2:30 p.m., than if they play sports, they have practice or games, so that could push it to 4:30 or even 5 p.m., than they come home, go to work (If they have a job) and get back at 8:00 p.m., than eat dinner, and do homework. By the time that that is all done, than its close to 10. That means its only giving them 8 hours of sleep, which, occurring to the scientists that say they need 9 hours, is not enough.

"Because of all the physical changes taking place in teenagers' bodies, teens require more sleep than adults." T. D. Todd, Central News. This exact fact proves my point that kids need more sleep than other people, but why start school later? It would just push things back more and more.

SCORE 1 – MINIMAL RESPONSE

(DOCUMENT-LEVEL SKILLS)

most teens go to bed late and have to wake up early . going to bed can be very easy and can be very difficult . kids also like to stay up late because their never tired but when they say its time to go to bed t can be a very long time befor acauly falling asleep . same as geting up . its easy and feels beter most of the time to get up late on weekends (sleeping in) untill 9:00 or 10:00 but durring school days its hard to get up .

SCORE 1 – MINIMAL RESPONSE – 2nd example (DOCUMENT-LEVEL SKILLS)

set 1 because of physical change during teenage years. There bodies go's through changes like puberty they would need more sleep because they are producing Melatonin it makes them go to bed later at night and that would make them wake up later in the morning.

set 2 most teenagers have less sleep because of school work when they come home they usually don't do there homework right away they want to go out side and play with friends. some ways to make kids learn is to start school at a later time that means they would sleep a little longer and wouldn't sleep during classes.

set 3 student should try to finish school work during school but students don't have that much in class we only have 45 minutes in a class before it ends and during that 45 minutes teachers are talking through half of are time and by the time the finish we only have like 15 minutes to do work and when we have time we don't do work because now your at home and can do what ever you like and not tell your parents that you have homework and if they did it is usually noisy at homes.

set 4 if we started school at a later time that would be a problem because we have 6 hours of school for example like at 1 we would get out at 7 that is no free time to play out side and it would be dark out and you cant do you home work because your so tired.

set 2 most teenagers have less sleep because of school work when they come home they usually don't do there homework right away they want to go out side and play with friends. some ways to make kids learn is to start school at a later time that means they would sleep a little longer and wouldn't sleep during classes. set 3 student should try to finish school work during school but students don't have that much in class we only have 45 minutes in a class before it ends and during that 45 minutes teachers are talking through half of are time and by the time the finish we only have like 15 minutes to do work and when we have time we don't do work because now your at home and can do what ever you like and not tell your parents that you have homework and if they did it is usually noisy at homes. set 4 if we started school at a later time that would be a problem because we have 6 hours of school for example like at 1 we would get out at 7 that is no free time to play out side and it would be dark out and you cant do you home work because your so tired.

"Teenagers & Sleep"TASK 2 – Summarize an article

SCORE 4 – EXCELLENT RESPONSE (CRITICAL-THINKING SKILLS)

Student's Response:

This paragraph had many facts about sleep and related issues. In these, "Scientific studies show that most teens need about 9 hours of sleep a night." This would mean that kids would have to go to bed a 9:00 to get 9 hours and get up at 6:00. 25% of kids only get 6 1/2 hours a night. "During the teen years, our brains experience a chemical change that affects the body's need for sleep. Melatonin is produced later at night in teens than it is for kids and adults, making it harder for teens to fall asleep on the same schedule." This show how big of a problem it is, "A couple hours of missed sleep a night may not seem like a big deal, but over time, this can become a serious problem."

SCORE 4 – EXCELLENT RESPONSE – 2nd example (CRITICAL-THINKING SKILLS)

Student's Response:

Over 25% of teens sleep less than 6 1/2 hours a night, they need 9 hours to be alert, if this becomes a recurring problem it causes a sleep deficit which can have serious consequences. During teenage years Melatonin is produced later at night than any other stage in life causing an inability to go to sleep early. When this is occurring teenagers are already more stressed and overwhelmed by the world, thanks in some ways to early starts at schools.

SCORE 3 – ADEQUATE RESPONSE (CRITICAL-THINKING SKILLS)

Student's Response:

The avereage teenager only gets around 6 1/2 hour or less of sleep and is too little sleep for waking up at school at 6:00 or 6:30 in the morning. That can cause to be sleepy or even fall asleep in class, which can make it difficult for teenagers to concentrate in class or study or even falling asleep in class. A chemical change is produced in the later years of teenagers that affects teens bodys need to sleep. "These changes come at a time when kids have more homework, more sports and other extracurricular activities, a busier social life and sometimes even a part-time job." All those things that teens do cause them to stay up later and not get enough sleep from waking up so early from school. School starting too early is what I think is causing teens to sleep in class.

SCORE 3 – ADEQUATE RESPONSE – 2md example (CRITICAL-THINKING SKILLS)

Student's Response:

Sleepy teens is about sleep deprived students or teens who can't keep up with their life. Studies show that an average teen needs 9 hours of sleep every night. Thats from 9 pm to 6 am. This is a growing concern, studies show that 25% of teens get less than 6 1/2 hours a night. Losing that much sleep can cause sleep deficit, sleep deficit is when you are unable to concentrate, study and work effectively.

SCORE 2 – LIMITED RESPONSE

(CRITICAL-THINKING SKILLS)

Student's Response:

Being a teenager in full throttle, I understand the difficulties with going to bed early and waking up early. As Dr. Pamela Smith has given, teenagers have Melatonin that is produced later in the night than kids and adults, which makes it harder on us to fall asleep at a regular time. Melatonin is a chemical in your brain that helps you to sleep. I have witnessed this myself, since I am a teen, it is harder for me to go to bed. Around 930-10:00, I still feel very energized and could spend the night awake. Unlike my Mother and Sister, who are older than me, they go to bed around that time. "Scientific studies show that most teens need about 9 hours of sleep a night to be alert and to perform their best", says the article. Well, if we need 9 hours of sleep then why don't we just go to bed, you ask. The lack of Melatonin at a regular hour is partially to blame, but it is also the stress that comes along with getting up at such an early hour for us to get ready for school.

SCORE 2 – LIMITED RESPONSE – 2nd example

(CRITICAL-THINKING SKILLS)

Student's Response:

This article tells you a lot of important point. It expresses very strongly the need of sleep in young adults and teenagers. It show many different components of why we aren't getting the sleep we need as well. For example it shows real life examples of why we don't get to sleep on time, including, jobs, social life, homework, sports, and all sorts of extra-curriculars. It gives a few explainations for it too. These explainations include schools starting too early. If you have to be up to get ready by six it says that you would have to go to bed at nine o'clock to get the recommended 9 hours, and that is unrealistic. It shows important points including why it happens, what the solution is, and how to fix it.

SCORE 1 – MINIMAL RESPONSE

(CRITICAL-THINKING SKILLS)

Student's Response:

To start off, the article talks about a teenager named jason who has a hard time getting up in the morning. He stays up late and wakes up early. In the begining of the school year, it was pretty easy for him. As the school year went on, the number of missed hours of sleep stacked up. This is very common of teens around his age. Because teenagers are falling asleep in class, their grades has dropped and they show lack of focause.

SCORE 1 – MINIMAL RESPONSE – 2nd example

(CRITICAL-THINKING SKILLS)

Student's Response:

It tells you about why students are so tierd in the morning. You need at least 9 hours of sleep a night. Parents do not get "How did the cheery toddler who used to bounce out of bed become a groggy teenager who struggles to wake up in the morning?"

"Teenagers & Sleep" TASK 3 – Write a well-researched article

SCORE 4 – EXCELLENT RESPONSE

(CRITICAL-THINKING SKILLS)

Does an earlier start time for schools equal a negative impact on grade performance, as well as hurt test scores and attendance records? Research has shown that there is a legitimate link. Surveys, studies and records are showing that by pushing back the starting time for schools-only by an hour-can see an increase in student performance. Of those schools that have already set back the starting time, results are quite amazing. Research is also showing that during the teenage years, sleep cycles differ from those of adults. In all, it is seems that good things could come from pushing back school starting times just an hour.

