Research, Consultation, & Teaching Program
Training Module No. 12

RCTP

Scoring Student Writing Using Intellectual Operations: Examining Student Understanding Within Content Areas

Dewayne E. Joehnk
Gerald Tindal
Victor Nolet

University of Oregon College of Education, Behavioral Research and Teaching,
Eugene, Oregon, 97403-5262

Published by
Research, Consultation & Teaching Program
Behavioral Research and Teaching
College of Education
University of Oregon

Editorial Staff
Abe Deffenbaugh, Editor
Jencina Butler
Traci Milliren

Joehnk, Dewayne E.; Tindal, Gerald; Nolet, Victor
Scoring Student Writing Using Intellectual Operations: Examining Student Understanding Within Content Areas
Training Module No. 12

Notes and Acknowledgments:
This project is funded by a Leadership Training grant (H023D30087–Developing Individuals who Reframe and Enhance Content Teaching) from the Office of Special Education Programs, Personnel Preparation, United States Department of Education. However, the opinions expressed herein do not necessarily reflect the position of the U.S. Department of Education or the College of Education at the University of Oregon. No official endorsement by the Department, College, or University should be inferred.

This module could not have been written without the enthusiastic support and work of the teachers and students who participated in the projects. We would particularly like to recognize several individuals who helped in the development of the scoring systems and in the preparation of this module: Lisa Howard for her help on the initial scoring systems, Abe Deffenbaugh for his overwhelming patience and guidance on the manuscript; Denise Swanson for her incredible organization of the data; Carrie Wegmann and Sandra McCollum for their invaluable input on the scoring systems; and Sevrina Tindal for her attention to detail in various aspects of production.

Layout: Abe Deffenbaugh
Cover design: George Beltran

"Writing permits us to understand not only the world but also the self. We discover who we are by writing"
(Lindemann, 1987, p. 7)

As Lindemann so aptly noted, writing is a complex culmination of skills and knowledge that provides a unique reflection of our individual perspectives. How often have we as writers, researchers, and teachers written something we thought clean and polished, only to later return and change the writing because of new experiences and new reflections we've had since initially laying down the words? Our growth, our learning, and our experiences shape the way we write, and in turn, these shapers of our writing may be seen through what we write.

In this training module, we present a series of scoring systems that use student problem-solving as an indicator of their understanding of content area concepts. Our eventual goal in this module is to develop a method of "scaling" student responses to content tasks, however, we are constrained by a number of different perspectives that must be addressed before describing the specific scoring systems. In this introduction, we will present an overview of these perspectives in order to establish a context for judging student understanding through specific writing tasks. Following this introductory section, we provide examples of student work in three areas of understanding that require students to either (a) make evaluative judgments and develop sound arguments, (b) arrange content information to arrive at conclusions and explain why, or (c) use content information to make predictions. At the end of each section is a short check-out section in which you may practice the scoring systems on several actual student writings. The examples of writing you will see were written by middle school students of various ability levels from both science and social studies classes.

Although the three operations we explore herein are certainly not an exhaustive list of the myriad possibilities of how student understanding may be reflected in writing, these three tasks can potentially be operationalized easily and in a low inference manner. As a result, the scoring systems are grounded in deeper constructs and can be quickly employed while excellent reliability is maintained.

THE CONTEXT

While educational assessment literature is replete with calls for performance-based measures, few if any studies have been conducted on strategies for scoring those performances. Most of the articles on performance-based assessments are conceptual, describing and suggesting strategies that "sound" good (have face validity), but have unknown psychometric underpinnings. As a result, performance-based assessments are continually under fire for poor development and even poorer reliability.

We begin with the assumption that to understand student understanding, the assessment system likely needs to be a production task. Although multiple-choice tests may be used to make summative judgments about students' knowledge (Bennett & Ward, 1993), we believe such measures fail to adequately reflect the ways in which students use their knowledge to arrive at conclusions, develop arguments, or make judgments.

We also argue that written problem-solving essays may be the easiest and most straightforward way to capture student understanding, although this assertion must be validated. At the very least, we believe that written responses to problem-solving tasks may be more "authentic" than multiple-choice type tests because the writing represents several critical skills for success in middle and high schools. However, for authenticity to accrue, "it must be intimately understood by teachers, students, and parents, so that it can help them strive for and achieve the learning goals it embodies" (Darling-Hammond, 1994, p. 25).

The first step to creating an assessment that is embedded within systemic understanding of student understanding, is development of technically adequate assessment systems that are both reliable and valid. This problem, however, is twofold. First, there must be emphasis upon determining the relative consistency and generalizability of the system (reliability), and upon the consequences of any decisions that are made using the system (validity). But secondly, there must be emphasis on establishing an empirical and evidential basis for each specific performance measurement. In the system explained in this training module, we approach this need for technical adequacy as minimally requiring that reliability, especially interjudge reliability, be well documented for any assessments used to make judgments. This is particularly true for assessments used in making high-stakes decisions. "Comparability becomes important when claims are made that students are meeting common performance standards or when results are used for external purposes such as certification or admission to college" (Linn, 1994, p. 12).

What we have found in our reviews of the literature is an incredible paucity of research linking performance assessments to a strong empirical base. For example, in a recent keyword search of the scores of articles listed in the ERIC databases (using "performance assessment" as a descriptor), we found only a handful of articles in which empirical data were reported. Further, within this handful of studies the results have not been encouraging:

> Ironically, the reliability of performance estimates in many alternative assessment systems has been low. For example, the Vermont portfolio assessment program reported very low consistency in the interjudge scoring of the writing and math samples; coefficients ranged from .33 to .43 (Koretz, Stecher, Klein, & McCaffrey, 1994). In the Oregon State-wide Assessments, there were few exact agreements on 6 writing traits (ideas-content, organization, voice, word choice, sentence fluency, and conventions); in general, fewer than 50% of the writing samples were given the same exact score by more than one reader (and this was true across all the grades of 3, 5, 8, and 11) (Oregon Department of Education Technical Report: 1989-1992)" (Tindal & Nolet, in press).

Such poor consistency is not just a function of the raters, but it may also result from task incomparability. Shavelson, Baxter, & Gao (1993) suggest that a range of 10-23 performance tasks in science may be needed to obtain a stable estimate of a student's performance.

One final problem with the current research on performance assessments is the lack of attention to the constructs underlying the student performance being measured. Presently, the major concern is with the tasks themselves, not with the requisite interactive intellectual or learning operations that underlie the student's performance. For example, most critiques of traditional measurement systems limit the debate to a comparison between multiple choice versus constructed responses (Bennett & Ward, 1993). Yet, within constructed responses, so many variations in task format exist that it is difficult to make comparisons among different studies or to consider discussion of any intellectual or learning operations inherent in those tasks. Nevertheless, consideration of those underlying operations is important for grounding student performance in "deeper" constructs that can, for psychometric purposes, be separated from content. However, in order to maintain instructional relevance, consideration of student content knowledge must also be possible.

With these two issues in mind—the technical adequacy of a performance measure and the format of the learning operation addressed—we sought to establish the beginnings of a scoring system flexible enough that teachers could use it within the classroom, yet steadfast enough to be psychometrically rigorous.

When we began, five premises focused our efforts, much like the criteria noted by Deno (1985) in the original development of the basic skills Curriculum-Based Measurement (CBM) system:

1. The system had to be capable of generating alternate forms so a time-series measurement system could be generated. In the area of content assessment, then, not only did the responses have to be examinable in a somewhat generic form, but the information upon which they were based had to be broadly defined in their domains. This requirement is probably the most important, and this is what distinguishes the system we describe in this module from other forms of criterion-referenced measurements (which are usually topically bound and incapable of more than one alternate form).
2. The measures had to be technically adequate, meaning reliable (consistent, stable, repeatable) and valid (related to other measures and useful in the classroom so the teachers could use the information to improve instruction or, as Frederiksen & Collins, 1989 describe it, the measures had to "systemically valid").
3. The system had to be fast and easy to use, making it "user friendly" and possible to use frequently in alternate forms (again emphasizing a time-series nature).
4. The scoring system had to be low inference with little reference to cognitive events or complex constructs that emerge only within a constructivistic framework.

5. The scoring had to include a response that was instructionally relevant, regardless of how "authentic" that response appeared outside the school setting.

With these five premises in mind, we developed an assessment system with three available performance options for anchoring the measures to interactive learning tasks that could be modeled by the teacher or taught to the students. As a part of this system, we developed a flowcharted scoring method that served to simplify decisions about a student's content area knowledge. These flowcharts also served to increase the scoring reliability measures. We focused our efforts on three specific learning and performance tasks, all of which use written responses within problem-solving formats. Although the construct validity of such differentiations may have yet to be established (Tindal, Rebar, Nolet, & McCollum, in press), we felt that the most productive way to proceed was by treating each of them separately.

The first performance task is *evaluation.* To perform an evaluation task, students must use their content knowledge to make a choice from among several offered possibilities that are equal in cost and benefit. The students must then defend their choice by using their content knowledge to create an argument for their choice and against the possibilities that were not chosen. The second performance task is *prediction.* In a prediction response, students use content information to devise a logical outcome for a prompted situation. The students must relate the outcome they predict to the prompted situation by carefully constructing a chain of events. The third task is *explanation.* In an explanation response, students must explain why a prompted outcome has occurred using their content knowledge to reconstruct the probable chain of events that led to that outcome.

These three intellectual operations were first described by Roid and Haladyna (1982) within the context of test development. They were later researched by Nolet and Tindal (in press, 1994) in the context of student production responses. Rather than looking for mere use of a key word or definition in student writing (reiteration of facts), by examining *how* the students use their content knowledge to create a response, student understanding of the content is addressed. Performance becomes a reflection of their understanding of the content by how they use the information within the framework of a "deeper" intellectual operation.

What follows in this module are three major divisions, each devoted to one of the intellectual operations: evaluation, prediction, or explanation. Within each major division are three sections: (a) an examination of the intellectual operation being assessed with scoring flowcharts and explanations of the criteria used for scoring, (b) multiple examples of student writings from two content areas with scoring analyses of each example, and (c) a check-out section for trying your hand at scoring several essays.

In each of the flowcharts, the lowest value response is one that is not in response to the prompt, and although indicative of something, it is uncertain what it reflects. Otherwise, the scoring systems reflect slightly different nuances of argumentation. In evaluation, we emphasize

multiple arguments both for the choice and against any nonchoices (Toulmin, 1958). In the prediction flowchart, we focus on logical outcome derivatives. And in the explanation scoring system, we concentrate on sequential delivery. In the end, we believe the flowcharts are a way to begin understanding student understanding that does not dissect the whole, while retaining the integrity of judgments with the least amount of inference.

SUMMARY AND CONCLUSION

The scoring systems presented herein are reflective of student performance on tasks that are grounded in the intellectual operations evaluation, explanation, and prediction. This grounding in "deeper" constructs allows a psychometrically rigorous examination of student knowledge and fluency. Using these intellectual operations as prompting frameworks, the students are assessed on their ability to use the content materials they have studied. Students must variously use their content knowledge to: (a) evaluate several prompted choices by establishing content-based criteria for making judgments, (b) predict a likely outcome of a prompted set of antecedent circumstances, or (c) explain a situation by describing antecedent conditions that were necessary to bring about the prompted situation. Each of these operations demand extensive use of content specific knowledge in various ways that are accurately quantifiable.

In our past experience with the presented scoring systems, we have continually measured inter-rater reliabilities of .85 to .90 for exact agreement on scores, and reliabilities of .90 to 1.00 when scores are compared within one point plus or minus. The inter-rater reliability of these measures is quite high and extremely consistent for performance assessments of this kind. This lends considerable credibility to the notion that higher technical adequacy on performance measures may be achieved through establishing the performance tasks in "deeper" constructs such as the learning operations.

Additionally, these flowcharted systems may easily be adapted to alternate forms to allow time-series measurements and broadly defined domains. This creates a well-grounded though flexible scoring system that may be used in a wide variety of activities. Teachers using the scoring system have reported that it is easy to use, rapidly employed, adaptable, and rigorous: causing them to focus on specific operations and content rather than upon a student's poor grammar and penmanship. It has been suggested that the flowcharts themselves may be useful for training students to write more effectively. It is hoped that teachers using this system will find other ways to use the materials herein. The scoring system is quite adaptable to course content making it, therefore, instructionally relevant by integrating content knowledge and writing skills while providing a more reliable measure of student understanding.

Evaluation Scoring

Section A

SCORING EVALUATION ESSAY RESPONSES

Evaluation is the analysis of a problem using appropriate criteria to make a judgment. The primary consideration when examining evaluation essay responses is to determine whether the student has made a choice between several prompted options by applying appropriate criteria to judge those choices. Implied in this statement are three constructs crucial to understanding the form of the evaluation essays. First, multiple options must be available to students from which a choice can be made. These options must be presented to the student via the prompt. If only one option is available or if all of the prompted information points strongly to one choice, students will merely be summarizing information on one concept. The prompt must force a choice, not shape it, so the students can display their knowledge about the concept being taught.

Secondly, students must choose the option they consider the best path to follow. To do this the students must establish criteria for analyzing the available options. These criteria reflect previous knowledge and principles, concepts, and facts that were presented to them through class content and curriculum. These criteria will be used to justify their choice. In essence, necessary criteria cannot be in place if the student doesn't know the content materials.

Finally, students must present their decision and explain their rationale. Ideally, student presentations should come as arguments that overtly weigh and compare possible choices using content derived criteria. Students must use content knowledge, make decisions through choosing criteria and a course of action, write an effective argument for their position, and support that choice by presenting evidence constructed through use of the criteria. This exercise becomes an overt display of their content knowledge and learning.

It must be noted here that evaluation operations go beyond what is typically known as a comparison and contrast question. Comparison and contrast questions are summarizations of factual information that only require students to recall information and compare and contrast similarities and differences.. Comparison and contrast questions do not ask students to use that information. This latter operation is only offered in an evaluation-type question.

The scoring system presented here reflects two dimensions of this process that are inherently necessary to the task: the position the student has taken and the student's use of the facts, concepts, and principles to support that position. By careful analysis along these two dimensions, three scores may be obtained. Students may be scored on their use of correct and incorrect facts as evidence for their position, and they may also be scored for the quality of their overall decision and argument. The student evaluation answer presented in this article has been examined and scored using the flowcharted scoring system appearing below (see Figure 1). The critical features of evaluation scoring

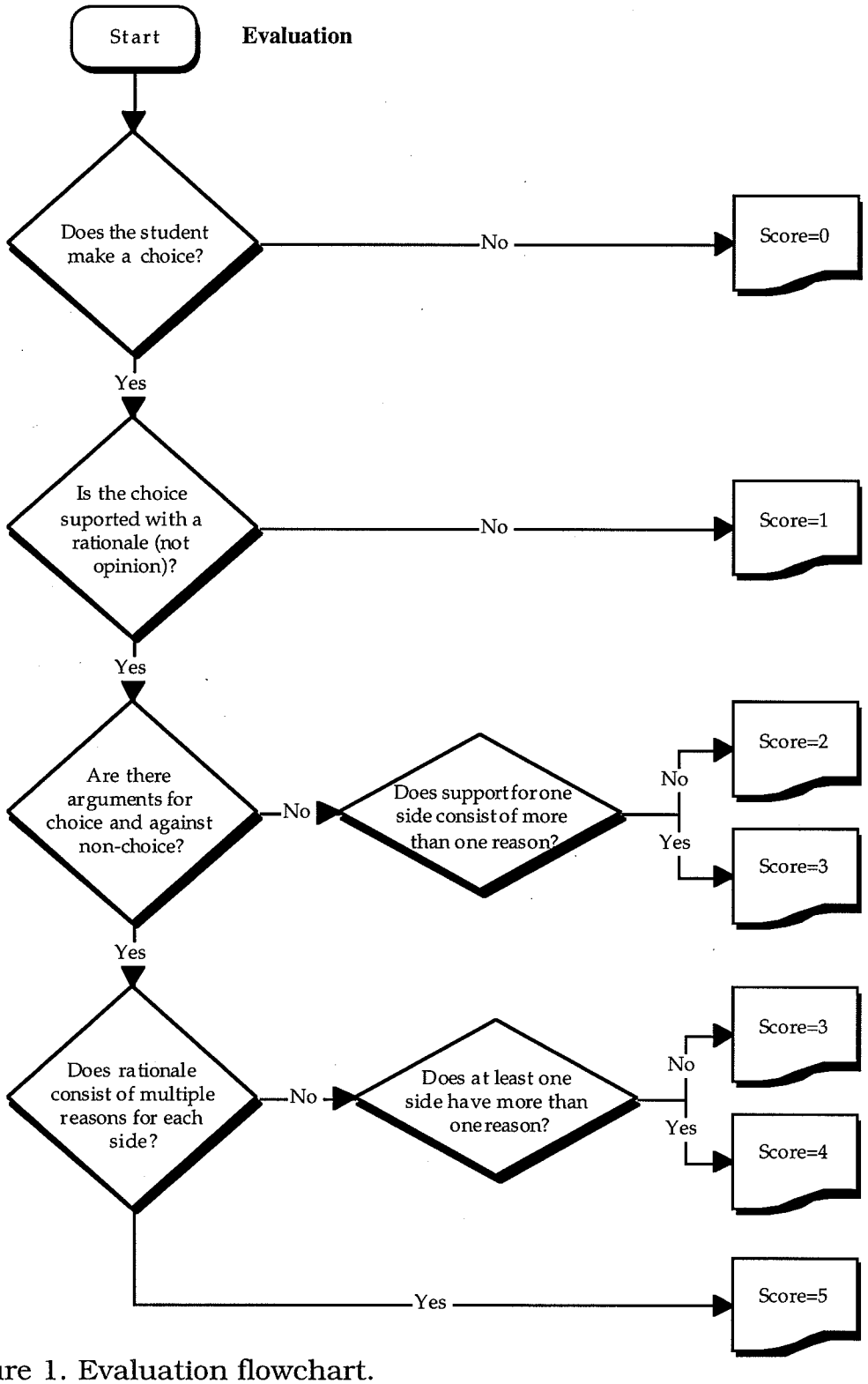


Figure 1. Evaluation flowchart.

are: (a) presentation of knowledge through the use of facts, concepts, and principles as criteria for judgment, (b) student presentation of a choice in response to the prompt, and (c) the presentation of a rationale for that choice, consisting of multiple facts supporting the choice and refuting other, unchosen possibilities.

