

Research, Consultation, & Teaching Program Research Report No. 7



A Case Study on Placing Talented and Gifted Students into the Appropriate Instructional Level and Monitoring Learning Rate

Gerald Tindal, Victor Nolet
Carrie Wegmann, Sandra McCollum

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

Staff
Gerald Tindal, Program Director
Jerry Marr, Editor
Abe Deffenbaugh, Assistant Editor

Copyright ©1993 University of Oregon. All rights reserved. This publication, or parts thereof, may not be used or reproduced in any manner without written permission. For information, write University of Oregon, College of Education, Research, Consultation, & Teaching Program, 237 Education, Eugene, OR 97403-1215.

Tindal, Gerald; Nolet, Victor; Wegmann, Carrie; McCollum, Sandra
A Case Study on Placing Talented and Gifted Students into the Appropriate Instructional Level and Monitoring Learning Rate
Research Report No. 7

Preparation of this document was supported in part by the U.S. Department of Education, grant number H029B20030. Opinions expressed herein do not necessarily reflect the position or policy of the U.S. Department of Education, and no official endorsement by the Department should be inferred.

Layout: Jerry Marr
Cover design: George Beltran

A Case Study on Placing Talented and Gifted Students into the Appropriate Instructional Level and Monitoring Learning Rate

Gerald Tindal
Victor Nolet
Carrie Wegmann
Sandra McCollum

University of Oregon

Abstract

This report describes procedures for placing talented and gifted (TAG) students into instructional levels (materials) and ascertaining their learning rates in middle school content areas. The report is organized into two sections. In the first, a model is described that addresses what behaviors to assess, how to assess them, and how to use this information to place students into instructional levels and assess their learning rates. Because the focus is on content instruction, two issues are worth noting: Emphasis is given to higher-order thinking skills and instructional delivery is centered around the content teacher, although a TAG specialist-coordinator may be used to supplement instruction, ensure level, and enhance learning rate. Then, a model is described for placing students into the appropriate instructional level and monitoring their rates of learning. In the second section, a case study is presented, which describes specific strategies for determining level (using a criterion-referenced approach) and ascertaining rate (with an individually-referenced system that tracks improvement over time).

In this report, only two decisions are being considered: placing students into instructional levels (which may also involve grouping students) and determining their rate of learning. Other decisions (i.e. screening and eligibility) are quite unrelated and are not considered.

Definitions

- *Instructional level* is comprised of the areas in which the student has adequate background and knowledge to successfully engage in material or activities, but lacks mastery or fluency in the material or activities.
- *Learning rate* implies two dimensions: (a) the material that has been learned, and (b) the amount of time needed for learning to occur.
- *Classroom-based assessments* must be used to make valid decisions about instructional placement and learning rates. Published, norm-referenced achievement tests lack (a) content validity for placement—curriculum and test items don't overlap—and (b) instructional validity—they are insensitive to instructional programs and changes in student performance.
- *Teacher-derived measures of performance and learning* are stressed in which technical adequacy (consistency and truthfulness) are ensured by careful development and implementation.
- *Behavior samples* may be diverse; however, well-developed assessments must focus on behavior, both during and after instruction.
- *Quantification and qualification of student performance* is highlighted. A wide range of information is needed, some of which is objective (requires counting amounts) and other of which is subjective (requires judgments of quality).
- *Core curriculum areas* from Oregon Department of Education are emphasized in the assessment process; however, other areas could have been considered.

A Model to Serve TAG Students in Content Classes

Current legal mandates require a system that allows students the benefits of instruction from content experts but with the TAG specialists' emphasis on learning strategies to accommodate them. The clear implication is that TAG educators are much less effective at helping students master specific content. As many TAG students reach middle school and beyond, the focus of their educational programs must shift increasingly toward acquisition and use of content knowledge. And it seems appropriate that they should receive instruction from those teachers who have the most expertise in the content areas. However, just as we should not expect TAG teachers to have the expertise required to teach in all the content areas, neither should we expect content area teachers to automatically be able to tailor instruction to fit the needs of every TAG student who comes to their class.