Of everything that is being said over the topic, why school starting times should or shouldn't change, research is showing why it might be such a good idea after all to go ahead with the push. As children grow into teens, their bodies-and minds-change signifigantly. One of the largest changes deals with the production of a chemical called "Melatonin." This chemical is linked to sleep, as it's produced later in the night by teens, making it more difficult to fall asleep. By pushing back school starting times back after around 8:00am, when teen bodies stop the production of Melatonin, it can give ample time for teens to prepare to perform their best accademicly.

And those schools which have already pushed back starting times, namely those in the states of Kentucky, Virginia, Minnesota and Conneticuit, have seen wonderful results. High schools in Kentucky whom have pushed back start times-from 7:30 to 8:40-have seen a jump in attendence records. Test scores have also been on the rise. In the latter three of the mentioned states which followed Kentucky's example, there have been similar results, as well as increasing grade averages, and a dip in behavioral issues and dropouts. The results from these schools are model examples as to why the starting times for schools should be set back.

Because of the sleep deficit, the first period lf school is effectively wasted, as students find it a challenge to combat tiredness and stay awake. Teens would doze off and miss important sections of class discussions, ultimately hurting the teacher when the class is not knolagable of an important topic and when certain components must be re-taught.

In all, it seems to be quite bennificial to schools to bump back starting times just an hour. Research has shown that teens need that extra time to wake up and be more prepared for school, as well as perform well. Schools whom have already set back starting times are prime examples of what benifits can come of it. And also, when teens are more rested, they can focus more on thier schoolwork and less on trying to stay awake.

SCORE 4 – EXCELLENT RESPONSE – 2nd example

(CRITICAL-THINKING SKILLS)

Teenager's across the nation would all benefit greatly if the start time of their schools were pushed later into the day, there for providing them with more sleep. Teenagers are rarely getting the amount of sleep that is healthy for their body and their attitude, and if we were to allow them to get extra sleep, improvements in academic performances would be seen. School that have tried this technique have been having very successful results.

Lack of sleep causes many problems to occur at school. Students who have trouble waking up in the morning usually end of being late to school. Also, if students are tired and dozing off during the school day, they will not be focussed in class at all. So, if students were given time for more sleep, they would be awake in and more alert in class, and would then be able to learn more during the school day. Not only is lack of sleep causing school issues, but the huge amount of sleep teenagers are missing is unhealthy.

Teenagers do not get the healthy amount of sleep they sohuld be getting at their age. The healthy amount of sleep teenagers should get is 9 hours each night. The average teenager only gets 6 hours or 6.5 hours of sleep on school nights, and there are many teenagers getting much less then that. Studies have shown that only 15% of teenagers are getting 8.5 hours of sleep on school nights. This is a major issue, and schools are acting on it and have seen positive results. Schools who have tried pushing the start time of the school day later have seen very encouraging results. *Kyle Dragger, Superintendent of Schools, Minnesota said* "Starting school later has been the very best thing we have done recently to improve students' attitudes toward school." In 2002, various high schools in Kentucky began starting school at 8:40 rather then 7:30. They were then rewarded with higher attendance records and higher test scores. Other schools in Virginia, Minnesota and Connecticut have seen similar results. They have also seen better student behavior, lower drop-out rates, and students' grade going up. These positive results are definately encouragin and should convince other schools to do the same.

The different functions of the teenage brain gives them odd sleeping patterns, and in the circumstances they are forced to live by with school and other activities, it is difficult for them to have enough sleep. Many teenagers are getting up 3 hours less sleep then they should be, and this causes many problems

in the school. Schools have been pushing the start time of the school back and have been seeing very positive results. Teenager's across the nation would all benefit greatly if the start time of their schools were pushed later into the day.

SCORE 3 – ADEQUATE RESPONSE

(CRITICAL-THINKING SKILLS)

The majority of teenagers feel tired and asleep during the morning, some of them blame of it on school starting at such an early time. Is this true? In many ways it is. Teens require at least 9 hours of sleep to be rested in the morning. According to a survey by the National Sleep Foundation 25% of teens don't even get 6 1/2 hours of sleep. Another factor is when melatonin is produced in a teenagers body. Melatonin is a chemical that helps our body decide when we are sleepy. During adolescence it is produced between 11:00 pm and 8:00 am. Since most schools start their days around 7:30 am or earlier this can cause problems with how a person thinks or learns. It has been proven by many medical institutions that that not getting enough sleep can severely negatively impact a persons cognitive abilities. Students sometimes try to catch up on their sleep during the first class of school, as much as 28%. If they even come at all. If they don't get enough sleep consistently for months they may become sleep deprived and have trouble doing some of the simplest of things. High schools in Kentucky tried a simple solution to this problem. In 2002 they started school at 8:40 am 1 hour 10 minutes later than the usual 7:30. Attendance and test scores immediately went up and have continued to rise. Behavioral problems have gone down significantly. Schools in Virginia, Minnesota, and Connecticut tried this approach after the success in Kentucky and have seen similar results. To conclude you can't blame students sleepiness on them staying up late. It's not their fault and schools could do so much more to fix it.

SCORE 3 - ADEQUATE RESPONSE - 2nd example

(CRITICAL-THINKING SKILLS)

Teens do not get enough sleep. The average hours of sleep teens get is 6.5, the required amount of hours is 9. In order to get 9 hours of sleep and wake up in time for school, all teens would need to go to bed at 9 pm. This simply does not work with many teens. Teens have sports, activities, social lives, and some have jobs. Most teens stay up to 10 or 11 pm and have trouble waking up in time for school. An hour extra of sleep would make a world of difference for teens' performance in school. At the teenager stage, a chemical change in the brain causes teens to stay up later and get up later. This chemical change is the producing of melatonin. Melatonin controls when we feel sleepy. Melatonin doesn't stop producing until 8 am, already an hour into school for some teens. Schools that start at 8 am or earlier are often a waste of time. Teens end up not paying attention or falling asleep in class. Some are so tired they don't show up. For the rest of the classes they are tired too. Starting school an hour later would mean that students would be alert and paying attention in all classes rather than sleepy and un-alert in all of them.

Some schools have already implemented this change. High schools in Kentucky start at 8:40 am. Attendance and test scores are much higher than other schools. Schools in other states such as Virginia, Minnesota, and Connecticut have also changed hours. They have less behavioral problems as well.

All schools should just give teens an extra hour of sleep. Their performance in school will improve drastically, while maintaining a busy social life and after school activity.

SCORE 2 – LIMITED RESPONSE

(CRITICAL-THINKING SKILLS)

I think that schools should start at the same time that they always have. If we change the time that school starts do you think that students will be able to get up? I think that if school starts later then they would want to stay out later and go to bed later than before. It really wouldn't help any thing. People from ages 11-20 years old need and average 9 hours of

It really wouldn't help any thing. People from ages 11-20 years old need and average 9 hours of sleep to be functioning fully. Because of the physical changes that your bodies are going through we need more sleep than adults. One study found that only 15% of teens are getting 8 1/2 hours of sleep a night. The average teen gets only 6.5 hours of sleep on a school night. Many teens snooze through their morning classes. If they had adequate sleep, they would learn more. Also because most teens go to sleep around 11pm that gives them only 7 hours of sleep if they get up at 6 in the morning. "Personally,

I speak for a lot of teenagers who do not want to school to start later. We have jobs after school. Get real. It's up to us to manage our own lives so that we get the sleep we need." says Kan Do age 14. Another reason i believe that school shouldn not start later is because of extra coricular activities, and many teenagers have jobs. Starting school later would restrict the free time they had even more. Students are already busy studying and just being active teens and spending time with friends. "In New England, it would be too difficult to start school later. That would mean giving up some after-school activities. A later start time would be really inconvenient for everyone." says Jose (age 15). These are the reasons that i believe school should not start later.