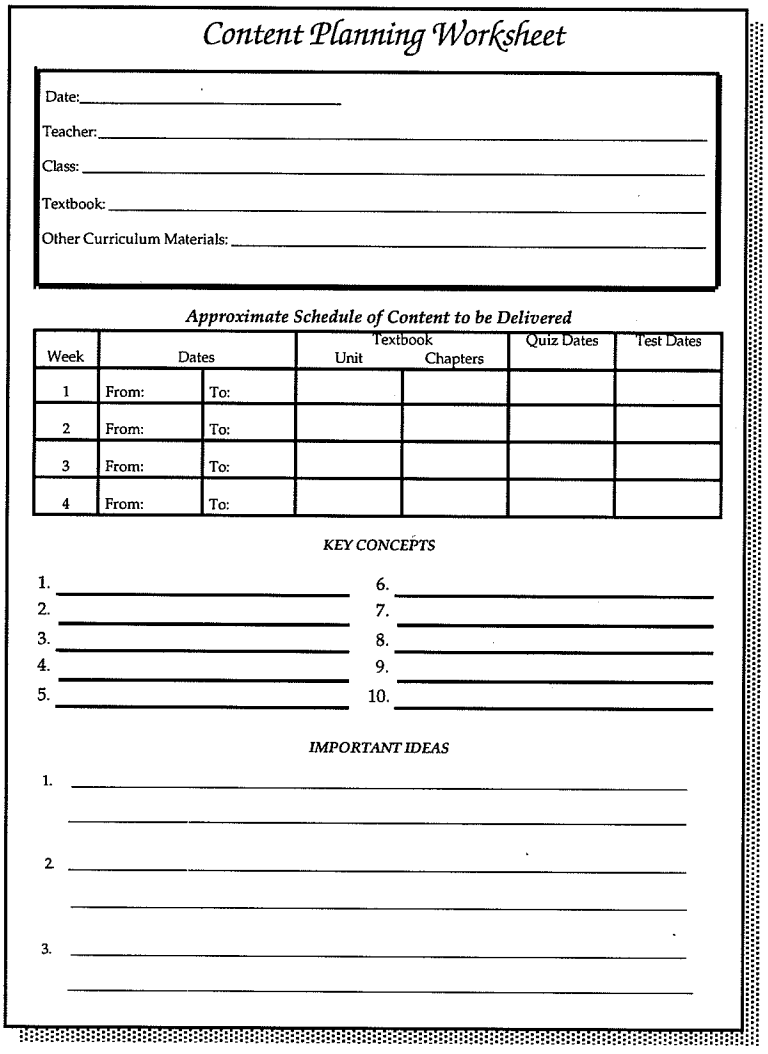
PRESENTATION OF KNOWLEDGE: FACT SCORES

Facts form the foundation for good evaluation essays. Items of content knowledge from the class or from previous experience are used as criteria to weigh choices and to derive a "best" choice from among multiple presented options. Students must examine the prompted information, process it in the light of their knowledge about the topic, make a choice, and present reasons for their choice by describing how it fits their criteria. At each step the use of known facts, concepts, and principles is critical to the development of an effective argument for their choice. This use of facts, concepts, and principles may be scored separately from the quality of the argument. Facts are most commonly presented in the development of this argument. Also reference may be made to concepts and principles. Facts represent only one example or instance. Concepts and principles represent complex constructs using multiple facts. In the case of student essays, they are rarely developed beyond a name or brief description. Because of this, for scoring purposes, concepts and principles are considered facts.

For example, in a social studies prompt students may be asked to choose a course of action for controlling immigration in the United States. The options may be quotas, strict health screening, or education requirements. There is no absolutely correct choice from among these options. Students must decide which choice would best fulfill the task according to their knowledge, then they must defend that choice. The student may choose quotas and present instances of previously used quota systems (facts). The student may also refer to the frontier thesis (principle) in the answer when describing a need for room to grow. Either the former facts or latter principle can be counted as facts in support of the student's choice. In scoring there are two kinds of facts that may be used in scoring: correct and incorrect facts.

CORRECT FACTS

Correct facts (FC) accurately reflect course content and correct previous knowledge. These items will usually be listed on the content planning worksheet (see Figure 2) or they will be nearly synonymous to content listed thereon. However, students can occasionally throw the scorer a curve by inclusion of correct previous knowledge that was unanticipated. Scorers should use their best judgment in deciding how such information is scored. Decisions on how to score such anomalous



Content Planning Worksheet

Date: _____

Teacher: _____

Class: _____

Textbook: _____

Other Curriculum Materials: _____

Approximate Schedule of Content to be Delivered

Week	Dates		Textbook Unit	Textbook Chapters	Quiz Dates	Test Dates
1	From:	To:				
2	From:	To:				
3	From:	To:				
4	From:	To:				

KEY CONCEPTS

1. _____ 6. _____
2. _____ 7. _____
3. _____ 8. _____
4. _____ 9. _____
5. _____ 10. _____

IMPORTANT IDEAS

1. _____
2. _____
3. _____

Figure 2. Content planning worksheet (Tindal, Nolet, Blake, 1992)

information should be arrived at before scoring. Note, that student opinions are not considered facts and therefore they are not scored as either facts correct or facts incorrect. If a student argues that their choice was made "because it looked good" or "because it was right," those statements do not add evidence to the argument and therefore do not count as facts in support of the argument.

INCORRECT FACTS

Incorrect facts (FI) are defined simply as inaccurate information. FI may represent mislearning, guessing, simple mistakes, or a genuine lack of understanding. FI will be considered in scoring only if they were used as building blocks for the logical argument in the qualitative portion of the score. FI used in this manner are not scored as FC but merely considered as a logical building block the student used to

strengthen the framework of the logical argument. FI can also give a teacher valuable insights on individual and global mislearning or misunderstandings. If several students mention the same incorrect facts, then it's a pretty sure bet that something was miscommunicated in the class: the students could have misunderstood the materials or task, the teacher or text could have presented an error, or perhaps a small group of students shared wrong information amongst themselves. By monitoring the level and type of incorrect facts that occur, a teacher can carefully monitor the accuracy of the learning taking place within the classroom.

SUMMARY OF FACT SCORING

The scores for FC and FI are mere tallies of the number of facts used in the essay. The tallies present limited information about the factual content of the student's essay. Consideration of facts is necessary in nearly all steps of evaluation scoring because facts are intertwined in the whole of the argument being presented. However, to assess the quality of the answer the framework of the argument must be examined. This overall view of the student's essay is examined when deriving the qualitative score.

PRESENTATION OF CHOICE: THE QUALITATIVE SCORE

The *qualitative score* (Q-score) is considered a direct reflection of the logical framework of the answer. Q-scores are derived from students' choices and how those choices are supported by the facts. Students must include some statement presenting their position on the prompted task, and there must be a framework of facts, possibly made up of both FC and FI, that is used to support that position. As stated earlier, the critical features of evaluation scoring are: (a) presentation of knowledge through the use of facts, concepts, and principles as criteria for judgment, (b) student presentation of a choice in response to the prompt, and (c) the presentation of a rationale for that choice, consisting of multiple facts supporting the choice and refuting other, unchosen possibilities. We have examined how facts appear in a student's answer. After such facts are counted, their use to create a logical argument of support can be examined.

THE STATEMENT OF CHOICE

The first step in analysis of the Q-score is to find a statement indicating that a choice has been made between prompted alternatives. That is, the student must take a position. This statement should be clear and the position taken definite. For example, students can write, "I think people should use fossil fuels," or, "People shouldnít use fossil fuels." The statement of choice becomes the hinge point of the student's evaluation of the problem. If the statement of choice is clear, the student immediately scores one point for inclusion of the statement. If

the choice is unclear or if no choice is presented, then no points are awarded for taking a position and further criteria are used to derive the score. This feature corresponds to Toulmin's notion of a proposition (1958).

Because of the nature of evaluation prompts, two unprompted choices may be found while scoring: definite decisions that no choice is best, or that all of the choices are good. These possible choices may make the answer seem equivocated or vague. Such answers can occur because the student does not understand the task or has no knowledge of the prompted situation, or because the student truly feels that is the best course and so does not want to commit fully to one of the offered choices. The former situation yields an answer that is vague. The latter situation will yield a clearly directed answer that is supported by facts. This type of answer can appear as (a) a presentation of a hierarchy of several choices, or (b) as a choice of taking no position because none of the offered choices are clear-cut.

As an example of a hierarchical presentation, if students are prompted to choose where people will settle from among several landforms, the student could write, "Most of the people would settle near the river but some of the people would settle in the mountains. Others would settle in the forest." This statement addresses several landforms, implying a hierarchy evident in the adjectives "most" and "some."

When students choose not to accept a single position it presents scoring problems. The student may say, "None of the choices are best." This, in fact, may be a clear position statement. Whether it is or not depends upon the support the student gives that statement. There are three ways in which this may be in evidence. First, the prompt itself may have been too vague or presented too much evidence for all of the choices (over or under prompted) making the taking of a single position too exclusive. Or perhaps the student has an excellent understanding of the topic and has over-analyzed the situation, realizing that taking a position in one way will go against other presented evidence. Both of these positions will be supported by a great deal of evidence. Finally, the student may even create a new choice or category of their own to facilitate mixed information. In any of these cases, the essay should be scored as containing a clear statement of choice. If the student uses such a statement and offers little evidence in support of that position, it may be because the student doesn't understand the prompt or is confused by it. That student's answer should be treated as having no clear position offered and scored accordingly.

FACTS AS SUPPORTING EVIDENCE

The student's rationale for making a particular choice is analyzed by examining the facts used to either support the statement of choice or lower the value of other choices. Scoring is accomplished by a simple count of the facts employed and whether or not other choices were considered. For the purposes of the Q-score it doesn't matter whether FI

or FC are used in support of the student's choice. What we are examining here is how those facts support the position.

The first decision to be made is whether or not the student addressed alternative choices. The assumption is that addressing both or all sides of an issue or decision creates the strongest argument. Therefore, students must present their reasons for their choice and their reasons for not choosing something else. A fact count of both sides of the issue is performed. If the choice is clear but the student presented only a single fact in support of one side of the issue, the Q-score is 2. If the choice is clear and the student offered one fact in support of both sides or multiple facts in support of one side, then the score is 3. If multiple facts were offered on one side of an argument that addresses both sides, then the score is 4. And ideally, if the student offers multiple facts in support of both sides of an issue, the Q-score is 5.

SUMMARY OF EVALUATION SCORING

In scoring student evaluation essays there are three scores that may be derived. The first scores, FC and FI reflect correct and incorrect facts, concepts, and principles that are presented in the essay. The scores are merely tallies of these facts. Correct facts are defined as facts that accurately reflect course content. Incorrect facts may be defined as facts that represent mislearning or misunderstanding and as such they should be monitored. Following tally of these scores, the Q-score is derived by examining the essay for a statement of choice that presents the student's position. The rationale consisting of these supporting facts is then examined for how the facts support the choice made, how other choices are refuted, and how a body of evidence is built through the use of multiple facts. What follows are examples of student answers accompanied by an analyses of those answers, demonstrating this method of scoring the evaluative and factual properties of student essays. The examples presented herein were chosen as representative of many of the issues that can arise during scoring.

ANALYSIS OF EVALUATION

SCORING EXAMPLES

The following essays were generated in two different content areas within several classes. The sixth grade science essays address the use of fossil fuels, while the seventh grade social studies essays address land forms and settlement patterns. The latter prompt included a map that students filled out in conjunction with their written essays. Each of the two sections consists of (a) a list of attributes used to identify facts for scoring the essays, (b) the prompt, and (c) several examples of student

answers divided into high, medium, and low groupings. Following these two sections there is an exercise section with student written responses.

SIXTH GRADE SCIENCE

The list below (Figure 3) was used to identify facts used in scoring the following science essays. If facts were presented in different parts of the answer as forms analogous or parallel to one another, then they were counted only once. For example, saying that fossil fuels cause pollution (general) and that there may be oil spills (specific), two analogous statements are made, and as such, they are scored as one fact. However, if a student points out that there may be acid rain caused by burning coal (specific) and that oil spills can occur (specific), two statements of fact are presented and two facts are scored.

SIXTH GRADE RESPONSES ON THE USE OF FOSSIL FUELS

USE FOSSIL FUELS

- Fossil fuels make life easier because of the capacity to do more work
- There are ways to use fossil fuels wisely and carefully (i.e., conservation)
- Trees may run out if they are over-harvested
- It is a waste not to use fossil fuels
- Fossil fuels may be easy and cheap to develop

LEAVE FOSSIL FUELS ALONE

- Fossil fuels are non-renewable resources
- Getting and using fossil fuels can cause pollution (i.e., acid precipitation, oil spills)
- There are good alternatives to fossil fuels available
- Getting fossil fuels may pose health risks (i.e., black lung, cancer)

Figure 3. Fact statements for and against the use of fossil fuels used in scoring sixth grade science essays.

WELCOME TO NEWTOPIA

Newtopia is a small planet in a galaxy not far from here. It is covered with large trees, the sun shines most of the time and a steady wind blows all year long. People on Newtopia cook and heat their homes with wood from the trees and use animals to do farm work. Gathering firewood and taking care of the animals is hard work so they have very little time to relax and enjoy the beauty of their planet. Recently, scientists have discovered large deposits of fossil fuels on Newtopia. There are oil reservoirs, coal seams, and natural gas deposits.

Some of the people on Newtopia would like use these fossil fuels to heat their homes, and cook their food. They also want to build refineries to make fuels to power tractors, cars, and trains. These people want to make life easier so they will have time for the finer things, like art and music.

Other people on Newtopia do not want to use the fossil fuels. These people are afraid of what will happen to their planet if people start burning coal, oil, and natural gas. This disagreement between the people who want to use the fossil fuels and the people who don't want to use them has caused many arguments and the people of Newtopia don't know what to do.

What do you think the people of Newtopia should do? Should they use the fossil fuels to run their cars and heat their homes or should they leave the fossil fuels in the ground and keep things the way they are now? Place an X beside the statement that tells what you think the people of Newtopia should do.

____ Use the fossil fuels.
____ Leave the fossil fuels alone.

Write an essay that tells why you made your decision. Tell what information you used to make your decision. If you think the people of Newtopia should use the fossil fuels tell why. If you think the people of Newtopia should leave the fossil fuels alone, tell why.

Figure 4. Prompt for the sixth grade science essay.

In the following examples, the statement of choice used in the qualitative analysis is italicized and fact statements used for that argument are underlined.

In scoring this essay, incorrect facts were not considered as facts and were not scored. They were, however, considered for logical argument purposes in the Qualitative score (Q-score) if they occurred.