Furthermore, if an important goal of secondary education is not just to get students to pass their content classes, but also actually to learn critical thinking skills and problem solving, three components of educational programs need to be in place:

1. The type of support by TAG personnel should accommodate (a) information provided by general education "experts" in specific content areas, (b) the format of knowledge contained in both the curriculum and instruction, and (c) learning (classroom) tasks which are "authentic" and allow students to manipulate information in a meaningful manner. These last two elements are essentially templates that can be used to organize content information.

2. This support needs to be directed at both the content area teachers (so they can teach students) and to the students directly in the form of strategies instruction (somewhat akin to study skills), so they can learn the information and manipulate it within learning tasks.

3. Instruction and assessment need to be jointly framed to focus on higher-order information (concepts and principles) and allow meaningful manipulation of information. Changes in instruction alone are not enough; likewise, changes in assessment are insufficient. Both must be changed concurrently.

We propose a model of TAG education in which content helps theoretically and conceptually to determine the purpose of curriculum and instruction. This model defines a new relationship between TAG coordi-

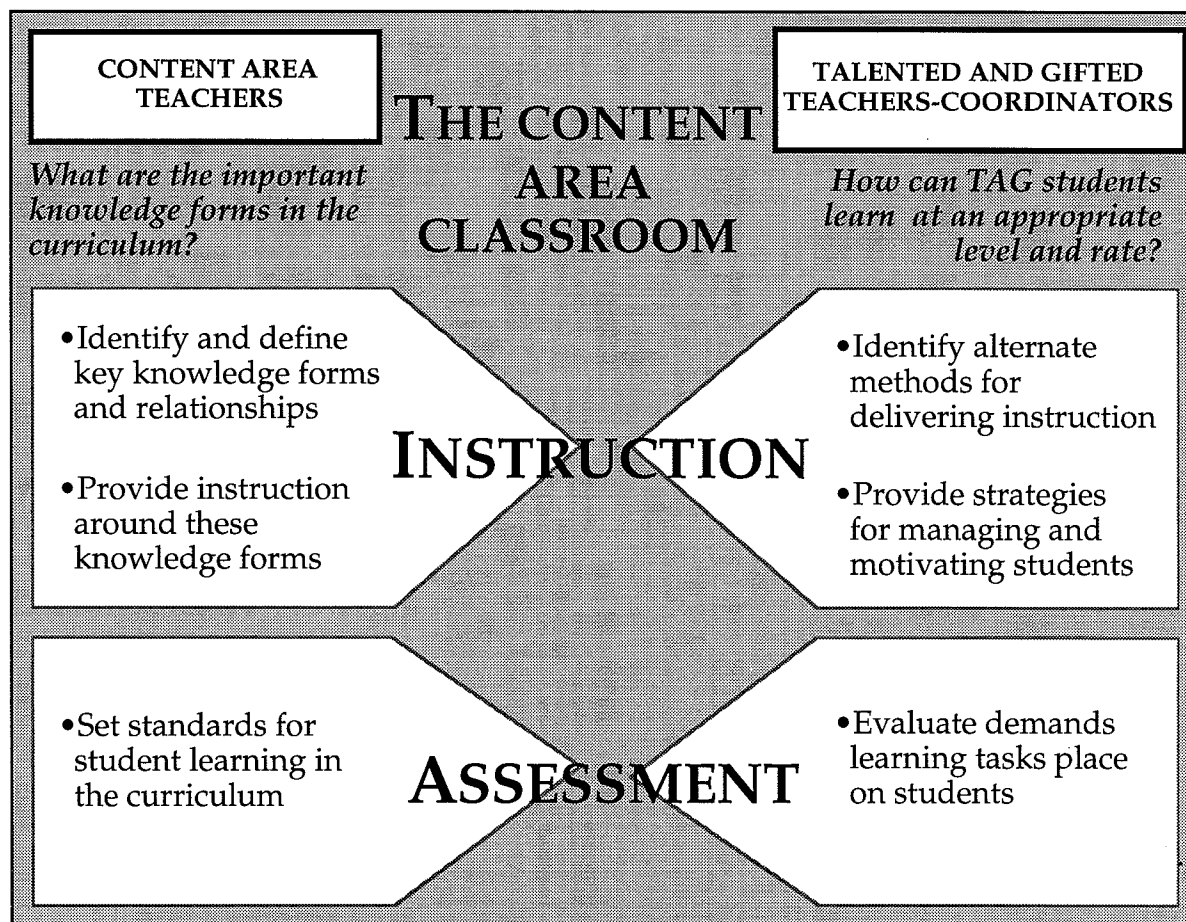


Figure 1. A model of TAG education in content area classes.