SCORE 2 – LIMITED RESPONSE – 2nd example

(CRITICAL-THINKING SKILLS)

Teenagers are always falling asleep in class or coming in tardy because they slept in. Many teens including myself do think that schools should start a little bit later than they do now many, many kids cant wake up at 6:00 a.m to get all ready and leave for school because we are not refreshed enough to do activities without being tired.

In middle school and in high school kids are put up to the task to do a lot of homework and studying but the main thing is waking up early to go to school. Teens are required to sleep 9 hours to be ready to go to school and learn. Most kids now don't really go to sleep at the time they need to, most kids have after school activities like sports and clubs and maybe studying for a big test also kids like to socialize with friends and family so many kids go to sleep late. Sleeping in late isn't kids faults, teen around the age of 11- 20 need more sleep then adults and younger kids. Kids need to at least have an another hour to get the sleep they need, get dressed and eat breakfast so then they can go to school well prepared and ready to learn what they need to know to be successful in life. Kids also have worries about school starting later because school would finish an hour later and kids have jobs and commitments they do after school. But as kids see they would still manage to do everything they do after school when they see how better it is to not always be tired or falling asleep in class. Kids would see a better world and not have to think oops i missed the bus no I'm going to be late for school.

Because of all the schools across the country that have changed for school to start a little earlier there test scores went up and students actually do better because they learned and not slept through classes.

SCORE 1 – MINIMAL RESPONSE

(CRITICAL-THINKING SKILLS)

most teens go to bed late and have to wake up early . going to bed can be very easy and can be very difficult . kids also like to stay up late because their never tired but when they say its time to go to bed t can be a very long time befor acauly falling asleep . same as geting up . its easy and feels beter most of the time to get up late on weekends (sleeping in) untill 9:00 or 10:00 but durring school days its hard to get up .

SCORE 1 – MINIMAL RESPONSE – 2nd example

(CRITICAL-THINKING SKILLS)

I have seen many teens fall asleep in class and I admit, I almost have a few times too. As we get older, school and sports become more demanding. Personally, I have been up till 12 o'clock at night doing homework. Also, sports make me tired, but I just can't seem to get to sleep on time. Having trouble getting to sleep may not just be in your teen years, but it can also greatly effect your adult years. "Over time, those nights of missed sleep can build into a sleep deficit. People with a sleep deficit are unable to concentrate, study, and work effectively. "(quoted from an article written by Dr. Pamela Smith) So when we have trouble getting to sleep, it is something to worry about.

Not getting enough sleep can also greatly effect school grades. Not getting enough sleep makes teens less likely to pay attention and give their work all the effort they have. I think schools should give kids an extra hour so that kids can be rested and ready to work.

"Teenagers & Sleep" TASK 4 – Create a Web page

SCORE 4 – EXCELLENT RESPONSE (CRITICAL-THINKING SKILLS)

Student's Response:

How many of you were tired your first period class? Raise your hand if you'd like to go to school later on in the morning rather than bright and early! Some schools have taken the plunge and decided to start school and hour later. A bunch of students have researched the topic to figure out if this is a positive idea or not. They have tossed all the info into an article that will be featured in the next edition of Teen Talk. You should definatly pick up a copy and take a peek at the article. There are so many questions about the topic that are answered in the article and a look at all the different side. This decision could affect you! Knowing about it can help you pick a side!

SCORE 4 – EXCELLENT RESPONSE – 2nd example (CRITICAL-THINKING SKILLS)

Student's Response:

Ever find yourself thinking about the fact that teenagers always look so tired? There's actually a reason for it. Teens usually never get enough sleep and end up falling asleep in class. Schools across the country have tried starting at a later time. Research was done by teens themselves about the topic, and in the next issue of Teen Talk, we'll tell you what they learned. If you're a teen, then you should definitly read the article to find information that could help you or your school change some things. Should schools start later, or should they keep depriving teens of their sleep? The country has different answers, and the one in your region could affect you. Watch for Teen Talk in June for the article, it could save you some sleep.

SCORE 3 – ADEQUATE RESPONSE (CRITICAL-THINKING SKILLS)

Student's Response:

During the teenage years, many things about us change, namely our sleep cycles. We often don't get the sleep we need to succeed, and thus our performance in school is compromised. Of the schools whom have already decided to start school later, all have seen dramatic improvements in student performance. Students have done reasearch on the topic, and will disect the whole matter in the next edition of Teen Talk. You should check ou the article to learn more about the hot-button topic. Should schools start later to allow teen the sleep they need, or not? There are different, view, opinions, and findings all across the board. It's up to you to read it for yourself and decide! Tune in to the June issue of Teen Talk to find out more.

SCORE 3 – ADEQUATE RESPONSE – 2nd example (CRITICAL-THINKING SKILLS)

Student's Response:

Do you always find yourself falling asleep in class, yawning and wishing you were at home in bed? If yes, then you are not alone. Many teens are not getting the right amount of sleep we need. Some schools decided to start classes at a later time. Do you think thats a good idea? A group of students throughoutly researched the topic and have been chosen to have an article in the next edition of Teen Talk and they will explain all of their knowledge on the subject. If you are interested in learning more about this issue then you should deffinitly take a peek at the article and choose your opinion. Should schools start later allowing teens to get the right amount of sleep? There are tons of different opinions, read the article in the June issue of Teen Talk and choose your side of the topic. There are tons of new things you could learn from reading the article. Remember you learn something new every day!

SCORE 2 – LIMITED RESPONSE

(CRITICAL-THINKING SKILLS)

Student's Response:

Students you may not know this but you nead more sleep then your teachers. Some schools are wonting to start school later to give you guys the extra hour of sleep so you guys can be awake in class. Who thinks thats good? If you research the topic then you can get your research in Teen Talk.

SCORE 2 – LIMITED RESPONSE – 2nd example

(CRITICAL-THINKING SKILLS)

Student's Response:

do you think that school should start later, do you want another hour to sleep do you want to destroy your alarm clock. Come read our next edition of Teen Talk and we will talk all about the problem with waking up early and what can be done to help. we want youre opinion, and who knows maybe you will be able to put youre alarm clock on snooze... forever. this article could greately effect youre grades and youre school year please join us.

SCORE 1 – MINIMAL RESPONSE

(CRITICAL-THINKING SKILLS)

Student's Response:

If you are like the million of people in this country, you get up, do the same rotation every day, and go to sleep. If you are fallin asleep in class, than what do YOU think the problem is? Not interesting topic in History? Already learned that in Math? If you fall asleep, than you are the type of person that could have sleep defecit. This occurs to many teens in Middle School and High School. So, if you think you need more sleep, talk to your principal about it, TODAY =)

SCORE 1 – MINIMAL RESPONSE – 2nd example

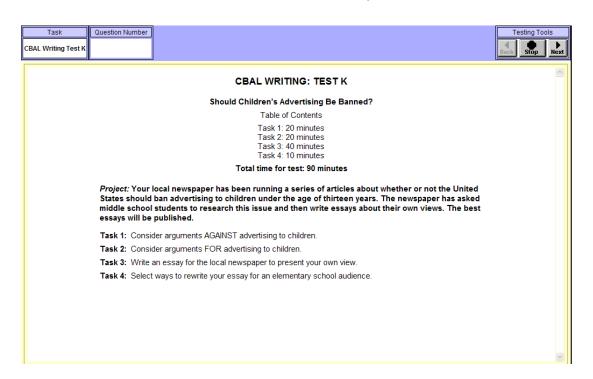
(CRITICAL-THINKING SKILLS)

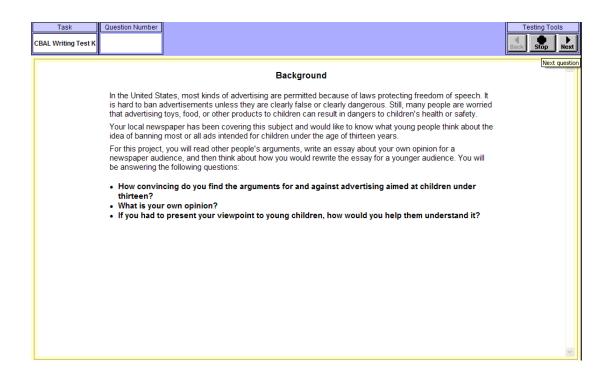
Student's Response:

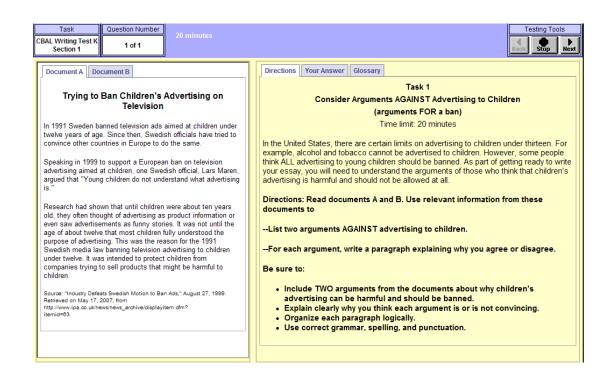
You know how people in school are sleeping? Well if you don't know you don't need to, because how would you like starting school later in the day, so you can sleep in and go to school 1 hr. later. It affects school by people who are tired and can't sleep. People like that need extra time on homework, and can't focus on after school activities like sports or other things.

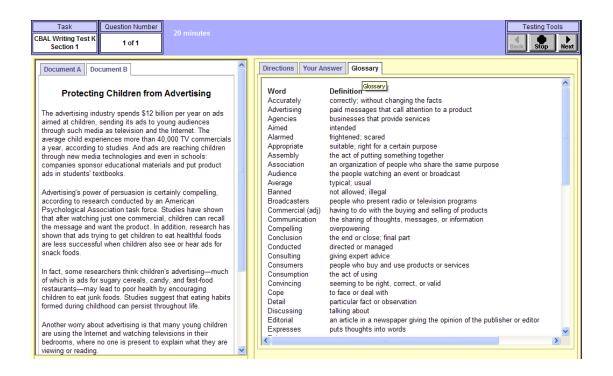
Appendix E1

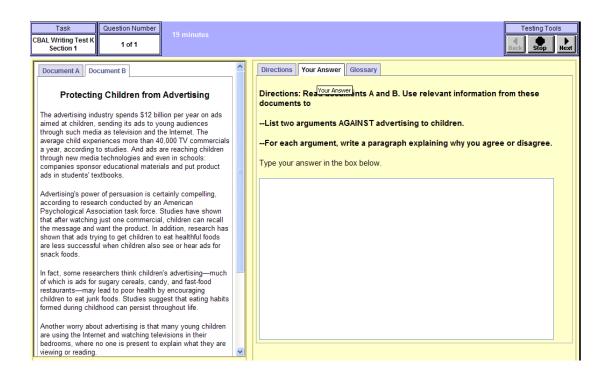
Screen Shots for Test Design #4, Persuasive Essay (*Advertising to Children Administered*Fall 2007; Scored/Analyzed in 2008)

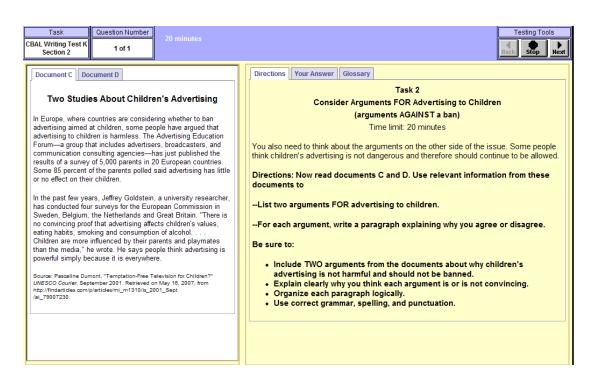


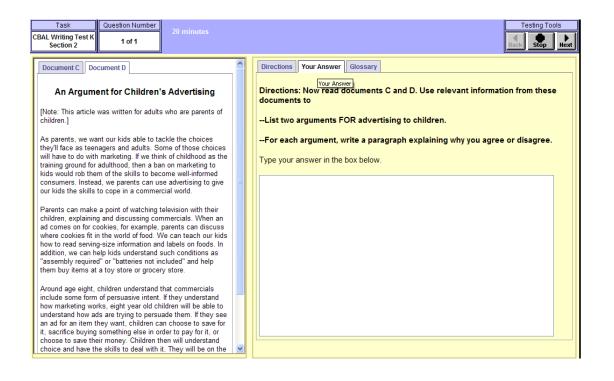


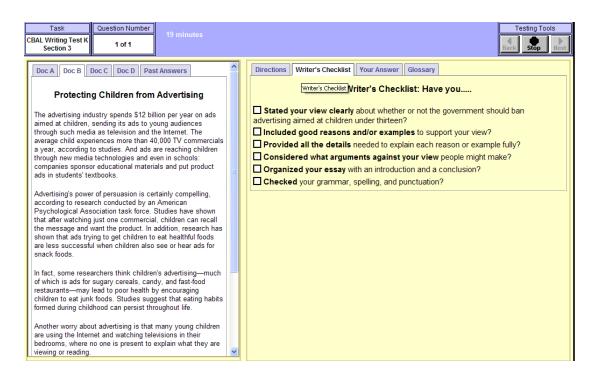


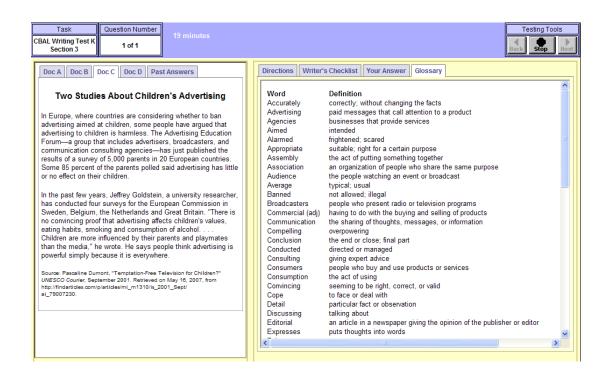


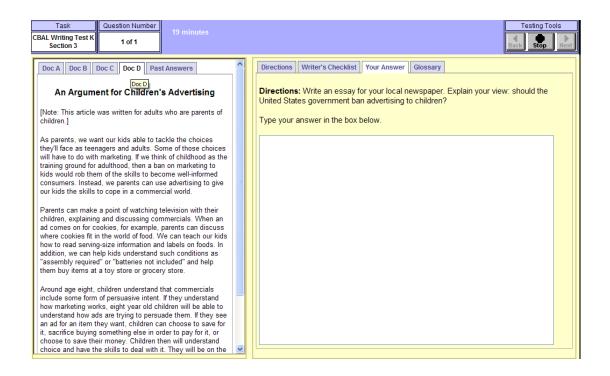












Appendix E2
Frequency Tables at the Strand Within Task Level for Test Design #4

Task 1 Strand I	Frequency	Percent	Cumulative Frequency	Cumulative Percent
1	5	10.87	5	10.87
1.5	10	21.74	15	32.61
2	12	26.09	27	58.70
2.5	9	19.57	36	78.26
3	6	13.04	42	91.30
3.5	4	8.70	46	100.00

Task 1 Strand II	Frequency	Percent	Cumulative Frequency	Cumulative Percent
0	1	2.17	1	2.17
1	4	8.70	5	10.87
1.5	7	15.22	12	26.09
2	14	30.43	26	56.52
2.5	4	8.70	30	65.22
3	13	28.26	43	93.48
3.5	2	4.35	45	97.83
4	1	2.17	46	100.00

Task 1 Strand III	Frequency	Percent	Cumulative Frequency	Cumulative Percent
1	14	30.43	14	30.43
1.5	8	17.39	22	47.83
2	11	23.91	33	71.74
2.5	6	13.04	39	84.78
3	4	8.70	43	93.48
3.5	3	6.52	46	100.00