STUDENT SAMPLES
HIGH SCORES

RANDALL

> The reason I chose to *leave the fossil fuels alone* is because if they use fossil fuels the planet will be ruend there will be air pollution caused by the gases in the fossil fuels. If they don't use the fossil fuels they may not have time to enjoy there planet but at least they will have good health. If they pollute it by using the fossil fuels there health will be poor or even died.

ANALYSIS

The statement of choice is clearly stated: "leave the fossil fuels alone." Multiple facts, addressing both sides of the issue, support this statement. Two facts support the choice the student made and one addresses a counter argument. The facts are:

- the planet will be ruend there will be air pollution.
- they may not have time to enjoy there planet.
- they will have good health.

The Q-score was derived from first having a clear statement of choice. In addressing both sides of the issue, the student created a stronger argument. Had the student considered one more fact to deny counter arguments, then the student would have scored "5" on qualitative measures.

Qualitative Score	Facts Correct
4	**3**

REBECCA

> *I think the people on Newtopia should use the fossil fuels*, so the can heat houses, or make tractors or cars so the have transportation around the planet. What are the going to do in the ground mite as use the more than waste the fossil fuel. even though the fossil fuels wil make polution.

ANALYSIS

A statement of choice is clearly stated: "I think the people on Newtopia should use the fossil fuels..." The answer contains multiple facts and addresses two sides of the issue. There are three legible facts presented in the argument. Two facts are presented for the positive side and a third addresses the counter argument. They are:

- heat houses. (Choice support)
- make tractors or cars so the have transportation. (Choice support)
- though the fossil fuels wil make polution. (Counter argument)

Because there was a counter argument that considers the reasons against the choice, the effectiveness of the argument was enhanced. In scoring, it was assumed that the answer, if more legible, would have suggested that rather than waste the fuel by leaving it in the ground, they may as well use it even though it pollutes the environment. Heating homes was treated as a separate fact because it was a totally different application for fuel from transport, addressing shelter needs.

Qualitative Score	Facts Correct
4	**3**

MEDIUM SCORES

JERRY

> The reason I chose *leave the fossil fules alone* is because <u>if people use the fossil fuels, their air with be poluted</u> with the fumes from the fossil fules. If they start using fossil fules they won't want to stop and <u>all of the fossil will be all used up.</u> Then they will have to <u>wait one million years so that the fossil fuels can develope.</u> People would get sick from the pollution in the air. They would get lazier and lazier as they use up the fossil fuels and when (fossil fuels) they are used up the people won't want to start taking responsibility's like they used to have before the fossil fuels were used up. That's why they should leave the fossil fuels alone.

ANALYSIS

A decision is clearly stated: "leave the fossil fules alone..." There are two facts given in support of the decision. They are:

- if people use the fossil fuels, their air with be poluted.
- all of the fossil will be all used up.

In the second fact, the student clearly understood that fossil fuels are a nonrenewable resource. In addition to the two facts, a third fact, though not entirely accurate, to "wait one million years so that the fossil fuels can develope," shows that the student understands that it takes a long time to develop fossil fuels. A fourth idea was presented, the people taking responsibility, however, it was opinion and unclear so it was not scored as fact.

Qualitative Score	Facts Correct
3	**3**

CHRIS

> I think the people of Newtopia should *leave the fossil fuels alone* because the fossil fuels would make easier but if I were a Newtopian I would rather die in a cool, clean, easy, climate than a hot, ugly, smoggy, polluted, high technology climate. I think beauty is more important than a T.V. or, hair dryer, or microwave. If they leave the fossil fuels alone then the planet will stay beautiful, If not then the planet will become a nightmare rather than a dream.

ANALYSIS

A choice is clearly given: "leave the fossil fuels alone..." Two straight-forward facts are presented:

- fossil fuels would make easier.
- hot, ugly, smoggy, polluted, high technology climate.

The first fact was assumed to mean life would be easier. This fact addressed the other side of the argument (for the use of fossil fuels). The second fact addressed the choice made by the student. Therefore, there was one fact offered about each side of the argument. A single reason on multiple sides. The remainder of the answer was merely opinion and was not counted as fact.

Qualitative Score	Facts Correct
3	**2**

LOW SCORES

LOWELL

> I think they should just *leave the fossil fuels alone* because some people like the smart ones know that if they do start buming the fossil fuels it well start pollution and make the beatiaful plante yuchy and if the people want to use the fossil fuels are unhappy they can just come to earth they want pollution just come to earth.

ANALYSIS

In this essay a choice was clearly presented: "leave the fossil fuels alone..." It is supported by one fact:

- buming the fossil fuels it well start pollution.

A strong statement of opinion was offered but was not considered for scoring purposes.

Qualitative Score	Facts Correct
2	**1**

WANDA

> FOSSIL FULES. I thik that they sould *leave the fossil fules alone.* The reason I thik that they sould leave the fossil fules alone is so they Don't ruened there nice imvierment. The fossil fules will make the weather to where it is not so good and it will create smog and bag air. fossil fules are hasertes to are imvierment becouse of there gas and the gas that is created when we use.

ANALYSIS

The statement of choice is obvious: "leave the fossil fules alone." In support, the student offers three statements:

- so they Don't ruened there nice imvierment.
- fossil fules will make the weather to where it is not so good and it will create smog and bag air.
- fossil fules are hasertes to are imvierment.

These are all statements that are parallel forms of the same thought (i.e., fossil fuels cause pollution). The statement concerning air pollution differs slightly because it is the only form of pollution mentioned. The student focussed on only this aspect of the issue and consequently, the statements were only scored as a single fact.

Qualitative Score	Facts Correct
2	**1**

WILL

> I think the people who want to use them than I think should use the fossil fuels. The people that don't want them. then they don't have to use the fossil fuels if they don't want too! because no one's forcing them to use them. No one shoud force someone to use or not to use the fossil fuels. So that's *my choice for some people to use the fossil fuels and some not to use the fossil fuels.*

ANALYSIS

A clear choice is offered: "my choice for some people to use the fossil fuels and some not to use the fossil fuels." There are no facts offered, only opinion. With only opinion in support of the position, this answer was scored as an unsupported choice, but because the student stated a position, their Q-score was "1."

Qualitative Score	Facts Correct
1	**0**

SEVENTH GRADE SOCIAL STUDIES

A map was included with this prompt, and students were asked to draw dots to show where they thought people would settle, placing more dots in areas of higher population density and fewer dots in less densely populated areas. This graphic representation of the students' answers was used to (a) verify what the student's answer implied if a scorer was unsure, or (b) if the student referred to the map in the answer.

In scoring this essay, parallel and analogous forms were scored as facts because of the differing features each landform presents. Incorrect facts were scored as both incorrect facts and as building blocks for the Q-score if those facts were used as a part of the logical framework. In the case of scoring this essay group, if a student wrote an opinion in support of the statement of choice, then that opinion was scored as an incorrect fact.

Because of the construction of the prompt, multiple statements of choice appear in some answers. The statements of choice usually reflected (a) a single major choice with a higher population density centered on one landform (most often near the river or bay), (b) secondary statements of choice where there were one or more smaller population groups in other areas (usually the mountains or forest), and (c) possibly tertiary statements about a sparse population or lack of population in the remaining areas (the desert or unidentified areas).

The list of facts developed for this prompt reflects landforms and land use. The students were taught that each landform offered people different things. The facts used in scoring are presented in Figure 5. Scores in the qualitative area were awarded for facts given in support of choosing a particular landform, and facts given for not choosing another landform.

LAND FORM FACTS FROM SEVENTH GRADE SOCIAL STUDIES FOLLOW-UP ESSAYS

RIVERS AND SEA
- Food–fishing, agriculture, irrigation, game
- Transportation, trade, and recreation
- Moderate temperature and rainfall

MOUNTAINS
- Game, water, recreation, mineral resources
- Cooler temperatures and rainfall
- Difficult transportation

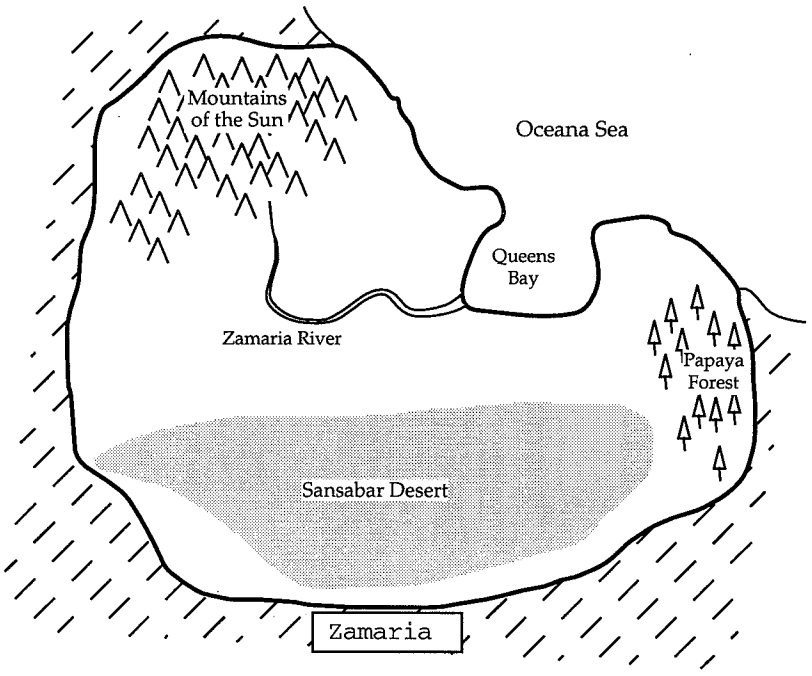
FORESTS
- Game, trees for heating and housing, recreation
- Rainfall

DESERTS
- Some natural resources
- Little water, game, or food
- Temperature extremes

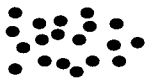
Figure 5. Land form facts for seventh grade.

ZAMARIA

Zamaria is a land that time forgot. No humans have lived there for many centuries. Recently a team of explorers rediscovered Zamaria and now people want to settle there. It is important to know where people will settle so that railroads, sea ports, and cities can be planned. Here is a map of Zamaria. Look at the map and decide where you think people will settle.



Show where you think people will settle by drawing dots to show population density. In areas where you think many people will settle, draw many dots, like this:



In areas where you think fewer people will settle, draw fewer dots, like this:

In the areas where you think almost no one will settle, don't draw any dots.

Write a short essay that tells why you completed the map the way you did. Tell why you think people will settle in the areas where you drew many dots. Tell what things affect where people settle.

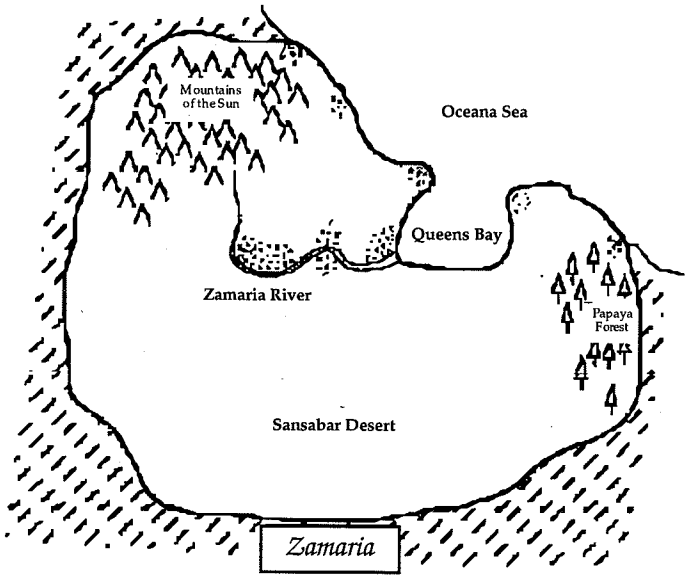
Figure 6. Seventh grade social studies prompt.

In the following examples, the statements of choice used in the qualitative analysis are italicized and facts used for that argument are underlined.

In scoring this essay, incorrect facts were scored and considered for logical argument purposes in the qualitative score.

HIGH SCORES

TRINA



I completed the map the way I did because *many people I think would settle around the Zamaria River* because it would suppily <u>water, transportation for them</u> and <u>transportation of goods to trade</u>. It would also <u>suppily food like fish</u>, and it would also keep land fertil because <u>it would bring minerals from up river and fertilize the land</u>. *I also think people would settle near Papaya Forest* because it would suppily them with <u>wood</u>, <u>meat from animals</u>, animal skins, <u>shelter and many other resources</u>. *I don't think as many people will live in or near the Mountains of the Sun or the great Sansabar Desert* because <u>their isn't as much fertil land, resorurces or water</u> they could use to help them survive for a long lifetime.

ANALYSIS

The argument presented in this answer is quite well developed. In it we see an initial statement of choice: "many people I think would settle around the Zamaria River..." Five facts are offered to support this statement alone. They are:

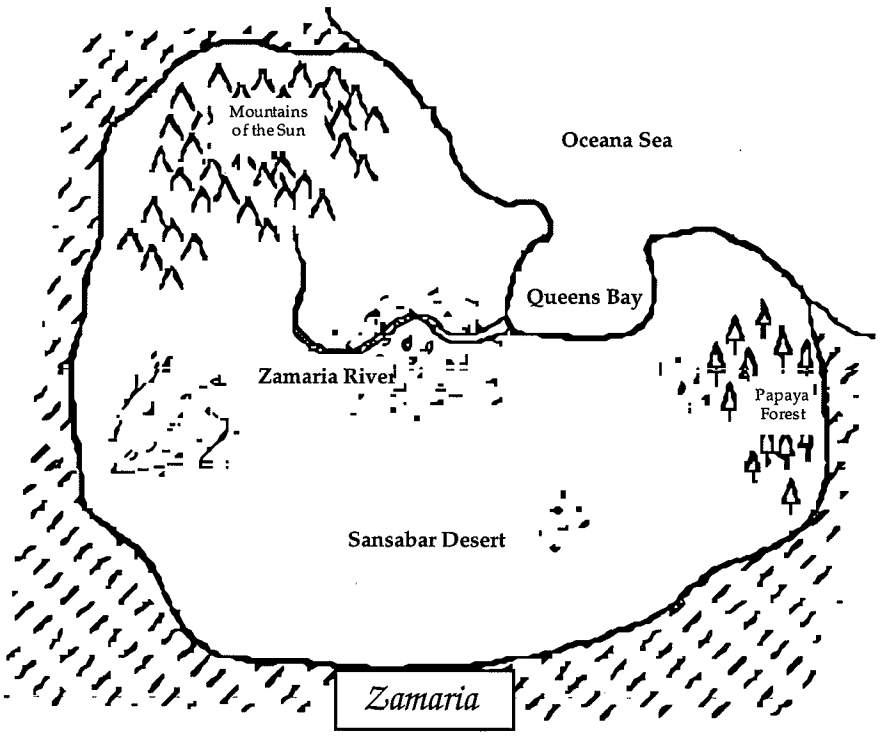
- water (for multiple uses including drinking)
- transportation for them (the people)
- transportation of goods to trade
- suppily food like fish
- it would bring minerals from up river and fertilize the land (agriculture)

A broad array of information is presented. The initial position is solidified by attention to each of the other landforms present on the map. This attention comes in the form of two additional statements of choice that allow the student to develop further arguments. First, "I also think people would settle near Papaya Forest," supported by mention of resources, wood, meat from animals, and shelter. Second, "I don't think as many people will live in or near the Mountains of the Sun or the great Sansabar Desert," supported by "their isn't as much fertil land...or water." This framework is solid in its analysis of the problem. The student presented an initial statement of choice, presented two statements of choice that solidified the initial choice, and then each of the three statements was supported with a considerable number of facts.

That the desert does not have fertile land and that rivers bring minerals from upstream could have been treated as parallel or analogous, but for the purposes of this student's response it would have made little difference. In addition to these facts are some points that are attempted as support but make little difference to the argument as a whole. Resources in general are referred to twice, and when writing of the game in the forest, the student mentions animal skins, a byproduct of hunting, along with shelter.