nators and general education teachers. The content area teacher brings to this relationship expertise associated with content knowledge of the particular domain. This discipline knowledge permits identification of key knowledge forms (facts, concepts, principles, and procedures) around which content instruction can be organized. The TAG coordinator, in turn, brings pedagogical expertise related to methods for designing instruction, classroom management, and motivational strategies effective with TAG students. Finally, both need to consider assessment in terms of the standards for learning (from the content teacher's perspective) and the demands made upon students (from the TAG coordinator's perspective). This model would redefine the working relationship between TAG and general educators. While it is assumed that students receive their primary instruction from the content experts, TAG coordinators have a role in helping support students through curriculum and instruction. They also must become familiar with the demands set by the content area teachers not only in terms of the criterion for success (i.e., performance on the test) but also the content of the assessments. Figure 1 illustrates this new relationship.

Model Components to Help Inform (In)dependent Variables

In describing a model for delivering programs to TAG students, we want to build from the empirical findings and theoretical positions that have been developed in the last decade. Much of this research is quite recent, with most of the investigations conducted in general education. However, if TAG education fails to remain aligned with general education, any success will be limited or nonexistent. Although a considerable amount of the research in five important areas has been conducted in general education classrooms, we believe this research should be directive, particularly because we propose instruction in general education environments for students with unique talents. TAG educators must be familiar with the research on teacher expertise, curriculum analysis, interactive teaching, learning strategies, and learning outcomes.

Content Teacher Expertise

Our research (and subsequent case study) is embedded within the current debate on teacher knowledge, sparked in great part by Shulman's (1986, 1987) perspectives. We have incorporated some of these views in both our rationale and treatments, in which content knowledge can be configured as subject-matter, pedagogical, and curricular. For example, Shulman (1986) writes that "teachers must not only be capable of defining for students the accepted truths in domains. They must also be able to explain why a particular

proposition is deemed warranted, why it is worth knowing, and how it relates to other propositions" (p.9). We argue that instruction for TAG students can bring learning of content subject matter alive so that subjects like history are not viewed simply as a matter of fact but as one of interpretation (Wilson, 1988). "Teachers must critically understand a set of ideas, a piece of content, in terms of both its substantive and syntactic structure...Teachers should also understand the relationships between that piece of content and the other ideas within the same domain as well as ideas in related domains" (Wilson, Shulman, & Richert, 1987; p. 119). In this respect, we agree with McEwan and Bull (1991) that "all subject matter is pedagogic" (p. 331).

Such views of content instruction place great emphasis on the role of teachers as content experts with whom students serve "cognitive apprenticeships" (Brown, Collins, & Duguid, 1989). This new role for content teachers goes beyond mere interpretation of curriculum and attends to the knowledge teachers bring to the classroom to inculcate students in a discipline. For example, in geography, students would be taught to "think like a geographer" (Harper, 1966; 1990) or in mathematics, to "think mathematically" (Schoenfeld, 1990).

In our view of subject matter knowledge, we place the content expert in a pivotal position, captured most eloquently by Wineburg (1991) in his study of the differences between true historians and advanced placement high school students of history. He found that historians and students were dramatically different on practically any criteria he selected (i.e., features to which they attend, knowledge forms they structure, interpretations of content, views of controversies, etc.). Although one might assume that topical knowledge (names, dates, and concepts) separated them most, in fact, some of the students actually performed better than the historians on questions requiring reiteration or summarization of factual information. Wineburg suspects that "knowing history is more complicated than answering such short answer questions. That students so rarely saw subtexts in what they read; that their understanding of point of view was limited to which 'side' a document was on; that they rarely compared one account to another, searching instead for the right answer and becoming flustered in the face of contradictions—all hint at something far greater than knowing more names and dates" (p. 510). This issue is especially important for students with unique talents and gifts: They must be allowed the opportunity for seeing the forest through the trees and not only knowing what information is important but knowing how and when to use information so it becomes more functional and operational.