Task 2 Strand I	Frequency	Percent	Cumulative Frequency	Cumulative Percent
0	2	4.35	2	4.35
1	4	8.70	6	13.04
1.5	10	21.74	16	34.78
2	10	21.74	26	56.52
2.5	5	10.87	31	67.39
3	10	21.74	41	89.13
3.5	5	10.87	46	100.00

Task 2 Strand II	Frequency	Percent	Cumulative Frequency	Cumulative Percent
0	2	4.35	2	4.35
1	8	17.39	10	21.74
1.5	5	10.87	15	32.61
2	12	26.09	27	58.70
2.5	8	17.39	35	76.09
3	6	13.04	41	89.13
3.5	5	10.87	46	100.00

Task 2 Strand III	Frequency	Percen t	Cumulative Frequency	Cumulative Percent
0	2	4.35	2	4.35
1	14	30.43	16	34.78
1.5	10	21.74	26	56.52
2	5	10.87	31	67.39
2.5	7	15.22	38	82.61
3	5	10.87	43	93.48
3.5	3	6.52	46	100.00

Task 3 Strand I	Frequency	Percent	Cumulative Frequency	Cumulative Percent
0	6	13.04	6	13.04
1	5	10.87	11	23.91
1.5	3	6.52	14	30.43
2	11	23.91	25	54.35
2.5	5	10.87	30	65.22
3	11	23.91	41	89.13
3.5	4	8.70	45	97.83
4	1	2.17	46	100.00

Task 3 Strand II	Frequency	Percent	Cumulative Frequency	Cumulative Percent
0	6	13.04	6	13.04
1	3	6.52	9	19.57
1.5	5	10.87	14	30.43
2	16	34.78	30	65.22
2.5	4	8.70	34	73.91
3	10	21.74	44	95.65
3.5	2	4.35	46	100.00

Task 3 Strand III	Frequency	Percen t	Cumulative Frequency	Cumulative Percent
0	5	10.87	5	10.87
0.5	1	2.17	6	13.04
1	5	10.87	11	23.91
1.5	7	15.22	18	39.13
2	11	23.91	29	63.04
2.5	7	15.22	36	78.26
3	6	13.04	42	91.30
3.5	4	8.70	46	100.00

Task 4	Frequency	Percent	Cumulative Frequency	Cumulative Percent
0	9	19.57	9	19.57
1	1	2.17	10	21.74
2	3	6.52	13	28.26
3	6	13.04	19	41.30
4	13	28.26	32	69.57
5	14	30.43	46	100.00

Appendix E3

Rubrics for the Advertising to Children' Test

Critical-Thinking Rubrics⁴⁸

KEY - TASK 4: Choose Ways to Revise for a Different Audience

Which of the following changes would be best for revising your letter to be read by fourth-grade students? Choose the FIVE correct revisions.

____ Add more information from research studies about advertising.
 ____X__ Find simpler words to use instead of long or difficult words.
 ____X__ Mention how you reacted to ads when you were the same age as these students.
 ____ Make the whole article much longer so that you can explain all your ideas in more detail.
 ____ X__ Give one or two examples of the kinds of ads that these younger students often see.
 ____ Focus the article on what parents think about children's advertising.
 ____ Make sure the whole article is funny and has lots of jokes in it.
 ____ Explain more about the differences between how products are advertised in the United States and how they are advertised in other countries.
 ___ X__ Make sure your main idea is stated very clearly at the beginning and at the end.
 ___ X__ Separate any long sentences into shorter sentences.

ADVERTISING TO CHILDREN TEST K - TASK 1

Consider Arguments AGAINST Advertising to Children

Task Directions:

- (1) Identify two arguments in the documents against advertising to children
- (2) Explain why you agree or disagree with each argument

4 EXCELLENT -- An "Excellent" response:

- Cites two arguments clearly, completely, and accurately
- Gives persuasive reasons for agreeing/disagreeing
- Includes no irrelevant or inaccurate information; may make insightful comments

3 ADEQUATE -- An "Adequate" response:

- Cites two arguments in a generally clear and accurate manner
- Gives relevant reasons for agreeing/disagreeing
- Includes little, if any, irrelevant or inaccurate information

2 LIMITED -- A "Limited" response is problematic in one or more of the following ways:

- e) Cites two arguments, but they are somewhat unclear or inaccurate
- f) Gives vague or somewhat unclear reasons for agreeing/disagreeing
- g) Includes a distracting amount of irrelevant or inaccurate information

1 MINIMAL -- A "Minimal" response is problematic in <u>one or more</u> of the following ways:

- a) Cites only one argument
- b) Cites two arguments, but both are seriously unclear or inaccurate
- c) Gives confusing and irrelevant reasons for agreeing/disagreeing
- **d)** Fails to cite any arguments from the documents

0 NO CREDIT -- A response receives "No Credit" for any one of the following reasons:

- k) Not long enough for critical-thinking skills to be judged
- 1) Not written in English
- m) Off topic
- n) Blank
- o) Only random key stroke

ADVERTISING TO CHILDREN TEST K - TASK 2

Consider Arguments FOR Advertising to Children

Task Directions:

- (1) Identify two arguments in the documents for advertising to children
- (2) Explain why you agree or disagree with each argument

4 EXCELLENT -- An "Excellent" response:

- Cites two arguments clearly, completely, and accurately
- Gives persuasive reasons for agreeing/disagreeing
- Includes no irrelevant or inaccurate information; may make insightful comments

3 ADEQUATE -- An "Adequate" response:

- Cites two arguments in a generally clear and accurate manner
- Gives relevant reasons for agreeing/disagreeing
- Includes little, if any, irrelevant or inaccurate information

2 LIMITED -- A "Limited" response is problematic in one or more of the following ways:

- a) Cites two arguments, but they are somewhat unclear or inaccurate
- b) Gives vague or somewhat unclear reasons for agreeing/disagreeing
- c) Includes a distracting amount of irrelevant or inaccurate information

1 MINIMAL -- A "Minimal" response is problematic in <u>one or more</u> of the following ways:

- a) Cites only one argument
- b) Cites two arguments, but both are seriously unclear or inaccurate
- c) Gives confusing and irrelevant reasons for agreeing/disagreeing
- **d)** Fails to cite any arguments from the documents

0 NO CREDIT -- A response receives "No Credit" for any one of the following reasons:

- a) Not long enough for critical-thinking skills to be judged
- **b)** Not written in English
- c) Off topic
- d) Blank
- e) Only random key strokes

ADVERTISING TO CHILDREN TEST K - TASK 3

Write an essay for the local newspaper to present your view.

Issue: Ban advertisements for children age 12 or younger?

4 EXCELLENT -- An "Excellent" response:

- Presents a well-defined position on the issue
- Supports the position with highly relevant reasons, examples, and/or evidence
- Is well reasoned, possibly including an effective critique of opposing points of view
- Argues in a clearly appropriate tone for the newspaper audience

3 ADEQUATE -- An "Adequate" response:

- Presents a clear position on the issue
- Supports the position with relevant reasons, examples, and/or evidence
- Is generally well reasoned, possibly including a critique of opposing points of view
- Argues in an appropriate tone for the newspaper audience

2 LIMITED -- A "Limited" response displays one or more of the following characteristics:

- a) Presents a vague or sweeping position on the issue
- b) Supports the position with few relevant reasons, examples, or evidence
- c) Gives a weak or undeveloped argument to support its reasons
- **d)** Argues in a somewhat inappropriate tone for the newspaper audience

1 MINIMAL -- A "Minimal" response displays one or more of the following characteristics:

- **d)** Does not present a position on the issue
- e) Supports the position with vague, irrelevant, or extremely limited reasoning
- f) Does not explain the reasoning
- g) Argues in an extremely inappropriate tone for the newspaper audience

0 NO CREDIT -- A response receives "No Credit" for any one of the following reasons:

- f) Not long enough for critical-thinking skills to be judged
- g) Not written in English
- h) Off topic
- i) Blank
- j) Only random key stroke

Appendix E4

Benchmark Responses for Advertising to Children Test

SCORE 4 – EXCELLENT RESPONSE (DOCUMENT-LEVEL SKILLS)

K_S3Q1

Will advertisements seen by young children effect them possitively? Parents around the country are wondering the same thing. Will the hours of comercials bennifit the children? No, advertisements should not be targeted at young children.