Qualitative Score	Facts Correct	Facts Incorrect
5	**10**	**0**

MINDY



I marked the map the way I did because *I think alot of people would live in the forest* because of the peace and quiet and *I think they would live in the Mountains* for the fresh air. *I don't think they would live in the dessert* because the climates are two dry and they would not be able to grow food

ANALYSIS

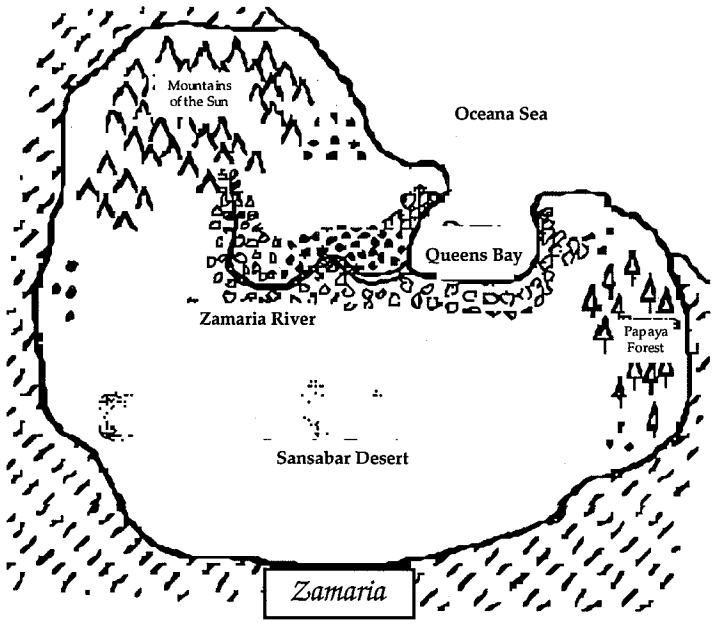
The main statement of choice in this answer is "I think alot of people would live in the forest..." and "I think they would live in the Mountains..." The choice is clear and distinct as also seen in the map drawn by the student. In support of this, the student offers facts that the forest offers "peace and quiet" and the mountains offer "fresh air." This answer is typical of one in which the student utilized incorrect facts in the form of opinion or unsubstantiated fact to support the statement of choice. The argument is hollow, but nevertheless, the framework for the argument is present. In addition, the student mentions, "I don't think they would live in the dessert because it is two dry and they would not be able to grow food."

In this case, the student scored 4 points on the Q-score even though half of the argument is based upon incorrect facts and opinion. The student gives multiple reasons for making the choice: the incorrect assumptions about peace and quiet and fresh air, and the correct fact against making another choice, the desert is "two dry" and it is difficult to grow food. Therefore, the student offers multiple reasons on one side of the issue and one reason on the other – the criteria for a score of "4."

Qualitative Score	Facts Correct	Facts Incorrect
4	**2**	**2**

MEDIUM SCORES

SCOTT



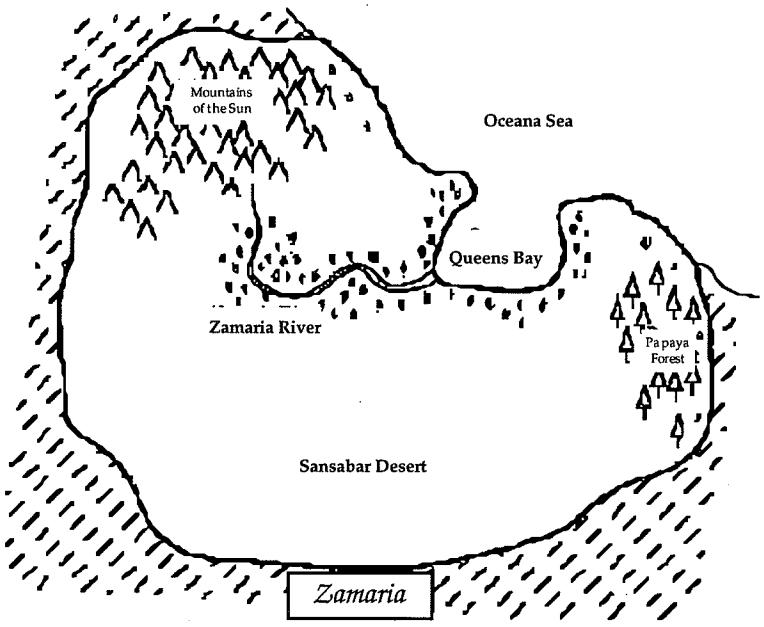
I think *people will settle in the areas that I put alot of dots* because they would have some resources, food & water. One thing that could affect people living by the ocean is that during a storm the ocean's tides could over flow land. But if you live by a river you can have water to water your crops.

ANALYSIS

The statement of choice, "people will settle in the areas that I put alot of dots..." is driven by the map and it is necessary to view it in order to understand the student's decision. Upon examination of the map it may be seen that the student put dots in the mountains, the forest, and around the river. In support of this, the student offers that "resources, food & water" are available in these areas. The second fact is derived from the idea of irrigation near a river. This is treated as an additional fact because it goes beyond the first fact statement. This student received a Q-score of "3" because multiple statements of support were given to the choice.

Qualitative Score	Facts Correct	Facts Incorrect
3	**2**	**0**

JENNY



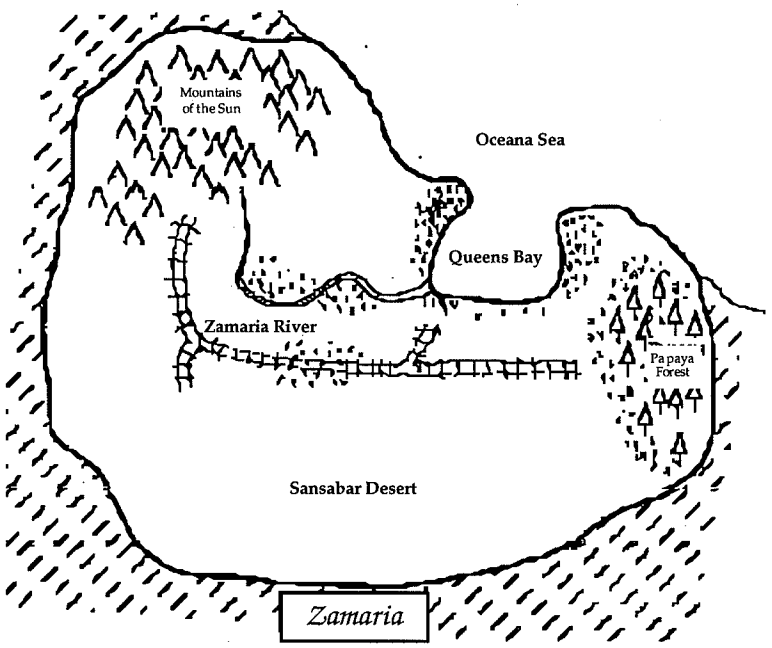
I *think lots of people will live around the river* because they need water for many things. *Lots of people might live around the Queens Bay* because there is alot of water there. *I don't think there will be any people living in the desert or the forest* but they could live around the mountains. The main place for them would be by the river, another reason is because they would build their cities around water to get things that are being shipped on boat.

ANALYSIS

This student addressed all land forms, however, statements about the mountains, the desert, and the forest were insubstantial. The reasoning for living near the river or Queens Bay was the availability of water for many uses. If this had been all of the support offered, the student would have scored a "2" on Q-score. Support was weak until the student added on that people build cities around water for shipping purposes. This last statement added substance to the student's argument to score a "3" on Q-score.

Qualitative Score	Facts Correct	Facts Incorrect
3	**2**	**0**

CARA



Zamaria would have *most of its population around the river and bays. People most likely would not choose to live near or in a desert* so a fewer density of people would live there. *Fewer people will pick to live in the middle of the Papaya forest* because they would miles from other civilization. The *density by the edge of the papaya forest would be very great.* There would be forestry, and water to help with the resources they needed. Although most likely the greatest density would be by the river and bay because there would be water, lush forests, places to build irrigation for farms, natural resources, and in the bay area that would probably be a big city which has boat landings, and other city needs. In between Zamaria river and the Sansabar Desert there would probably be a small city and train station. Up in the mountains would be small hick towns, except towards the bottom.

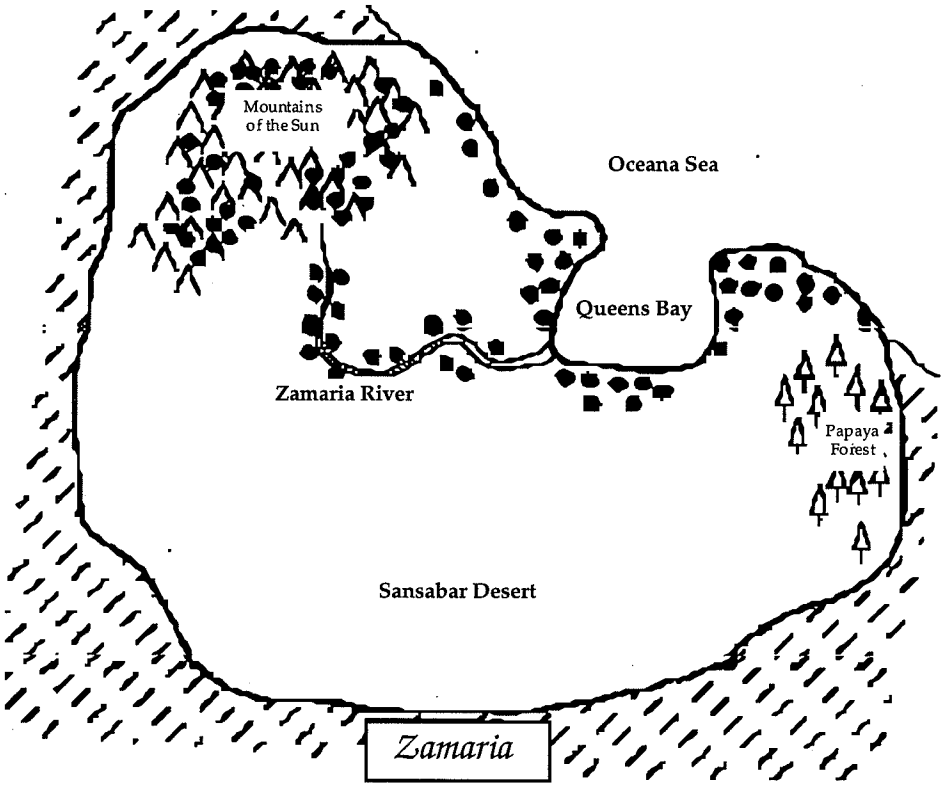
ANALYSIS

This student provides an unusual analysis of the situation. Seven facts are included in the argument, however, the facts focus only on the choices for a denser population rather than addressing reasons for not choosing other areas. The facts that address the forest are the available "forestry and water." The river offers water, which was not counted again as it was included in with forests, "lush forests [fertile land], places to build irrigation for farms, natural resources," and "boat landings," which seems to address shipping and transportation. Had the student been more specific about how the argument was laid out, the Q-score would have been much improved.

Qualitative Score	Facts Correct	Facts Incorrect
3	**7**	**0**

LOW SCORES

ROBERTA



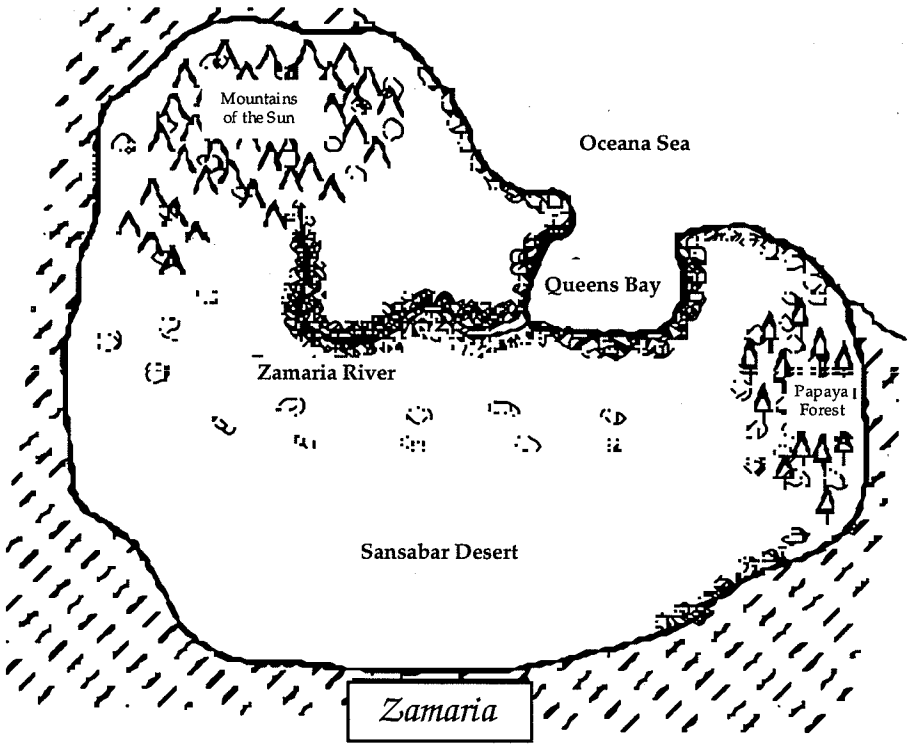
I think *thail Live in plases by water river's and mountions.* The mountions bring water. too. The water can help the paple to survive. people ned food and water to survive.

ANALYSIS

In this example, the student's statement of choice is clear, "thail Live in plases by water river's and mountions." There is one solid fact in support of this statement: "people ned food and water to survive." There is a statement about water in the mountains that is unclear, and therefore not scored. Comparing the statements to the map, the student put in a great deal of time placing dots on the map. The placement of the dots shows a general understanding of the need for resources and water, or perhaps it is just that the student recognized the lack of resources and water in the desert. In either case, the task seemed a bit unclear to the student.

Qualitative Score	Facts Correct	Facts Incorrect
2	**1**	**0**

AMY



I completed the map the way I did because. The people that settle in those places because *the Zamaria river is by a big population* of people. The Papaya Forest is by a small population of people. And the Sansabar desert is by a small population. The mountain will affect the big population. And the Desert and the Forest will affect the smaller populations of people.

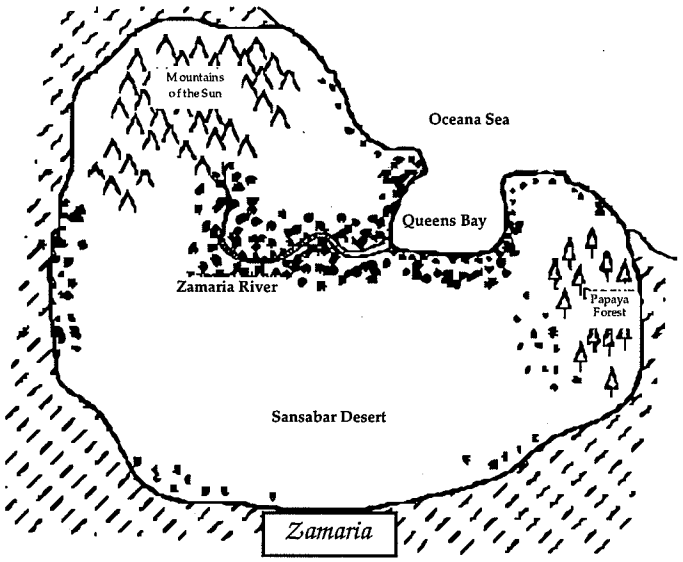
ANALYSIS

The student offers "the Zamaria river is by a big population" as a statement of choice. Though the student addressed several landforms, there were no facts offered in support of the statements of choice. In initially scoring this essay, there was discrepancy between raters and a Q-score of "0" was given. The issue was whether the dot placment on the map was random, or if there was reason. Upon further examination, the answer received a Q-score of "1."

Qualitative Score	Facts Correct	Facts Incorrect
1	**0**	**0**

AN ANOMALY

NOEL



> Near water is the first thing *people want to live near water.* Some people might live in the mountains some in the Forest There might not be any lady in the desert.

ANALYSIS

The scorers had a difficult time with this response. While the student suggests "people want to live near water" as the statement of choice, some scored it as having facts and others did not. The qualitative score is "2" which implies that a fact is needed, otherwise the Q-score would be only "1." Even more curious, upon inspection of the map, anomalous dots appear around the border of Zamaria. This was not the only map that sported such dots. This may indicate confusion as to what the hash-marked area indicates - whether it is ocean or land. This student, as well as others, assumed that Zamaria was an island. It is obvious from the density of the dots elsewhere that the river was highly favored population-wise, therefore some scorers were giving the student the benefit of the doubt by awarding a point for the use of water as a fact. The important thing to note here is that the map, designed to be a help, to some students may have been a distraction.