Wineburg (1991) describes the difference between historians and advanced students using the courtroom

as a metaphor. Historians worked through the documents as if they were prosecuting attorneys, actively interacting with testimony, locating discrepancies, and actively questioning sources. Students, however, were like jurors, "patiently listening to testimony and questioning themselves about what they heard, but unable to question witnesses directly or subject them to cross examination. For students, the locus of authority was in the text; for historians, it was in the questions they themselves formulated about the text" (p. 511).

This view of the content expert is central in our model of teaching and learning and has great bearing in the eventual use of texts and tests to teach and measure content comprehension. Suddenly, comprehension becomes subject-matter bound, with the discipline shaping both the content and standards for judging competence. Teaching becomes more than a mere collection of methods but the transformation of knowledge, in which expertise is absolutely requisite. And the expertise of teachers is commanding in its influence on both content and format.

Content Curriculum

Teachers become even more important when viewing the centrality (and limitations) of the curriculum in most content instruction. Curricula in the different subject areas cannot be relied upon as the primary instrument of instruction, for in many instances, disciplinary knowledge is misrepresented (Fitzgerald, 1979; Hashweh, 1987; Kantor, Anderson, & Ambruster, 1983). As Ball and McDiarmid (1990) note, "Concepts are often inadequately developed, with just one or two examples given and an emphasis on 'hints and reminders' to students about what to do...learning from textbooks may also contribute to the perpetuation of thin or inaccurate representations of subject matter." (p. 445).

And while most middle and high school courses in the social sciences typically are structured around textbooks (EPIE, 1977; Weiss, 1978), other research on curriculum materials in general-education content area classes has consistently shown that these textbooks are "inconsiderate" (Armbruster, 1984), cover material superficially (Crabtree, 1989; White, 1988), present many facts with little integration of material into coherent representations (Beck, McKeown, & Gromoll, 1989), and contain few unifying concepts or principles (Nolet & Tindal, 1990). "For every topic we pursued, we discovered that secondary school texts presented relatively superficial treatments, without sufficient concepts and depth of supporting knowledge to allow the development of deep understanding" (Baker, Freeman, & Clayton, 1991, p. 138).

While the goals of social science instruction may seem applicable for TAG students, they receive little formal instruction that fosters critical thinking. In a study examining the extent to which geography in-

struction prompted critical thinking, Nolet and Tindal (1990) examined textbook lessons, supplemental curriculum materials, and teacher-delivered instruction in two seventh-grade world geography classes over a period of 3 weeks. They found that an average of 340 separate facts were taught in each chapter. Only 7.5 concepts and 3 principles were presented. Passages such as the following were typical in the textbooks encountered:

Afghanistan is a remote, mountainous country to the west of Pakistan. In size, it is a little smaller than Texas. There are about 60 people per square mile [23 per sq. km]. Most of the people are farmers or herders. Afghanistan has good mineral resources, but they are mostly undeveloped. In 1979, the Soviet Union invaded Afghanistan in support of the pro-Communist government. The government was in danger of being overthrown by rebels. (Patton, Rengert, Saveland, Cooper, & Caro, 1985, p. 360)

In this single, 73-word paragraph, the location, population density, economy, landforms, and recent political history of an entire country are summarized. Eleven separate facts are presented (e.g., "Afghanistan is west of Pakistan"). More important, at least 10 separate concepts are labeled, (e.g., "mountainous country"), but none were actually taught. And this finding was not isolated. These same outcomes were found when the analysis was expanded to five different social studies curricula (Tindal, Nolet, & Blake, 1992). The ratio of facts to concepts per chapter was 1098:1 in 23 pages of a 1989 Heath curriculum at the worst extreme to a ratio of 317:8 in 18 pages of a 1991 Macmillan book.