Advertisements can hurt childrens health. Most comercials targeted at kids are about food, specifically junk food. The marketing companies are clever and put in things kids like along with the food. Now kids will associate harmful foods with their favorite TV character or colorful song. Their eating habits will change do to this. They will now want the food because of the character on the box rather than the value of the food. This can effect their decision making for life.

Advertisements can effect childrens decision making. Kids will be exposed to comercials about all kinds of toys, foods, video games, movies, and other items. It's natural for the kid to now want these items since he/she has seen them in such a good light, with kids like them playing with them and smiling. But the kid will be sadly disapointed to know he/she can not get all these things. They will want the one that they like best, nag their parents to buy it, and find out it is no where near as cool as it seemed on TV. They will think the advertising is lies and never associate themselves with it again.

Advertsiements can confuse children. They will see things like cookies and pizza being eaten with fun, while seeing brocoli and vegatables being thrown into the garbage. Their parents do the exact oposite, so they will be at a delema on who to believe. If the child watches to much TV, they will grow to love and trust it. They will memorize jingles and comercial characters. But when the dentist says not to eat those foods, the kid will wonder why this ocmpany makes out the Cheeto Cat to be so great, if the dentist says not to eat his foods.

Overall, advertisements can hurt young childrens health, negatively effect their decision making, and confuse them overall in life. Young children should not exposed to advertisements.

SCORE 3 – ADEQUATE RESPONSE (DOCUMENT-LEVEL SKILLS)

K_S3Q1

One arguement for advertising to children is that the parents can watch television with their kids and explain the advertisements to their children. I agree with this. For one, if children see a cookie comercial, their parents can explain to them exactly what a cookie is and how healthy it is. Kids will now know not to eat too many cookies for future reference. For two, parents can help kids understand what certain terms mean like "assembly required." This is valuable because later on in adulthood they will see comercials and know what these terms mean.

Another arguement for advertising to children is that children will become smart buyers by watching advertisements, I agree. For one, kids will see all kinds of things they want in; comercials including toys, foods, video games, and others. They will slowly realize they can't have everything, and that they need to choose which item they want most. They develop valuable decision making skills at a young age. For two, they will be able to tell how good a product is from its comercial once they've seen alot of them. They can tell if a company put a lot of work in to their product, or if they are just trying to get your money

SCORE 2 – LIMITED RESPONSE (DOCUMENT-LEVEL SKILLS)

K_S3Q1

Children should be exposed to advertisements.

Advertisements are harmless to children. Kids are more likely to be influenced by their parents and peers than the media. Besides, commercials are fun to watch for kids, and it gives them a sense of new technology and ideas in the world.

Sure, some people say that some commercials tell children to nag their parents to buy the product, but parents should not stand for that. If the child nags the parents and throws a tantrum, the parents should be able to tell their children the commercial was not okay to say that, and the way their child is acting is also not okay.

SCORE 1 – MINIMAL RESPONSE

(1 b)

(DOCUMENT-LEVEL SKILLS)

K_S3Q1

I think that children's advertising should not be banned because there are commercials for kids that show healthy cereal that taste good so when the kids see other kids eating the cereal then they kids on t.v. say that the food is the best cereal they ever ate, then the kids watching the comerial will want to buy that cereal because it tastes good, and the parents will buy the food for them because the kids want something that tastes good and the adults want something that's healthy for their children. Most people would say that the companies are targeting the children to buy unhealthy products like candy and sugary junk food and popcorn even though popcorn is one of the mos healthy snacks, well after carrots that is. The comercials advertissments for popcorn would be like pocorn act2 or popcorn act3. The only difference between those two popcorn packages is that the company that makes that popcorn puts different amounts of butter and salt to make it taste better. So I think that advertising towards children should not be banned.

'ADVERTISING TO CHILDREN' TASK 1 - Consider arguments against ads to children

SCORE 4 – EXCELLENT RESPONSE (CRITICAL-THINKING SKILLS)

One argument against advertising to children is that children will remember the advertisements for snaqck foods and not healthy foods, thus leading to poor health. I agree with this argument for many reasons. One, children do not understand the effects these foods have on their bodies. They only want these foods for maybe the theme song of their commercial, or the color, or just the taste. There are many more snack food commercials than healthy food commercials. Children will watch the junk food ones more and want the foods.

Another argument against advertising to children is that young children don't know what adcertising is. I agree with this for the following reasons. Firstly, young children will watch cartoon channels like Nickelodean and Cartoon Network. These channels do a lot of advertising for snack foods. The kids don't realize that the companies are trying to get you to buy these, they just think they are amusing commercials with food. Since the children don't realize what the companies are trying to do, I find it unjust to target this population of people.

SCORE 3 – ADEQUATE RESPONSE (CRITICAL-THINKING SKILLS)

Two of the reasons why some people want to bann advertising to children are becasue most of the commercials that they see are trying to persuade them to but junk food, sugary cereals, or other sweet snacks. Studies have been taken and they show that what people eat when they're kids effects what they eat when they're grown up. Another reason is that some commercials try to get children to buy something by nagging their parents about it. The commercials bribe the kids by using character from popular kid's T.V. shows such as SpongeBob and Blues Clues.

The first argument against advertising to children, I don't agree with. I think that it just depends on the kid whether or not what they eat when he or she is young makes an impact on what htey eat when he or she is an adult. Everyone is different and some people find out a healthy way to eat once they've grown-up even if they didn't eat as healthy foods as they did when they were young.

I agree that using bribes- cartoon characters- for commercials isn't right. Also, the commercials that persuade the kids to have to nag their parents isn't good either. That can just make there more work for the parents and sometimes weaken the relationship between the child and his or her parents. For example, if the parents don't have enough money to buy their kid some toy they saw on a commercial it turns the parent into the bad guy if they have to say no. I don't think it's right for advertisers to do that.

SCORE 2 – LIMITED RESPONSE – 2 b (CRITICAL-THINKING SKILLS)

I think that we should absolutely not advertise to children. If we advertise to children possible outcomes could occure. For example, some researchers think that advertising to children, which is mostly about candy and junkfood, amy lead to very bad health by encouraging children to eat candy and junkfood. Many researchers suggest that the way children eat can stay with them throughout their lives. By looking at the facts, children should not be advertised until they are about over 12.

Another reason why we should not advertise to children is because research has shown that children about ten years old often think that advertising as product information or they even thought of advertisments as funny shows or stories. Research also shows that children that are about 12 years old fully understand what advertising is. In 1991, the Sweden media made a law banning television advertising to children under the age of 12. This law was made to protect children from companies that sell bad things. This is another reason why we should not advertise to children.

SCORE 1 – MINIMAL RESPONSE – 1 d

(CRITICAL-THINKING SKILLS)

Children should not be advertised to. When you advertise to children they are going to think that that product is apsalutly impossibal to live without. They just have to have it or it they won't be "cool". Its all about self assteem and peir pressure because you if you show a little kid a comercial on a certain ceral and it shows lots of kids eating it and saying "yumm" then the little kid is obvisly going to want to have it. When you advertise a child you are pretty much minipulating them.

What child do or see is completly different than a teenager or adult or even an elderly person would see. If an old woman saw a comercial for a weelchair that was really cheap and they needed one they would probally want to get it because it worked for all those other old peaople why not them? If an adult saw a comercial of ways to keep your life on the right track or something and they knew they were slipping up then they would probally want it and if you show a little girl a barbie doll thats skinnier than a pencil and boobs the size of the grand canyon of course shes gonna want to get it because then she can try to look just like her.

Advertising just shows people what they "should" be not who they are. I honestly dont think that there should be advertising for anyone let alone kids.