Qualitative Score	Facts Correct	Facts Incorrect
2	**1**	**0**

CHECK-OUT OF EVALUATION: SIXTH GRADE SCIENCE

In scoring these essays, incorrect facts should not be scored. They should, however, be considered for logical argument purposes in the qualitative score. Count analogous forms as one fact.

HALLIE

I think they should leave them alone because It is going to have ~~[illegible]~~ pollution all over. It is going to destroy these air. They well have cokes and then they can't work. Now if we want to have it then they should let us are land is all most all destroyed. There land has not been hurt yet. It will hurt are land more but people think they need it.

What is the statement of choice?_____

What facts are presented to support their choice?

_____

_____

What facts are presented to insubstantiate other possible choices?

_____

_____

_____

Qualitative Score	Facts Correct	Facts Incorrect

SANTANA

I think the people of Nautopia should not use the fossil fuels on thes planet because it would polute their planet.

It could cause harm to the miners and the people by the mines and the explosions.

It could ruine their land by having the fossil fuels dug. Many of the people would not be ablev to see the land when they are not busy. People could become lazy.

The smog and polution could make ther lungs blach and they could die many diseases are cdused by fossil fuels like blach lung.

It could kill some of the animals and crops so no one would have any food. It would polute water and there would not be any clean water to drink or swim in. The water could get polutede by oil spills getting fossil fuels from place to place

What is the statement of choice?_____

_____

_____

What facts are presented to support their choice?

_____

_____

_____

_____

What facts are presented to insubstantiate other possible choices?

_____

_____

_____

_____

_____

_____

_____

_____

Qualitative Score	Facts Correct	Facts Incorrect

RICKIE

I think that the people of Newtopia should leave the fossil fules alone, keep things the way they are. Because fossil fules are bad for the environment to use them it causes polution. I think that the people that don't want to use them are very smart because they don't know what is going to happen to their planet if they use the fossil fules. I also think that people would get bord of just sitting around and doing nothing. Even if they used fossil fules to do their work they would get lazy and when the fossil fules run out they won't want to do the work to heat their homes and cook their food.

What is the statement of choice?_____

_____

_____

What facts are presented to support their choice?

_____

_____

_____

_____

What facts are presented to insubstantiate other possible choices?

_____

_____

_____

_____

_____

_____

_____

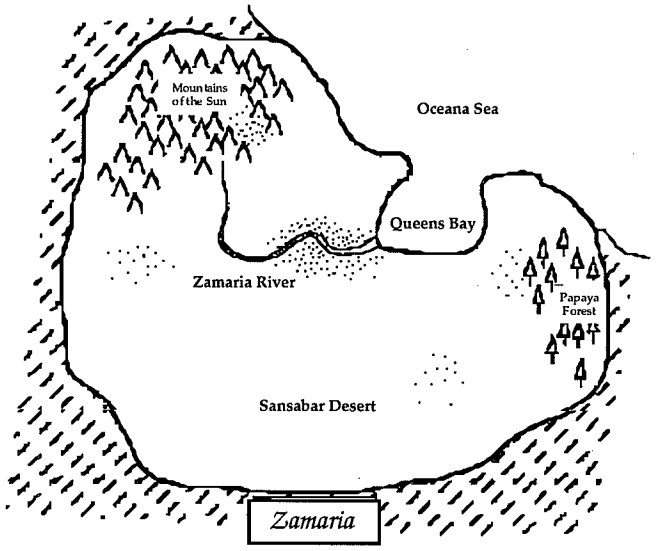
_____

Qualitative Score	Facts Correct	Facts Incorrect

CHECK-OUT OF EVALUATION: SEVENTH GRADE SOCIAL STUDIES

In scoring these essays, incorrect facts should be scored and considered for logical argument purposes in the qualitative score. Count analogous forms as multiple facts.

KELLY



I think more people would want 2 live by the water, because they can use it every day & not have to walk a long way. I don't think many people would want to live in the desert or mountains or forest, cause these would not be much water and would have to walk a long way to get to a river.

What is the statement of choice?_____

_____

What facts are presented to support their choice?

_____

_____

_____

_____

What facts are presented to insubstantiate other possible choices?

_____

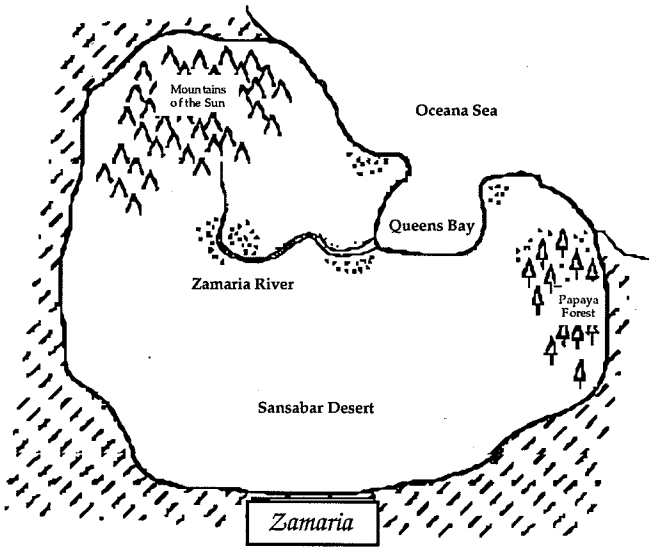
_____

_____

_____

Qualitative Score	Facts Correct	Facts Incorrect

SANDRA



The way I did this is because. The most people live near the river becouse. They can ship out goods but not be flooded by a storm or some then.

What is the statement of choice?_____

_____

What facts are presented to support their choice?

_____

_____

_____

_____

What facts are presented to insubstantiate other possible choices?

_____

_____

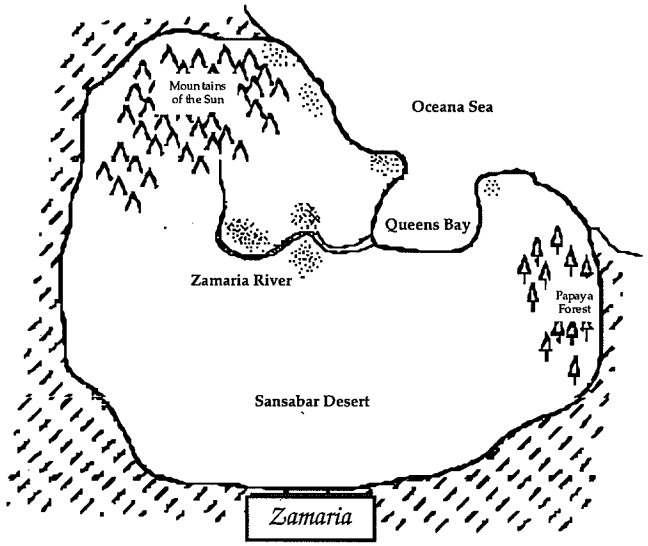
_____

_____

_____

Qualitative Score	Facts Correct	Facts Incorrect

DERRICK



I Pick The Places Because People won't survive As GOOD it The DesserT Because The soil is Too Dry AND THey WOULDNT Be Able To Grow cRops so THey will HAve To move ArouND To GATHer Food BUT iF THeyre By THe Sea or rivir THey'lleBe Able To irrigATe And Grow Food

What is the statement of choice?_____

_____

What facts are presented to support their choice?

_____

_____

_____

_____

What facts are presented to insubstantiate other possible choices?

_____

_____

_____

_____

Qualitative Score	Facts Correct	Facts Incorrect

ANALYSES OF CHECK-OUTS: SIXTH GRADE SCIENCE

ANALYSIS OF HALLIE'S ESSAY

The statement of choice is: "I think they should leave them alone." The student offers a single fact:

- It is going to have pollution all over.

There is only the one fact in support of the student's stance: fossil fuels pollute. There is a second statement mentioning air pollution that is analogous to the first. The student then turns to ungrounded opinion and becomes quite creative in suggesting that the people of Newtopia send Earth their fossil fuels because our planet is already polluted. The Q-score is "2," a statement of choice supported by one fact.

Qualitative Score	Facts Correct	Facts Incorrect
2	**1**	**0**

ANALYSIS OF SANTANA'S ESSAY

The statement of choice is: "I think the people of Newtopia should not use the fossil fuels."

In support of this statement are these facts:

- it would polute their planet.
- It could cause harm to the miners...Many diseases are caused by fossil fuels like black lung.
- People could become lazy.

The first fact is again straightforward. The second fact is actually two statements that are parallel and are therefore treated as a single fact. The third fact is a complicated statement that is, in reality, a misstatement of the fossil fuel capacity to do work. It is interesting to note that this statement may be a product of a discussion that took place in that classroom. Several of the students in the class, including the next student we will look at, Rickie, mentioned laziness as well. This could represent mislearning or students not fully understanding the definition of work used in physics. There are three facts that are presented and they are all in support of a single side of the argument, therefore the Q-score is "3." If you scored the third fact as incorrect, it is understandable but the Q-score is still "3" either way.

Qualitative Score	Facts Correct	Facts Incorrect
3	**3**	**0**

ANALYSIS OF RICKIE'S ESSAY

The statement of choice is: "I think that the people of Newtopia should leave the fossil fules alone."

In support of this statement are these facts:

- to use them it causes polution.
- fossil fules run out.
- if they used fossil fules to do their work they would get lazy and...won't want to do the work to heat their homes and cook their food.

The first fact is straightforward: the use of fossil fuels causes pollution. The second fact, that fossil fuels are non-renewable, is also apparent. The third fact, however, is convoluted and easy to miss. Its essence is that fossil fuels make life easier because of the capacity to do work such as heat homes. Because of its recognition of the opposite argument, (fossil fuels would help by greater work capacity), the student received a higher Q-score. Facts 1 and 2 argue the student's side of the argument, and fact 3 recognizes the other side of the argument: multiple facts on one side and a single fact on the other.

Qualitative Score	Facts Correct	Facts Incorrect
4	**3**	**0**

ANALYSES OF CHECK-OUTS: SEVENTH GRADE SOCIAL STUDIES

ANALYSIS OF KELLY'S ESSAY

The statement of choice is: "I think more people would want 2 live by the water." In support of this statement are these facts:

- they can use it [the water] every day & not have to walk a long way.
- these would not be much water and would have to walk a long way to get to a river.

The first fact refers to the need for water and the convenience and availability of water near the river. The second fact describes the relative lack of water and the need to travel farther to get water in other types of landforms. Had the student addressed each landform separately, they would have been considered separate facts, but because the desert, the mountains, and the forest were addressed as one, only one fact is scored. This may also be indicative of either the student's lack of understanding about maps (refer to accompanying map), or mislearning that has taken place in the class, for ample water may exist in the mountains or in the forest.

Qualitative Score	Facts Correct	Facts Incorrect
3	**2**	**0**

ANALYSIS OF SANDRA'S ESSAY

The statement of choice is: "The most people live near the river." In support of this statement, the student offers:

- They can ship out goods.

The student clearly understands the value of shipping goods via waterways. Unfortunately, the rest of the answer is vague. Perhaps the student was referring to the sheltered river and bay as being safer from storms at sea though the idea of flooding seems more appropriate to the river. Because of this lack of clarity there is really nothing that can be done with this portion of the answer. Therefore, there is one fact in support of a clear statement of choice and the Q-score is "2."

Qualitative Score	Facts Correct	Facts Incorrect
2	**1**	**0**

ANALYSIS OF DERRICK'S ESSAY

The statement of choice is: "I Pick The Places" (desert, river, and sea). In support of this statement are these facts:

- The soil is Too Dry.
- THey WOULDNT Be Able To Grow cRops.
- THey will HAve To move ArouND To GATHer Food.
- THey'lleBe Able To irrigATe And Grow Food.

Upon examination of the map it is clear that the student understands the value of land near the water and other natural resource areas. The student demonstrates understanding of the concepts by describing the desert as dry and having poor soil for growing crops. In addition, the student mentions nomadic life as the people of the desert must travel to find food. Finally, the student mentions irrigation by the sea or by the river and growing food. There are three facts that support the idea that the desert is a difficult place for people to live and one in support of living near the river. Therefore the Q-score is "4", multiple facts in support of one side of the statement of choice and one in support of the other.

Qualitative Score	Facts Correct	Facts Incorrect
4	**4**	**0**

Explanation Scoring

Section B

ANALYSIS OF EXPLANATION

SCORING EXPLANATION ANSWERS

The following student answers have been examined and scored using the flowchart scoring system (Figure 1). The critical features addressed in explanation scoring are: (a) the link between the answer and the prompt, (b) the number of events mentioned by the student, (c) the sequencing of mentioned events, and (d) the accuracy of the events listed.

The first level of questioning in the scoring system is whether or not the student's answer is an attempt to respond to the prompt. An answer such as "The moon is big," would not address a question about the American Revolution. Other answers of this nature would be "I don't know," or answers that had been left blank. Such answers are assigned a score of "0."

LINK

The link is a demonstration of the student's understanding of the prompt itself. It is a specific and topical reaction to the prompt that focuses and frames the student's answer and relates it to the prompt. It should be explicit in nature. Many teachers ask their students to include a portion of the prompt in their answers. This reiteration can be used to focus the student on the question being asked, to create a more coherent answer, and to force the student to begin writing. The link is usually given at the beginning of an answer as an introduction to the student's thoughts, however, the student may give a delayed link that occurs in the middle or end of the answer. For example, in response to a question that requires the student to mention the four forces influencing flight, the student may provide the link, "The four forces influencing flight are..." at the beginning of their response. "Drag and gravity are two forces that must be overcome in order to fly. This is accomplished by the two other forces that influence flight..." is an example of a middle link, while "...so those are the four forces influencing flight," is an example of an end link.

An answer that is mere reiteration is considered to at least be linked for scoring purposes.

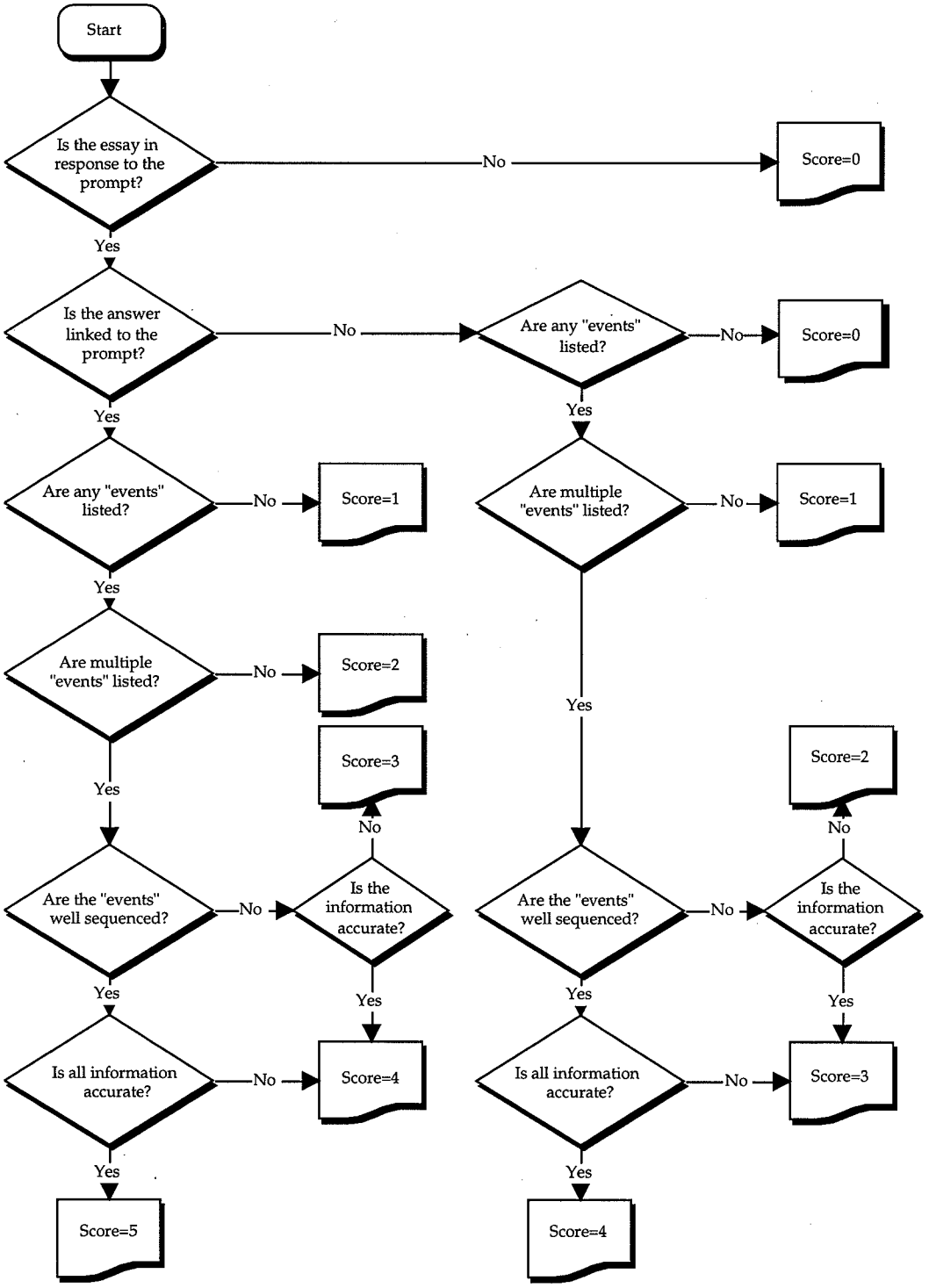


Figure 1. Explanation flow chart.