Interactive Instruction: Planning, Implementing, and Evaluating

Although we want to maintain a curriculum basis to our views of instruction in content areas, we also argue that teachers play an important intermediary role in the transmission and translation of information. While we begin with the teacher as the expert, and then consider the influence of curriculum materials, we now need to consider the actual instructional episode(s).

Our interest in the phases of instruction extends the research on teacher thought processes. However, unlike most of this research, we focus primarily on teachers' content knowledge and the knowledge forms of the content they are teaching. Furthermore, we are more interested in their preactive-postactive planning (thoughts), rather than their interactive thoughts, which, Clark and Peterson (1987) argue, appear to be substantially and qualitatively different. While research on teacher thought processes has addressed both of these categories of experience, much less research has been done on teacher knowledge in secondary content subject matter. The area of research that best approximates

this orientation to teacher expertise has been on their beliefs and theories, most of which have been content-free. While a considerable amount of research has been done on the differences in planning between novice and experienced teachers, this continuum of expertise usually refers to experience in the classroom only, not the depth of their knowledge in any content domain. The research on teacher planning is "almost exclusively descriptive and deals primarily with the planning of experienced elementary teachers" (Clark & Peterson, 1987, p. 267).

The key content issues are: what should be taught, how should it be taught, what supplemental materials need to be co-taught, and how should students be assessed? Our focus on content knowledge is a direct result of our focus on teachers as content experts, and the subsequent influence this aspect has on what is taught and how it is taught. For example, in Banks and Parker's (1990) review of social studies teacher education preparation, they concluded that "this research [the Stanford studies such as Gudmundsdottir, Carey, & Wilson, 1985; Shulman, 1987] indicates that the subject-matter preparation of teachers has a strong influence on the way teachers conceptualize, organize, and teach social studies subjects...teachers use the disciplinary framework in which they were trained to view and structure the social studies courses they teach" (p. 677). In her reviews of the research on teacher expertise, following Wilson and Wineburg's (1988) investigation of history instruction, Carter (1990) notes that "teachers' backgrounds hold great sway on their conceptions of the role of factual knowledge, the place of interpretation, the significance of chronology and continuity, and the meaning of causation in history, and ultimately, the processes and content of their instruction" (p. 306). Finally, Carter (1990) describes Grossman's (1987) research which likewise revealed that English teachers' orientations to subject matter influenced how they planned to carry out content in their classrooms. "Differences in teachers' disciplinary knowledge, background, experiences, and orientations have a significant impact on how teachers organize instruction and represent the substance of the curriculum to students...and that pedagogical content knowledge is...grounded in disciplines and formulations related to school curriculum and the collective wisdom of the profession" (p. 306).

Without teacher expertise, knowledge is likely to be viewed as "certain, neutral, and separate from its moorings in history and practice. That it is not discovered, like wealth by a lone miner panning for gold, but constructed, is not understood; nor is it noted that its construction is social and manifests the interests of dominant groups living in particular historical settings" (Banks & Parker, 1990, p. 682).

Assessment of Student Knowledge: Declarative, Conditional, and Procedural Knowledge

While we want to emphasize authentic thinking tasks, including problem-solving and critical-thinking skills, we also realize that, as Cheney (1987) reports, "the culprit is 'process'—the belief that we can teach our children how to think without troubling them to learn anything worth thinking about, the belief that we can teach them how to understand the world in which they live without conveying to them the events and ideas that have brought it into existence" (p. 5). Student knowledge is viewed as existing along a continuum from declarative to conditional and procedural knowledge (Alexander & Hare, 1989; Paris, Lipson, & Wixson, 1983; Ryle, 1949). "When we know something (be it content, linguistic, or otherwise), we can know not only factual information about it (declarative knowledge), but also how to use such knowledge in certain processes or routines (procedural knowledge). We can also understand when and where this knowledge would be applicable (conditional knowledge)" (Alexander, Schallert, & Hare, 1991; p. 323). These three types of knowledge are generally viewed as being distinct: Acquisition in one form does not automatically translate or transfer into acquisition in the other forms. "Thus it is certainly possible to know the what of a thing without knowing the how or when of it" (p. 323). Similarly, Skemp (1978) distinguishes between relational understanding (knowing what to do and why) and instrumental understanding (knowing rules without reasons). We also have borrowed a conception of knowledge structure from Alexander, Schallert, and Hare (1991) in which content knowledge is divided into domain (broader and more general information that is subject bound) and discipline (specific and highly specialized information that is acquired only over time and study), hierarchically organized (discipline is a subset of domain). In our research, we primarily are focused on domain knowledge for students and discipline knowledge for teachers (in which three characteristics need to be present: a domain, a set of generalizations, and a history).