TASK 2 – Consider arguments for advertising to kids

SCORE 4 – EXCELLENT RESPONSE (CRITICAL-THINKING SKILLS)

A servay of 5,000 parents in 20 European countries says that 85% of those parents say that their children are not influenced by advertising. In my opinion some children may not be influenced but I think alot of them will be influenced, especially because of the big corperations that play to the kids and send the message that if you have this product you will be cool. If i child isn't influenced then I would say good job parents because the parents are most likley the ones that do sit down with the kid and explain how the ads are trying to persuade you to do somthing. Now if that same child were left to their own devices then there is no doubt in my mind that the child will fall victum to the marketers messages and that is why I don't think there should be advertising for kids because not all kids get the needed guidence.

In document D it says that parents can make a point of watching tv with their kids and explaning terms such as "assembally required" and "batteries not included" and that way the kids will get the good messages. I would agree that if the parent did that the child would bennifit, but again you have to think of the kids that just sit infront of the tv all day while theri parents are working or out doing somthing. If ads aimed at kids were banned then the parents could expline politocal ads or adds aimed at adults if they really wanted to and then the kids that don't have parents wouldnt be sucked in buy the evil empire of toys-r-us or McDonalds.

SCORE 3 – ADEQUATE RESPONSE (CRITICAL-THINKING SKILLS)

I agree with the fact now that there should be commerials and advertisements on television. In document C there isn't a lot of data and not a lot of info so when I read at first I disargee that you should put ads for children on television. There wasn't any research or anything that would show me that puting ads for children on t.v. is good. Also what happen if children aren't influenced by their parents and playmates than the media and why did you survey just Europe why didn't they survey people all over the world in countries that have a lot of children that watch t.v. like the United States, China, Canada, or any other country. This was did not persuade me and that is because they could of made this document longer or get more info.

In document D is where I now argeed about puting ads for children on television. When they said that parent can help children be smart at what to do when ads come on and be smart of buying a product of a company. They have excellent reasons but they don't have anything that would support this arguement, no research, no tests only reasons. But it did persuade me and it is a good arguement. It doesn't show what children would do on there own and how they know this and didn't do anything to prove this

SCORE 2 – LIMITED RESPONSE - 2 b (CRITICAL-THINKING SKILLS)

Advertising to children has little or no effect on them. A university researcher has conducted 4 surverys for 4 different countrys, nearly everyone said that advertising doesn't affect kids wants, eatings habits or drinking consumption. The kids mostly will get those habits from their parents.

There is no risk of kids ending up eating unhealthy foods or drinking alcohol or smoking, as long as their parents are a good influence to them.

Kids need to learn from their parents or school what some of the bad things that there are and what is on T.V. that could harm them. Kids also need to learn what labels on the back of some food are, like what's healthy to eat and what is good, but not healthy to eat.

SCORE 1 – MINIMAL RESPONSE – 1 d (CRITICAL-THINKING SKILLS)

Kids should watch some T.V but not to much like "ONE HOUR A DAY" but some parents allow there kids to watch T.V as much as they want to. True that Ads are every where even in school or outside when you walk home you will see some signs and then eventualy you would want what ever it tells you on T.V signs school so kids like eight years old if they are smart they can wether decide if they want to fall for the trick or not. Its all about the kids at that time.

A. Parents shouldn't listen to there kids sometimes if they wanna buy something like food they should have like an allowence 5 dollers every month so they dont waste money on junk food.

B. Parents should listen to parents and agree with them not like the parents will go soft on kids and they shouldn't like always get what there kids want they should like see if the kids do good at school home have good manners.

Document D.

lets say if a kid goes to a store and just buys a HUGE bag of Choclate chip cookie and dosen't look at the labels or signs just buys it everyday after a week the kid gains fat because he didn't look at the back side it tells you about the calories, fat, protien, Carbs...

So i mean if kids look at these stuff they can help them selfs and there parens

Just lookin at t.v all day will not help you do anything you will be lazy sleep watch t.v come eat go to school and then you fail grades thinking bout food and when they get home.

ADVERTISING TO CHILDREN TASK 3 – Write an essay to the newspaper

SCORE 4 – EXCELLENT RESPONSE (CRITICAL-THINKING SKILLS)

Will advertisements seen by young children effect them possitively? Parents around the country are wondering the same thing. Will the hours of comercials bennifit the children? No, advertisements should not be targeted at young children.

Advertisements can hurt childrens health. Most comercials targeted at kids are about food, specifically junk food. The marketing companies are clever and put in things kids like along with the food. Now kids will associate harmful foods with their favorite TV character or colorful song. Their eating habits will change do to this. They will now want the food because of the character on the box rather than the value of the food. This can effect their decision making for life.

Advertisements can effect childrens decision making. Kids will be exposed to comercials about all kinds of toys, foods, video games, movies, and other items. It's natural for the kid to now want these items since he/she has seen them in such a good light, with kids like them playing with them and smiling. But the kid will be sadly disapointed to know he/she can not get all these things. They will want the one that they like best, nag their parents to buy it, and find out it is no where near as cool as it seemed on TV. They will think the advertising is lies and never associate themselves with it again.

Advertsiements can confuse children. They will see things like cookies and pizza being eaten with fun, while seeing brocoli and vegatables being thrown into the garbage. Their parents do the exact oposite, so they will be at a delema on who to believe. If the child watches to much TV, they will grow to love and trust it. They will memorize jingles and comercial characters. But when the dentist says not to eat those foods, the kid will wonder why this ocmpany makes out the Cheeto Cat to be so great, if the dentist says not to eat his foods.

Overall, advertisements can hurt young childrens health, negatively effect their decision making, and confuse them overall in life. Young children should not exposed to advertisements.

SCORE 3 – ADEQUATE RESPONSE (CRITICAL-THINKING SKILLS)

I do not agree with the U.S. government banning advertizing to children. I think that children are more influenced by their parents so if the parents tried to eduacate their children and exsplain how the advertizer try to persuade people and that thing arent always the same as they are on television I think that it would influence their children to make good choices about what the advertizment is telling them.

Researchers are alarmed that commercials persuade children to want the product so they nag their parents to buy it. I think that if an adult was not eduacated and saw a commercial that appealed to them and they didn't have money and they had people that took care of them and bought them the things they needed they would do the same. It's not just kids who are influence by advertizing I think that all age groups from children to elderly people without education on advertizing and the media could be persuaded by the media. The kids just need to be educated and it's the parents job to educate them about stuff like this, not the schools. Snack food and candy are advertized showing fun things and famouse charactors like Spongebob Squarepants, like the cheetah for Cheetos, Tony the tiger the list could go on and I've never seen healthy food ads with mascots and fun ads. I think thats why maybe the snack food and candy ads are more apealing.

In europe a group made a survey for parents and 85% said that advertizing has little or no effect on their children. The other parent in the 25% should try to eduacate their children about advertizing and it persuasive messages, they could watch television with them and exsplain commercials to them like if a candy or snack commercial came on parents could exsplain where candy and snacks come in the world of food.

So I think that advertizement to children should not be banned but the parents should watch out for their children and educate them on things the school dosn't since it's their job to raise their child and teach them commen sence and how to deal with the world.

SCORE 2 – LIMITED RESPONSE – 2 a c (CRITICAL-THINKING SKILLS)

In my opinion, I think that kids should be allowed to watch advertisments on TV. Some advertisments have fancy words in them, that kids may not understand. When a child watchs TV, they are subjected to TV, whether they like it or not. Advertisments also broaden a child's imagination. The sad part is, that when it comes for a chld to use their imagination, it is often not accepted by adults, such as teachers. If a, adult does not like a child's idea of something, the adult may tell the child that. If a child hears something like that from an adult, it might crush the child's dream, and the adult won't even realize it. This society has always been the subject of advertisers.