EVENTS

Events are the individual steps, points of argument, or ideas students mention in the process of answering a prompt. They should be listed on the content planning worksheet as the concepts, key words, attributes, examples, logical steps, or ideas that have been included in the class curriculum. The events are offered by the students as evidence for their position. This evidence becomes the scaffold of their argument or explanation. Acceptable events should be listed on the concept sheets by the teacher before scoring begins. Occasionally, however, students may offer correct ideas or concepts of their own that were not listed on the worksheet. These ideas should be included as correct events. Incorrect events should not be scored although they should be monitored, for such incorrect notions may show where mislearning has occurred in the classroom.

When multiple scorers are working on a series of student answers, possible deviations from concepts given on the worksheet should be discussed before scoring. Scorers should decide whether deviations from the worksheet will be accepted or rejected, or whether scorers will use their best judgment on a case-by-case basis. Scorers should also decide whether redundant events are scored as singular parallel ideas, or as re-occuring events. For example, they must decide whether the statements, "Deserts have little food," and "Forests have much game for food," are scored as synonymous rephrasing of a single event (food = one point), or as separate events (food + game = two points). With prompts that give a student a wide range of possibilities for their answers, the former scoring would be preferred. In other cases, the prompt may be narrow in focus so the latter scoring is necessary. Such scoring nuances must be discussed prior to scoring for the sake of inter-rater reliability, although some students will still surprise scorers with an obscure or highly unusual answer.

SEQUENCING

Sequencing is the listing of events by the student in a careful, step-by-step manner so all events, information, and their relationship to the prompt are clearly defined. Sequencing may also be viewed as the linking of one event to another in a careful chain of steps that is coherent to the reader.

For scoring purposes, the sequence is viewed as a logical progression from one event to the next. For example, in answer to a prompt asking students to explain the process of convection, a student may say, "Convection is the process that occurs when a fluid is heated. Heated fluid rises up, pushing the heavier cool fluid out of the way. The cool fluid sinks down and is heated so it rises." This answer clearly relates the process of convection and each event naturally progresses from one step to the next. In contrast the student who writes, "The fluid heats up, rises, and then drops down," has given an answer that is poorly

sequenced, for the step-by-step process of heating and cooling is not presented.

ACCURACY

Often students will present events and ideas which are clearly in error or are not related to the prompt. As mentioned previously, these non-events should not be scored but monitored by the teacher to prevent mislearning. There are cases, however, in which the events may be questionable as to their importance to the final outcome mentioned by the student. For example, in answer to the prompt on convection, a student may say that the glass beaker that contains the fluid is cold. Though possibly true at some point in an experiment, it is not pertinent to convection except as an initial condition and it has little bearing on convection outside of the experiment. Therefore, the statement is accurate, but not important to the prompt, and consequently not scored as an event.

Sometimes students will attempt to build arguments on opinions. For example, in response to a prompt asking students to explain why a small country would adopt a democratic form of government, a student may offer, "Because it would make them happy." The student fails to present concrete events that would have led to the prompted outcome. Because these statements reflect opinion or belief systems that are not verifiable, they should be monitored to prevent mislearning, but they should not be scored as events.

CONCLUSION

When scoring explanation answers, the link between the answer and the prompt, the number of events mentioned, the sequencing of those events, and their accuracy, are used to examine the quality of the answers. It is important to list concepts, key vocabulary, critical features, and facts on the content planning worksheet *before* scoring so focus can be maintained during the scoring process. When multiple scorers are involved in the task, several decisions must be made prior to scoring to maintain reliability and make scoring easier (i.e., what to do with deviations from the content planning worksheet and redundant events?).

TRAINING PROMPTS

SCIENCE

The science prompt used in this section was only one of several questions that appeared in the assessment (Figure 2). The prompt read, "What happens at night to make the top of your car wet when we go outside in the morning?" Because of the nature of this assessment question and its placement in the instrument, it is necessary to refer to the preceding question; "Explain how temperature might affect the

relative humidity of an environment." This is necessary because the two questions are inextricably intertwined. Many students addressing the relative humidity question answered it fully, but then they answered the car prompt without elaborating their answers. Therefore, both questions must be considered for the purposes of this module. The content of the preceding question was used as a check for understanding only. The relative humidity question was scored on its own merit and it is used as a reference point for clarification in the explanation problem presented in the "car" prompt.

SCIENCE PROMPT

18. What happens at night to make the top of your car wet when we go outside in the morning?

Figure 2. Science essay prompt.

The concept being assessed is dew point. The car is cooler than the surrounding air and the cool surface of the car causes the air immediately surrounding it to cool. This means that the water vapor in the air reaches dew point. As dew point is reached, the humidity must decrease so it condenses on the cold surface, or forms dew. The list of events below (Figure 3) was created as a guide for scoring student responses to the science prompt. It is by no means complete, for answers that were analogous to these ideas were also acceptable.

RESPONSES ON THE DEW POINT SCIENCE PROMPT

Events listed among the answers offered by the students were:

- Change in dew point
- Relative humidity
- Condensation
- Water vapor turning to liquid
- The cool surface of the car
- Lower temperature

*Also acceptable were analogous statements such as, "The air near the car cools."

Figure 3. Events used in answering the science prompt that were acceptable for scoring.

In the following analyses, each answer is presented as it appeared in the student's answer. Following the student answer is the scorer's response to that answer and an analysis, with the flowchart path italicized. The assigned score follows.

HIGH SCORES

MIKE

> If there was water on the car condensation accoured. Water vapor in the air around the car was cooled to its dew point and condensed on the car to make dew on the car.

ANALYSIS

This answer is a *well-linked answer of the prompt.* The student mentions *several facts* that point out some specific features and vocabulary of the concept. The events are *clearly linked* and *accurate.* The previous answer is fully correct, so the understanding of dew point demonstrated there is clearly carried over into this question. There is a solid link between temperature, dew point, condensation, and relative humidity.

Score
5

KRISTIN

> It wet because the moisture. The car is cold so It will cool the air around It. The processes is gas to liquid so the car gits wet

ANALYSIS

The answer is an *attempt to address the prompt.* It is *clearly linked* to the prompt through addressing the illustration (the car). There are *two events listed*, the car cooling the air immediately around it, and the process of gas changing to liquid. The *link between the two is weak*, however, *degrading the sequence* of the events. The events scored are *accurate*, and though unnecessary to do, upon referencing the preceding answer, it is evident that the student understood the concepts.

Score
4

OLAF

> When the top of your car is cold and the air is warm water vapor in the air will condense on the cool surface of the car.

ANALYSIS

This answer is a solid *attempt to answer the prompt* and it is *well linked* and grounded to the prompt. *The sequence of events present in this answer is accurate*, though because of the lack of attention to the technical aspects of the dew point, the *sequence suffers.* Attention to the preceding question reveals that the student did not fully understand the concept. The student claimed that a lower temperature would mean a lower relative humidity. However, the concept of condensation is understood.

Score
4

LAURA

> As the air gets cooler it gets more relitive humidity so after it gets saturated so of the air will condense on the top of your car.

ANALYSIS

The answer is a *well-linked attempt to answer the prompt.* The *events and evidence are accurate.* There is attention to some of the

vocabulary and it is used correctly. The concept of saturation also enhances the answer. However, clear presentation of why condensation occurs on the car is not presented, so the *link between suffers* slightly. Referring to Laura's previous question, it is apparent she didn't fully understand the concept for it is stated that, "on hotter days the relative humidity would go up." No link between temperature and condensation is presented.

Score
4

GEORGE

> Condensation makes the top of your car wet because the air reaches it's dew point, and dew has to have a surface to condense on. That makes the top of the car wet.

ANALYSIS

The answer is a *well-linked attempt to answer the prompt.* The *sequence of events is clear and accurate.* The events given are *accurate* and the vocabulary used is correct. However, the answer does not explain why the dew condenses on the car, only that it condenses. Again, temperature is not mentioned. Upon examination of the previous answer it is not clear that there is a connection between temperature, dew point, and condensation. Because of this flaw the answer suffers.

Score
4

MIDDLE SCORES

CARIE

> the air reaches it's dew point so it creates Dew.

ANALYSIS

This answer is an *attempt to address the prompt.* The student mentions the process of the air reaching its dew point and creating dew. The answer is *negligibly linked* in this prompt, and *two events are listed*, the air reaching dew point and dew being created. The few events that are listed are well sequenced but a large gap exists between the prompt and the answer. The *accuracy is unquestionable,*

for upon examination of the previous answer it can be seen that the student did understand the concept. It is stated that the hotter it gets, the lower the relative humidity and that when the air cools, the relative humidity gets higher.

Score
3

MITCH

> Dew the air is satrated so it makes moisture on the ground and all around.

ANALYSIS

The answer is an *attempt to address the prompt.* It is *not clearly linked* to the prompt by addressing the question directly. The student refers only to "the ground and all around." *Multiple events are listed;* saturation and making moisture on the ground. The events are *not well sequenced,* although they are *accurate.* By referring to the student's previous answer it is apparent that the principle is clearly understood, but poorly relayed in this answer.

Score
3

KAMALA

> The air cools off and dew formes on the car.

ANALYSIS

This *answer is an attempt to address the prompt.* Although the student mentions the process of the air cooling and dew forming, there is *no link* between the *two events listed.* Consequently, the *sequence suffers.* The student could have mentioned the attainment of dew point for a higher score. Examination of the preceding answer is unnecessary, for it was left unanswered. As to *accuracy, a question* remains whether the student understood the concept.

Score
2

LOW SCORES

NADAR

> The dew from the cold.

ANALYSIS

This answer is an *attempt to address the prompt*, however, it is not *clearly linked* to the answer. This makes it difficult to see whether the student understood the process of condensation clearly. There is only *one event listed.* Examining the preceding question does not show knowledge beyond the prompt.

Score
1

SOCIAL STUDIES

The second prompt (Figure 4) to be examined is a Social Studies explanation question in which students are asked to explore reasons why the American colonies changed from three distinct regions into a more united country. Scoring of this prompt was particularly difficult and the initial inter-rater reliability was quite low. The problem was the prompt itself: there was discrepancy in the wording that created confusion among the students. The prompt reads:

> SOCIAL STUDIES PROMPT
>
> Your grandpa grew up when the colonies were divided into three regions. You are a 14 year-old on vacation. When you travel up and down the east coast on vacation you don't see the differences that your grandfather had talked about. You come across the term "American" wherever you go, but you remember the stories from your grandpa about the three separate colonial regions. What happened to get people to change? In a 10 sentence paragraph, explain the most important reason for this change.

Figure 4. Prompt for social studies explanation essays.

The wording of the prompt was not sufficiently open-ended. The intent of the question was to force the students to write in depth about several factors. They were first asked to consider the factors that led to the change in people, and then about a single factor. Some students wrote general statements mentioning several events that were disjointed and contained little specific information. Other students concentrated on single events, bringing few details into their essays that could be used to build a solid explanation case. Because of the confusion from this prompt, answers primarily scored either high or low. There were few middle scores.

There were few events offered by the students as explanation for the unification. Most appear in Figure 5.

RESPONSES ON THE SOCIAL STUDIES PROMPT ABOUT AMERICA

Events listed among the answers offered by the students were:

- The Great Awakening
- Broader travel
- A reduced Native American population
- The American Revolution
- Increased trade

*Also acceptable were other analogous statements such as, "religious movements."

Figure 5. Responses found to be acceptable for social studies essays.

As with the science explanation answers, each of the following answers is examined using the scoring flowchart for critical features. Addressed in explanation scoring are: (a) the link between the answer and the prompt, (b) the number of events the student mentions in the sequence, (c) the linkage of events, and (d) the accuracy of the events listed.

In the following analyses, each answer is presented as it appeared in the student's answer. Following the student answer is the scorer's response and analysis of that answer with the outline of the flowchart path italicized. The score assigned follows each essay.

HIGH SCORES

PAUL

> The most important reason of this change was people. People began to change because of the Great Awakening, and the Revolution. As different people moved into different areas, they changed & adjusted to this. Different cultures in areas affected the way people there thought and did things. People began to accept these, making the Colonies unite.

ANALYSIS

The answer is *linked* in the first line. There are *several reasons for the change listed* within this answer. First, the student suggests that the Great Awakening and the Revolution united the people. The student then goes on to elaborate by saying that a mixing of cultures in different areas changed the way people thought. The student carefully examined several factors that fall under the same theme of people. While the *events are accurate*, the *sequencing is quite poor*. Had the student listed the events in order, the score would have been better.

Score
4

CAROL

> People traveled. They traded in other colonies. That meant they had to travel. Soon individual colonies would unite. Slowly they did. The "Great Awakining" also helped this. People wanting start new religeons moved together and made bigger towns. After a while they did unite and they were big towns. The Indians also contributed to this. People in big groups could fend off the Indians.

ANALYSIS

This student *responds to the prompt* and offers trade and the Great Awakening as *accurate events* for change. Also suggested is that religious beliefs and the threat of attack brought people together into communities. These ideas are *not explicitly linked* to the prompt. The *sequence of the information is visible* in the way the student gave time references such as "After a while."

Score
4

DENISE

> The people changed because England no longer owned them. They made their own laws. The colonies became states. They became United as one. The three different groupings came together. England no longer owned them because they had a war. The americans won. They were free and the Englanders no longer controlled them. They made a flag with 13 stripes and 13 stars. Once they were free they decided to only add a star for each state after that. .

ANALYSIS

This student *linked the answer to the prompt* with the first four words, "The people changed because..." This link creates a strong base for the remainder of the answer. The reasons offered by the student are related to the revolution: (a) colonies became states and united against England, (b) people made their own laws, and (c) reduced English control and ownership. While the accuracy of the second event is questionable, the first and third are considered *accurate.* The *sequencing is weak* because the student didn't make it clear whether the colonies united before or after the war.

Score
4

MIDDLE SCORES

GEORGE

> Cities were getting bigger and there were more farming. Because there keeping getting bigger.

ANALYSIS

There is *no overt link* between the prompt and this answer, however, there are *two events listed.* The *sequencing of the answer is vague* unless the second sentence is considered a link. The main ideas, that the cities were growing and that more farming was taking place, are *correct,* however, they are *not well sequenced.*

Score
3

ROBERT

> All the indains died off so, villages started spread out. Eventually those villages grew into cities. Now you can't tell where the 3 colonies were.

ANALYSIS

The answer is *linked to the prompt* by the last statement. There is, however, only *one accurate reason offered*; the villages grew into cities. Had the student offered more reasons the score would have improved considerably.

Score
2

TRAVIS

> The most important reason for this change is that the people learned that they had to unite. To become a free country, not just a couple little could triumph over a rather large country at the time. So the people started wising up and the became one army together.

ANALYSIS

This *answer is linked to the prompt.* This answer is at first difficult to score, however, upon closer examination it is evident that the student was trying to say that a couple of small countries could not have defeated the British, so the colonies united to become one army to fight for their common good. Only *one accurate event is listed* in this answer and it is difficult to pull out. Had the student addressed more reasons for the change, the score would have improved. Perhaps the problem stems from the prompt, as was discussed in the introduction.

Score
2

LOW SCORES

JOE

> The people thought that they were living to simple. They needed to make some progress to open up their villages to more new comers. When to many people started to come they decided to build little cities. After a while, the people got divided, causing them to form different colonies of their own. The people then started to live in small groups. When that happened, people had to fend for themselves. They were still a colony, but they were a lot smaller. Then, after a while, they decided to break up and live on their own

ANALYSIS

This student *attempted to answer the prompt*, but the answer becomes inaccurate. The prompt is *weakly linked* because the trend described in the response is one of breakup, not unification. This answer creates a problem: Do inaccurate events get scored? Do you score the questionable link? Certainly it is accurate to say that as people came to the colonies, the cities grew, but how did that affect the unification of the colonies? In taking this answer step-by-step through the scoring system, it is possible to say that *there is a link*, and that there is at least *a single event listed.*

Score
1

PATRICIA

> People changed because of all the work that they had people doing.

ANALYSIS

This answer is an *attempt to answer the prompt* and is *linked.* However, the *single reason* given is not clear and should not be scored as an event.

Score
1

MARK

> because the inenterd sevants came and they had they own religen and then came the slaves and they have there religen and the americans had ame religen.

ANALYSIS

An *unsuccessful attempt is made to link this answer to the prompt.* There is an attempt to give events that created unification, however, *nothing concrete* is offered. Therefore, the answer is not linked and no events can be scored.

Score
0

CHECK-OUT OF EXPLANATION: SCIENCE

MIKE

> During the night, your car is warm from all the driving you did in the afternoon, the cool night air making dew form on the extrior of your car.

What is the link?_____

What events are presented to support their explanation?

_____

_____

_____

Is the explanation carefully sequenced? _____

Is all of the information accurate? _____

Score

LINDA

> The air cools down and Dose condintion by turning The air around, it ~~condenses~~ cooler TURNING IT TO Gas TO liquide and The car wet (Dew point)

What is the link?_____

What events are presented to support their explanation?

Is the explanation carefully sequenced? _____

Is all of the information accurate? _____

Score

LOREN

The moisture from the air connects with the ~~the~~ temperature of the car which is condensation and ~~h~~ er the car ends up on the roof of ~~t~~ the car

What is the link?_____

What events are presented to support their explanation?

Is the explanation carefully sequenced? _____

Is all of the information accurate? _____

Score

CHECK-OUT OF EXPLANATION: SOCIAL STUDIES

PHILLIP

one of the reasons why it changed is because of the great awakening, and another is because they wanted to become one big region. I don't know

What is the link?_____

_____

_____

What events are presented to support their explanation?

_____

_____

_____

_____

_____

_____

_____

_____

Is the explanation carefully sequenced? _____

Is all of the information accurate? _____

Score

BETH

> Im traveling along the coast. I dont see any differences My grandfather told me there would be changes. I see the word "American" over and over, I remeber my grandfather toll me about 3 different colonies. I wonder what happened to make the people change. I get tired of writing this! I fall asleep. The bell rings. I wake up.

What is the link?_____

What events are presented to support their explanation?

Is the explanation carefully sequenced? _____

Is all of the information accurate? _____

Score

CAROLINE

Travel was what made the change. People visiting different colonies and started sharing cultures. The Great Awakening helped people start sharing their religious beliefs with one another. Due to travel people shared the beliefs with people in all the colonies. Everyone started to move around and everyone with different beliefs spread out and everyone came together.

What is the link?_____

What events are presented to support their explanation?

Is the explanation carefully sequenced? _____

Is all of the information accurate? _____

Score

ANALYSES OF SCIENCE CHECK-OUTS

MIKE

The answer is an *attempt to address the prompt.* It is *linked in a sequence that is carefully drawn.* The student did understand that dew is what makes the car wet, however, it is questionable whether the student clearly understands the concept. It is falsely suggested that the warmth of the car and the cool night air makes dew form on the car. Therefore, *only one event is scored as accurate.* The student is not clear on how the concept works though he knows it has something to do with the temperature. Examination of the preceding answer shed no light on this one, so nothing could be added to the argument.

Score
2

LINDA

The answer is an *attempt to address the prompt.* It is *clearly linked* to the prompt as can be seen by the reference to the prompt. Due to spellings and grammar, it is difficult to tell whether the student fully understands. If the phrase is taken to say "The air cools down and does condensation by turning the air around it cooler, turning the gas to liquid and the car gets wet," it could mean that the student understands that as the temperature drops, dew forms. There are at least *three conceptually important items* contained in the question: as water vapor cools, it turns to liquid, it is called condensation (misspelled), and it has something to do with dew point. The latter point is too vague to award a point, but the former points are *accurate.* The answer is fairly well *sequenced.* Reference to the preceding question does not help. Had dew point been addressed more directly, this would have scored "5."

Score
4

LOREN

The answer is an *attempt to address the prompt.* It is *clearly linked* to the prompt through addressing the illustration (the car). The answer is difficult to read because the student weakly defined condensation by suggesting that the moisture from the air connects to the temperature of the car. There are *two events suggested*; moisture is in the air and it condenses. *The sequence is considered poor* because it is unclear whether the student understood the concept. The *accuracy of the answer remains questionable.*

Score
3

ANALYSES OF SOCIAL STUDIES CHECK-OUTS

PHILLIP

This student *linked the answer to the prompt.* The *reasons given* for the changes are the Great Awakening and their desire to become a larger body or single region. The sequence is unclear. The second reason is unclear, therefore its *accuracy is questionable.*

Score
3

BETH

While this student created an *elaborate link* to the prompt, upon closer examination of the answer it is apparent that nearly every word is a reiteration of the prompt. There is no expansion upon this information and no events are mentioned. The student suggests that there is a new movement to unify America: "I see the word 'American' over and over." This, too, is a reiteration of the prompt and is not clarified, therefore the *accuracy is questionable* and it is not scored as an event.

Score
1

CAROLINE

Several well-linked events are listed in this *response to the prompt.* First, it is suggested that travel was a major cause of the change. This links the answer and offers one event. Mixing of cultures is also a factor. The Great Awakening brought unification in religious causes and the spread of evangelism. Finally, the student suggests that the spreading of these beliefs helped to unite the country. *The sequencing of the events is strong,* with indication that the Great Awakening was a process of religious belief growing and spreading. The *events listed are accurate,* and as a result, this answer scored high.

Score
5

Prediction Scoring

Section C

ANALYSIS OF PREDICTION

SCORING PREDICTION ESSAYS

Each of the following answers has been examined using the flowchart scoring system appearing in Figure 1. The structures of prediction prompts and essays are quite similar to the structures of explanation prompts and essays. Addressed in prediction scoring are: (a) the presence of a prediction, (b) whether the prediction is related to the prompt, (c) whether the prediction is linked causally to the prompt, (d) whether the outcome can be logically derived from the prompted situation, (e) the number of causal steps or events that have been listed, and (f) whether all causal steps in the chain of events are clearly sequenced from prompt to predicted outcome.

ADDRESSING THE PROMPT

Predictions are inextricably related to notions of cause and effect. The student is asked to tell what will happen (effect) given an initial condition (cause). The first level of prediction scoring is to ask whether the student offered a prediction in response to the prompt. An answer such as "The moon could blow up," clearly does not address a prediction prompt about the American Revolution. Other answers of this nature would be "I don't know," or answers left blank.

RELATIONSHIP TO THE PROMPT

The student's answer must be examined for a connection to the prompt. Usually the relationship to the prompt is clear and easy for the scorer to see. Nevertheless, sometimes students will offer a final outcome with few details or events between the prompted situation and outcome. This may appear as a cryptic answer that seems to have little bearing on the prompt.

For example, if a student is asked to predict the results of heavy pollution in the air over the Antarctic, the student may respond "Farmers in Nebraska will go out of business." At first glance the answer appears to have little bearing on the prompt. The student may have taken a leap in their thinking from the prompt to an outcome without attention to the bridging details. For example, weather changes could cause crop failures. There are multiple steps of logical development missing from the answer so it appears to be unrelated to the prompt; however, in an obscure manner it does relate.

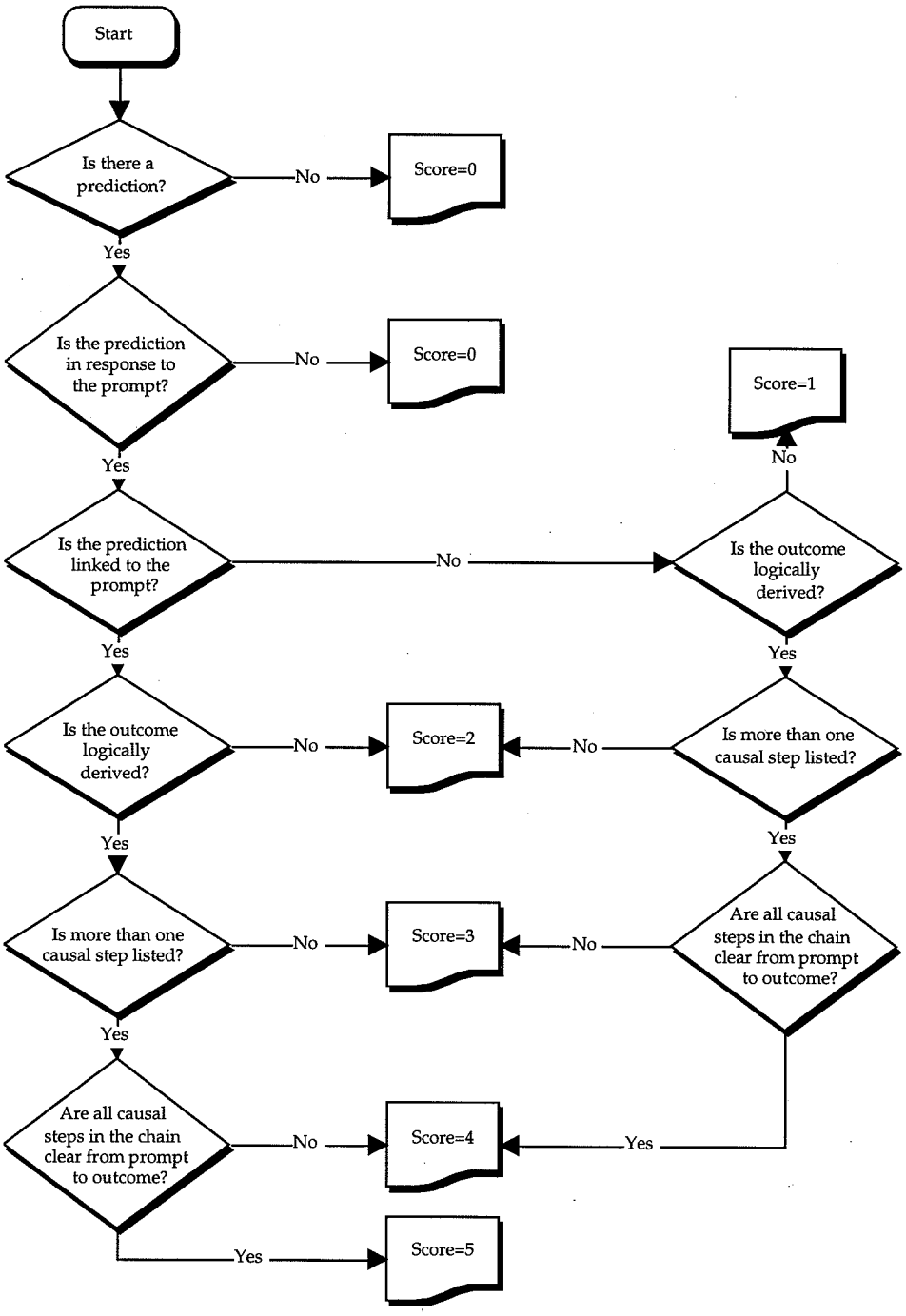


Figure 1. Prediction scoring flow chart.

LINK

As in explanation essays, the link is a demonstration of the student's understanding of the prompt itself. It is usually given at the beginning of an answer as an introduction to the student's thoughts. Students may give a delayed link that can occur in the middle or even at the end of their answer. For example, students may give the link, "I predict that..." at the beginning of their response; "and then I think the next thing to happen would be..." in the middle of their response; or "So that is what would happen if..." as an end link. The link is a specific and topical reaction to the prompt that focuses and frames the student's answer and relates it to the prompt. It should be explicit in nature. As in explanation scoring, a reiteration of a portion of the prompt is considered to at least be linked for scoring purposes.

LOGICAL OUTCOME

For this step in the scoring process, the question that must be asked is whether the outcome the student presents can realistically happen as a result of the prompted situation. For example, in response to a prompt in which students are asked to predict what would happen if NASA received more funding, a predicted result stating that Captain Kirk would travel to Vulcan to save Spock may be linked, but highly illogical. More logical results of increased funding for NASA would be industry in space, the building of a lunar base, a manned mission to Mars, the capture and mining of asteroids, or creation of O'Neill colonies.

THE NUMBER OF STEPS

Events are the concepts, key words, attributes, examples, logical steps, or ideas that have been listed on the content planning worksheet as being part of the class curriculum. They are the individual steps, points of argument, or ideas a student mentions in the process of answering a prompt. The events are offered as evidence for a position. This evidence becomes the scaffold of the student's prediction. Occasionally students may offer correct ideas or concepts that were not listed on the worksheet. These ideas should be included as correct events. Incorrect concepts, attributes, examples, and ideas should not be scored as events although they should be monitored, for such incorrect notions may indicate mislearning.

The count of events is a simple two-level count of the number of related steps in a causal chain that the student has listed in the prediction. This count includes the outcome and any related events listed in the response, but it does not include the prompted situation. The question to ask when scoring is whether there is a single event (the prediction only) or multiple steps listed.

SEQUENCING

As discussed in the section on explanation scoring, sequencing is the listing of events by the student in a careful, step-by-step manner so all events, information, and their relationship to the prompt and outcome are clearly related. Sequencing may be viewed as the linking of one event to another in a careful chain that is coherent to the reader.

The student-generated sequence may be in (a) a series format of different events given in a chronological sequence or (b) a parallel format whereby the student builds a case for a single step by presenting several simultaneously occurring events. As examples, consider the NASA funding prompt once again. A *series format* would be laid out in chronological order: "NASA would work on the space plane, bring people into orbit to work, and build a space station." An answer using a *parallel format* would address several issues for each step of the answer, but the chronological sequence would be less overt: "NASA would develop the space plane, mass drivers, and a space colony. People would go into space to go get an asteroid to bring back to Earth for processing."

CONCLUSION

When scoring prediction answers, the presence of a prediction, its relationship with the prompt, its causal link, logical derivation, the number of causal steps or events listed, and sequencing are considered. It is important to list concepts, key vocabulary, critical features, and facts on the content planning worksheet *before scoring* so focus can be maintained while scoring. When multiple scorers are involved in the task, several decisions must be made prior to scoring, such as what to do with deviations from the content planning worksheet and what to do with redundant events. Such prior decisions will maintain inter-rater reliability and make scoring easier.

TRAINING PROMPTS

SCIENCE PROMPT

The prompt for science read, "Predict where a drop of water going down your sink will be in 7 months and describe at least 3 changes it may go through." This prompt was straightforward and easy to read. Students were asked to predict and describe. The interesting part of the prompt is that it taps heavily into the students' imaginations. They could go anywhere with their answers as long as they touch upon three changes in the state of water.

The students were assessed for their use of the concepts condensation, evaporation, freezing, and precipitation. The idea was to let students freely explore the phases of water and let their imaginations wander. The events that counted for scoring purposes are listed in Figure 2.

RESPONSES ON THE SCIENCE PROMPT
ABOUT A WATER DROP

Events listed among the answers offered by the students were:

- Evaporation
- Condensation
- Precipitation
- Liquid
- Clouds of water

*Also acceptable were other analogous statements such as, "It will fall as rain."

Figure 2. Possible correct responses in scoring science prompt.

In the following analyses, each answer is presented as it appeared in the student's answer. Following the student answer is the scorer's response to that answer and an analysis. The flowchart path is italicized with the score given following the analyses.

HIGH SCORES

TRENT

> 3 changes the water may go through are evaperation, liquid, and ice. I think it might go down to the ocean and evaperated in the clouds and come down as snow some where over (say) Italy.

ANALYSIS

A prediction is offered by mention of snow over Italy, and the answer is *clearly linked* to the prompt. The student addresses the three states. The prediction has a logical outcome, more than one step is listed, and all steps are clear.

Score
5

HELEN

> It will change from gas, liquid, solid and it will go through my sink to the ocean and back up to clouds then back down.

ANALYSIS

This answer *addresses the prompt* and *all three states are listed,* however, *the explanation of the states is weak.*

Score
4

TONY

> The drop of water would be in any river or ocean & then it would evaporate, condense & then fall back to the ground.

ANALYSIS

This answer *addresses the prompt* and *all three states are listed,* but, *the explanation of the states is weak.*

Score
4

MIDDLE SCORES

BETH

> It will be in the ocean It would come out of a pipe fall into another I would go trogh pipe to swer to ocean.