It seems to me that there is a cycle going on. Children watch ads and think, "Hey! Maybe I could invent something like that one day." When the child starts to grow up, the adult society is starting to teach the child about more important things, such as good ways to get into college. When a child tells an adult about an invention, that the child thinks, might work, the adult disagrees with the child. The adults in this world think that there are more important things, than listen to a child's idea of something. Then what happens, the child grows up to own a company that sells the child's invention. At some point, the adult that was supposed to hear the child's idea, sees an ad selling the child's invention on TV, and the adult thinks "Why didn't I listen to that child, when he/ she telling me about that invention?" The adult probably would have made a lot of money selling the child's invention in the first place. It is basic karma.

I don't think that the United States should ban advertising to children. Our youthful society needs to be subject to more than just school and work. I think that a lot of great inventions have come from people that have seen other ads on TV, and used those ads to put something even better, together.

I guess what I am trying to say is that, TV ands won't turn a child's mind to mush, it will help the child succeed.

SCORE 1 – MINIMAL RESPONSE – 1 b c (CRITICAL-THINKING SKILLS)

You should ban advertisment because all these crazy things like dont do drugs kids dont care about not doing it. They say hey what the heck with it lets try it. kids at are school might think where coolbut then the they would try it once more then they would get hooked on it. Almost 60% of high schools does drugs because of stress of family problem the Commercial does not help they see it on tv and from friend that saw it on tv there probly say hey that's a good way to take away the pain but it doesn't most of teen death is from doing drugs and driving and where does it come from ADVERTISMENT's on tv.

Appendix F

Principles of the CBAL Initiative

- Accountability tests, formative assessment, and professional support will be derived from the same conceptual base. That base will be built upon cognitive research, state standards, and curricular considerations.
- CBAL assessments will consist largely of engaging, extended, constructed-response tasks that are delivered primarily by computer and, to the extent feasible, automatically scored.
- Because of their nature, CBAL tasks should be viewed by teachers and students as worthwhile learning experiences in and of themselves. Ideally, taking the test should be an educational experience and preparing for it should have the effect of improving student domain competency, not just improving performance on the test.
- Accountability assessment will be distributed over several administrations throughout the school year so that: (1) the importance of any one assessment and occasion is diminished; (2) tasks can be more complex and more integrative because more time is available for assessment in the aggregate; and (3) the assessments provide prompt interim information to teachers while there is time to take instructional action.
- For accountability purposes, estimates of student competency will be aggregations of information collected over time. In addition to these competency estimates, the accountability tests will offer some formative information to teachers, students, parents, and local policymakers.
- Results from the accountability assessments will serve as initial information for more extensive formative or diagnostic assessment, indicating such things as the competency level, area(s) in which follow-up is suggested, and initial formative hypotheses. (However, CBAL formative assessments will never be used for accountability purposes.)
- CBAL formative assessment will be designed to help teachers engage students in a structured process that reveals evidence about what students know and are able to do, helps teachers and students identify the characteristics of proficient performance, and moves students toward developing competency. CBAL formative assessment will include classroom tasks and activities, resource materials, and diagnostic tests. Most components of CBAL formative assessment should be adaptable by the teacher for use when and how the teacher sees fit.
- CBAL assessments should be designed to help students take an active role in their own learning and the assessment of it.
- While the three CBAL components derive from the same conceptual base, they should be able to function independently of one another. So, for example, states should be able to purchase the accountability system without also having to purchase the formative system.

Appendix G

Administration Procedures

The writing test consists of one 90-minute section, to be given over two class periods. 5-10 minutes before the end of the first class period, ask students to finish up what they're immediately working on. When they come to a stopping place, ask them to wait for 2 full minutes before they close out of the test, to allow the system time to save their work. On the second day, follow the instructions below for reconnecting. Check every student to make sure that they have followed instructions and clicked on the checkbox, or they will overwrite their work from the previous session.

Logging In

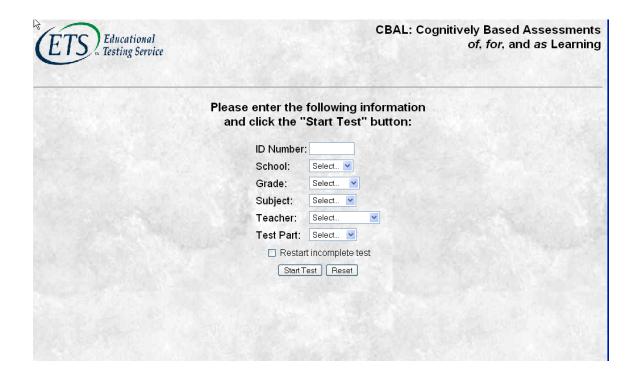
<u>Step 1:</u> Students must be using Firefox for the test to open correctly. The URL to log into the test is **http://naeptba.ets.org/cbal.aspx** If you see a screen that has an Open Test button, click on the button and the Log-In Screen will appear.

Attention

If you do not see a new window with the test sign in screen, you may have a pop-up blocker that is preventing the window from opening. Click on the button below to open the testing window.

If you accidently close the testing window, click on the button below to reattach to the testing session.

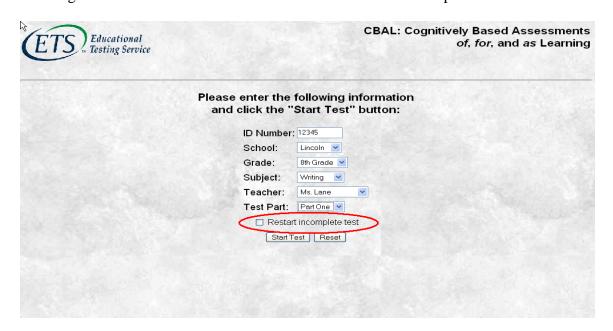
Open Test



Step 2: Students will type in their 4 or 5 digit Student ID number.

<u>Step 3</u>: Use the drop down menus to select the School, Grade, Subject, and Teacher. There is also a drop down menu for "Test Part". Students taking the Writing test should not select anything from this menu. Students taking the Math or Reading tests should select "part 1" on their first day of testing, and "part 2" on their second day of testing.

Warning!!! Make sure students **DO NOT** click the "restart incomplete test" checkbox.

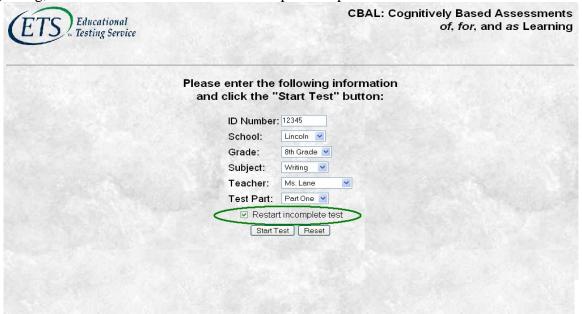


Step 4: Click the "Start Test" button.

Reconnecting

If students are disconnected from the system, follow Steps 1-3 for logging in. Then click the checkbox for "restart incomplete test", before clicking the "Start Test" button. Students should return to the last question they were working on.

Warning!!! If you do not click the "restart incomplete test" checkbox, it will start over at the beginning, and the work the student has done up to that point will be overwritten.



Quick Hints



Student's Screen Freezes: to refresh a student's screen use

Student is unable to type on the writing test after having finished part of an essay:



A Printscreen will be put on the student's desktop and you can paste this into Appleworks so that the student's work up to that point will not be lost.

Appendix Notes

- ⁴³ Thus far, human scoring of "diagnostic" categories keyed to the lowest levels of the competency model has not been particularly reliable; see discussion below for each individual pilot.
- ⁴⁴ The descriptions have been somewhat compressed; however, the elements outlined here, and outlined in the corresponding sections in Deane et al. (2007) are intended to match as described.
- ⁴⁵ This version of the rubric for sentence- and document- level skills has been revised to reflect changes in the competency models. All other tests were used with scoring guides in place from earlier versions of the rubric.
- ⁴⁶ The same benchmark examples for document level skills were used for both the service learning and the "advertising to children" PAAs.
- ⁴⁷ Sentence level and document level skills (Strands I and II) were only assessed for the essay task on the Teenagers and Sleep test.
- ⁴⁸ Other rubrics for so-called Strand I and Strand II are the same as for the service learning test.