ANALYSIS

There is a *definite link to the prompt*, but the student only addresses the logical outcome of the drop of water's journey to the ocean as liquid. *Other states are not addressed.* The outcome is *questionably logical* and is certainly *realistic.* The problem with this answer is the lack of expansion beyond the liquid form. The prompt is meant to pull out the triple state of water, but there is only a *single state listed.*

Score
3

JORDAN

> The drop of water that drops down the sink will be in the water again. The water during its journey will change from liquid

ANALYSIS

The answer is incomplete and therefore is questionable as to whether the answer would be logical. Clearly, the answer is an *attempt to answer the prompt* and is *clearly linked.* Because the answer is incomplete, scoring conflicts can arise over the student's logic. There is a *listing of one step* and if the student is given the benefit of the doubt, it is possible to derive a score of "3." However, if the outcome is viewed as illogical, then the score is "2."

Score
2/3

ED

> First, the water might evaporate into a cloud heading east, then after in becomes saturated. Might become a snow flake in minnesota.

ANALYSIS OF ED'S ESSAY

The student clearly tried to answer the prompt but the *link is weak.* The outcome is *logical* and *multiple steps* are addressed, however, the explanation of those steps is *unclear.*

Score
3

ALEX

it would go throw ground water and all kind of water

ANALYSIS

Alex *attempted to link* the answer and gave a *logical outcome*, however, the *three states of water are not addressed*, only liquid is mentioned.

Score
2

LOW SCORES

JENNIFER

evaporation, precipitation, condensation.

ANALYSIS

This student offered *no link to the prompt*. The *outcome is logical*, but it is *not developed* as was prompted and is *not a prediction*. The student offers only vocabulary without definition so it is highly questionable whether real understanding is present.

Score
0

SOCIAL STUDIES PROMPT

The social studies prompt read, "What would have happened to Northern industry if slavery had been abolished in 1850 and no more cotton was produced in the South?" The wording of this prompt was extremely tricky for some students. With a cursory examination of the prompt, the natural assumption a student would make would be to discuss the impact of abolition on southern industry. In fact, some students did interpret the prompt in this manner. This prompt gives the typical question a twist so students must explore the interconnectedness of industries in the North and South during the mid-nineteenth century.

The students were assessed for their use of concepts discussed in class. Because of the nature of the prompt, students were able to go beyond the scope of the typical questions and address both sides of Civil War industry. Some events chosen by students are listed in Figure 3.

RESPONSES ON THE SOCIAL STUDIES PROMPT ABOUT THE CIVIL WAR

Events listed among the answers offered by the students were:

- Purchase cotton from another country
- Raise prices of clothing
- Limited clothing
- Ruin northern textile industry
- Effects on workers
- Factory closures
- Economics

*Also acceptable were other analogous statements such as, "People in the North would end up losing their jobs."

Figure 3. Events that were acceptable for scoring as found in student social studies essays.

Student answers to this prompt were quite good. Several students scored either four or five points while the majority of the class scored three points. There were no low scores in the class. Considering the nature of the prompt, with the twist in thinking required, it is surprising that the students did so well.

In the following analyses, each answer is presented as it appeared in the student's answer. Following the student answer is the scorer's response and an analysis. The flowchart path is italicized and score given.

HIGH SCORES

HEIDI

> Well the north wouldn't be growing and making more factory. The north would ask the britsh for help. Things would be very hard. Very few jobs and there wouldn't be enough money .

ANALYSIS OF HEIDI'S ESSAY

This answer has been *carefully related to the prompt.* The outcome mentioned is a *logical* scenario of the prompted situation. There are *multiple steps listed* and these steps are *clearly linked* from step to step. Questions arise when the student assumes that "the north" equals "Northern industry."

Score
5

CRAIG

> If the slaves had been abolished in 1850 the Northern industry workers would have had to shut down the factorys or try to get cotton from Europe planters.

ANALYSIS

The answer is *clearly related* to the prompt and *linked* by the first few words. The outcome is *logical* and there are *two steps listed. There are obvious breaks in the causal chain.* The relationship between Northern industry and King Cotton is present, and mention is made of finding other sources for cotton, but it is *not clear* why this would ruin the Northern industries.

Score
4

TIMOTHY

> What would happen if slavery had been abolished in 1850 and no more cotton was produced in the South would be that the south would go through an industrial revolution and the North would be in compitition.

ANALYSIS OF TIMOTHY'S ESSAY

This answer contains a unique outcome that, although it is perhaps not the most logically derived outcome, could have been a possible outcome of the prompted scenario. The answer is *clearly linked to* the prompt and it *could have been a logical outcome* had the South taken the abolition as a competitive move and put their energies into restructuring. Implied in the answer is that the North would have suddenly found itself in a competitive situation. Nothing is said of the direct effect on Northern industry when cotton suddenly disappears from the market. There are *multiple steps* to the outcome but they are not presented as clearly as they should have been. The number of steps is considered multiple because the student's premise is built upon the Southern reaction. Certainly this could have also been scored as having a single step and given a score of "3" because only one step deals with the Northern industry, however, due to its uniqueness and development, the student shows an understanding of the interrelatedness of the industries. The student could have addressed the type of industries and the effects on the Northern industries to a greater extent. Had the student been more clear in the linkages between the steps, the score could have been a "5."

Score
4

MIDDLE SCORES

KELLY

> There pobably would be no factorys

ANALYSIS

This answer *is linked to* the prompt because of the mention of factories. *It is a logical outcome*, however, there is only the outcome presented and *no intermediate causal steps* are listed. This answer is an example of a student who answered quickly, skipping steps and leaping in their thoughts without relaying the information to the reader. Instruction in how to answer this kind of prompt would be recommended.

Score
3

DANNY

> I think the North would be very poor and the South wouldent have very much clothing. So the South and the North would both be out of money and cotton. But I think the North would find another way to buy cotton.

ANALYSIS

This answer is *not clearly linked* to the prompt. The outcomes are logically derived from the situation but only a *single intermediate causal step* is mentioned. Because of the lack of causal steps, there is no clear link between the outcome and the prompt. Two outcomes are presented: lack of money and clothing, and the North finding another source of cotton. Because neither of these outcomes is presented through a series of events, the strength of the answer suffers. Had the student made more overt the relationship of the prompt to Northern industry, the answer would have received a score of "4".

Score
3

JACOB

> I think that the south would have to do something easle to make money and I think people would have about two pairs of clothes.

ANALYSIS

This answer fails on several points. First, it is *not clearly linked* to the prompt. The student did not address Northern industry but instead addressed the South. Because of this the student did not link the answer to the prompt. *The outcome is logical for the situation but it is not related to the prompt. Only a single step exists*: limited clothing. There is room for variability in this score. If the scorer looked at the answer as having no logical outcome, the score, because of the lack of a link to the prompt, would be "1." If the scorer looked at the outcomes as logical, but as having only one step being addressed, the score would be "2." Certainly this answer could be looked at as having multiple steps listed (lack of money and lack of clothing) and thus be scored a "3," however, this would be pushing the issue to an extreme for it is clearly directed at the South rather than the North.

Score
1/2

CHECK-OUT OF PREDICTION: SCIENCE

Score the following student essays using the prediction scoring flowchart.

KEVIN

from Evaporation liquid to Gas .
Condonation Gas to liquid and freezing liquid to solid

What is the outcome predicted?_____

_____

What events are presented to support their prediction?

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____

Is the prediction carefully sequenced? _____

Score

SARAH

a drop of water going down my sink would be somewhere in the water cycle and it could have been evaporated, condensed, and fell as precipitation

What is the outcome predicted?_____

_____

What events are presented to support their prediction?

_____

_____

_____

_____

_____

_____

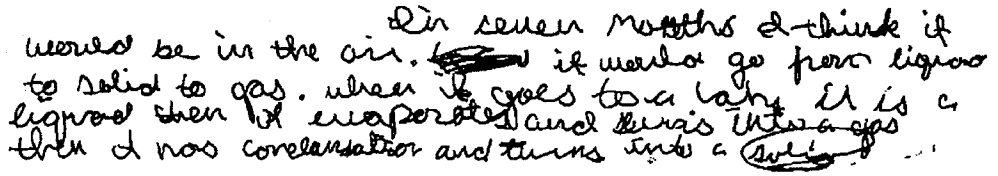
_____

_____

Is the prediction carefully sequenced? _____

Score

DENNIS



What is the outcome predicted?_____

_____

What events are presented to support their explanation?

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____

Is the prediction carefully sequenced? _____

Score

CHECK-OUT OF PREDICTION: SOCIAL STUDIES

Score the following student essays using the prediction scoring flowchart.

TARA

> If slavery had been abolished in 1850 and no more cotton was produced in the South the Northern industry would be practicaly ruined. It would be ruined because the Northerners would have to get cotton from another country. Getting cotton from another country would cost alot more because the Northerners would have to pay for the trip over to America. Since getting cotton cost so much the Northerners would have to raise the price of material, so only rich people could buy it. The Northern industry would be practicaly ruined.

What is the outcome predicted?_____

_____

What events are presented to support their prediction?

_____

_____

_____

_____

_____

_____

_____

Is the prediction carefully sequenced? _____

Score

T.J.

IF there'S no more slaves to pick cotton the southeners would have to hirer payed worker The price in cotton would go up because the southeners would have to pay the workers the Price in the factories production would go up.

What is the outcome predicted?_____

What events are presented to support their prediction?

Is the prediction carefully sequenced? _____

Score

RICHARD

There would not be any cotton clothes for people to wear and get warmer, because if they didn't have slaves go and pick the cotton there would be no reason to have a cotton gin

What is the outcome predicted?_____

_____

What events are presented to support their prediction?

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____

Is the prediction carefully sequenced? _____

Score

ANALYSES OF CHECK-OUTS: SCIENCE

ANALYSIS OF KEVIN'S ESSAY

The student understands the three states of water, but these states are *not linked* to the prompt. There is only an explanation of the three states. *There is no prediction.* This student's answer could have been caused by a lack of understanding of the task. It would be a good idea for the teacher to review proper answering of prediction prompts with this student. Had the answer been linked, the score would have been considerable. The concepts were clearly understood.

Score
0

ANALYSIS OF SARAH'S ESSAY

There is a prediction that is related to the prompt and it is *logically derived,* there is *more than one causal step* listed, however, those steps were *not fully developed* and described as the prompt suggested. Because of this lack of development, the score suffers. The student obviously understands, but didn't take the time to answer fully.

Score
4

ANALYSIS OF DENNIS' ESSAY

Dennis *did not completely address the prompt.* It is understood that water is a liquid in the lake and a gas in the air, but it is also suggested that condensation is solid. Clearly there was a misunderstanding of the prompt or some mislearning has taken place. In this case, it would be a good idea to double-check the student's understanding of solids. As for scoring, the *prediction is linked,* it is *logically derived,* and there are *multiple steps listed,* however, the *chain is unclear* because of the missed information.

Score
4

ANALYSES OF CHECK-OUTS: SOCIAL STUDIES

ANALYSIS OF TARA'S ESSAY

This response is quite good. The student has developed the argument carefully and taken us step-by-step through the prediction. Following the scoring routine, the *prediction has been carefully related to the prompt* and it is *well linked.* The outcome, the near ruin of Northern industry, is *logically derived* from the prompted situation, and *several causal steps* take the scorer from the prompt through the conclusion. All causal steps are clear from the prompt to the outcome, although the answer could have been better had the student addressed the impact of additional workers in the Northern work force.

Score
5

ANALYSIS OF T.J.'S ESSAY

This answer presents the scorer with a problem. The *outcome, prices rising, is logical,* but the *steps that link it are weak.* Further, the answer *does not address the prompt.* The student addressed the South rather than examining the North as prompted. The student also ignored the prompted scenario's point that "no more cotton was produced in the South." The student was trying to keep the South a cotton producer. Though this is more logical than a complete stoppage of production, it is not the prompted situation. Because of this, the student's *link to the prompt was weak.* The outcome is *logical* and there are *multiple causal steps* listed, but the *steps aren't clear* as to how they relate to each other and Northern industry. It would be understandable if a scorer gave this answer a "4." The score would be dependent upon the scorer's interpretation of the linkage occurring in the causal chain.

Score
3

ANALYSIS OF RICHARD'S ESSAY

This student took a great deal for granted that the scorer must interpolate. The chain, when presented clearly, would read, "If there were no slaves, cotton wouldn't get picked, there would be no reason for having cotton gins, and people would have no cotton clothes." This answer addresses Northern industry by the indirect mention of the cotton gin. There is *no clear link* between the two. No mention is made of getting clothing or cloth or cotton from other sources.

Because of this, the link between the prompt and the answer is weak. *The outcome is logically derived*, and there are *multiple steps* presented. Had the student more clearly linked the answer to the prompt and made more overt the connection between their answer and Northern industries, the essay would have received a "5."

Score
4

Appendices

REFERENCES

Bennett, R. E., & Ward, W. C. (1993). *Construction versus choice in cognitive measurement.* Hillsdale, NJ: Lawrence Erlbaum Associates.

Darling-Hammond, L. (1994). Performance-based assessment and educational equity. *Harvard Educational Review, 64*(1), 5-30.

Deno, S. L. (1985). Curriculum-based measurement: The emerging alternative. *Exceptional Children, 52*(3), 219-233.

Frederiksen, J. R., & Collins, A. (1989). A systems approach to educational testing. *Educational Researcher, 18*(9), 27-32.

Koretz, D., Stecher, B., Klein, S., & McCaffrey, D. (1994). The Vermont portfolio assessment program: Findings and implications. *Educational Measurement: Issues and Practices, 13*(3), 5-16.

Linn, R. L. (1994). Performance assessment: Policy promises and technical standards. *Educational Researcher, 23*(9), 4-14.

Lindemann, E. (1987). *A rhetoric for writing teachers (2nd ed).* New York: Oxford University Press.

Nolet, V., & Tindal, G. (1994). Instruction and learning in middle school science classes: Implications for students with disabilities. *The Journal of Special Education, 28*(2), 166-187.

Nolet, V., & Tindal, G. (in press). Essays as valid measures of learning in middle school science classes. *Learning Disabilities Quarterly.*

Oregon Department of Education (1989-1992). *Technical report: Oregon Statewide Assessment.* Author.

Roid, G. H., & Haladyna, T. M. (1982). *A technology of test item writing.* New York: Academic Press.

Shavelson, R. J., Baxter, G. P., & Gao, X. (1993). Sampling vriablity of performance assessments. *Journal of Educational Measurement, 30,* 215-232.

Tindal, G., & Nolet, V. (in press). Curriculum-based measurement in middle and high schools: critical thinking skills in content areas. *Focus on Exceptional Children.*

Tindal, G., Rebar, M., Nolet, V., McCollum, S. (in press). Understanding instructional outcome options for students with special needs in content classes. *Learning Disabilities: Research and Practice.*

Tindal, G., Nolet, V., & Blake, G. (1992). *Focus on teaching and learning in content classes (Training module no. 3).* Research, Consultation, and Teaching Program, College of Education, University of Oregon.

Toulmin, S. E. (1958). *The uses of argument.* Cambridge: The University Press.

EVALUATION SCORING SHEET

What is the statement of choice?_____

What facts are presented to support their choice?

_____

_____

_____

_____

What facts are presented to insubstantiate other possible choices?

_____

_____

_____

_____

_____

_____

_____

_____

_____

Qualitative Score	Facts Correct	Facts Incorrect

CONTENT PLANNING SHEET

Content Planning Worksheet

Date: _____

Teacher: _____

Class: _____

Textbook: _____

Other Curriculum Materials: _____

Approximate Schedule of Content to be Delivered

Week	Dates		Textbook Unit	Textbook Chapters	Quiz Dates	Test Dates
1	From:	To:				
2	From:	To:				
3	From:	To:				
4	From:	To:				

KEY CONCEPTS

1. _____
2. _____
3. _____
4. _____
5. _____
6. _____
7. _____
8. _____
9. _____
10. _____

IMPORTANT IDEAS

1. _____

_____

2. _____

_____

3. _____

_____

EXPLANATION SCORING SHEET

What is the link?_____

_____

_____

What events are presented to support their explanation?

_____

_____

_____

_____

_____

_____

_____

_____

_____

Is the explanation carefully sequenced? _____

Is all of the information accurate? _____

Score

Prediction Scoring Sheet

What is the outcome predicted?_____

_____

_____

What events are presented to support their prediction?

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____

Is the prediction carefully sequenced? _____

Score