Organized Knowledge Forms

Domain-specific information in content classes can be organized into three distinct categories: facts, concepts, and principles. These represent the Organized Knowledge Forms that are examined in various contexts in our research. Each of these knowledge forms is defined and illustrated below.

Facts are simple associations between names, objects, events, places, etc., that use singular exemplars. The distinguishing feature of a fact is that the association is very limited and not generalized across a range of names, objects, events, places, etc. The following statements are representative of facts from history, literature, and science:

- William Shakespeare wrote *The Tempest*.
- Harry S. Truman was president when the Korean War began.
- Karl Marx stated, "From each according to his ability, to each according to his need."
- Hydrogen is an element with an atomic number of 1.

Facts are very difficult to remember without an organizing scheme to link them. Yet, they are the basic building blocks to more complex information and are necessary, for example, in developing a vocabulary that can be used to work with concepts and principles.

Concepts are clusters of events, names, dates, objects, places, etc. that share a common set of defining attributes or characteristics. A concept may be thought of as "a category of experiences having a rule which defines the relevant category, a set of positive instances or exemplars with attributes and a name (although this latter element is sometimes missing)" (Martorella, 1972, p. 7) In this definition, rules provide the basis for organizing the attributes of the concept; these attributes, in turn, provide the criteria for distinguishing examples of the concept from nonexamples.

Concepts really form the bedrock for a great deal of teaching and learning in all classrooms. In a middle school math class, students may be taught the following examples of the concept "quadrilateral": rectangle, rhombus, trapezoid, square, and parallelogram. In a political science class students may learn about communism, socialism, and democracy—all of which are complex concepts.

Principles indicate relationships among different facts or concepts, more often the latter. A principle usually represents an if-then or cause-effect relationship, although this relationship may not be stated explicitly. A principle generally involves multiple applications in which the fundamental relationship among the relevant concepts is constant across virtually all examples of the concepts.

Principles form the foundation for organizing information into meaningful chunks. Since principles sometimes involve complex connections among objects, events, or other constructs, it is often difficult to recognize when a principle is operating to organize content. In addition, content area textbooks often state principles implicitly—if at all—rather than explicitly. Occasionally, principles underlying the organization

of content are easily identified. In the physical sciences for example, the law of conservation of matter, the laws of thermodynamics, the various gas laws, and so on serve as unifying principles for much of the teaching and learning that goes on in lectures and lab activities. Similarly, in social studies, economic principles such as the law of supply and demand and the law of diminishing returns can help students begin to make sense of the operation of a free market economy. In the biological sciences, the nitrogen cycle is a complex principle that is often shown in an intricate graphic display as a way to capture the many relationships present.

Procedures often associated with various study skills may be thought of as principles, in that they operate as a series of "if-then" rule relationships. For example, a strategy for finding a word in a dictionary may involve the rule "IF the word begins with the letter 'a', THEN look in the front of the dictionary." Similarly, many reading and writing strategies can be framed as if-then relationships.

Interactive Learning Tasks

Analysis of Organized Knowledge Forms provides a structure for information presented in content classrooms. However, it also is important to identify specific ways students can use facts, concepts and principles. While teachers must be clear on the materials and knowledge forming the core of their instruction, they also must be certain of the behaviors they expect students to exhibit to reflect proficiency. Learning requires more than simply passive receipt of information; a thing is known only when it can be used. Students need to interact with content information, manipulate it, and use it to solve authentic problems.

A variety of frameworks have been described pertaining to the tasks a learner performs, the most widely recognized of which is the one proposed by Bloom and his colleagues (Bloom, Englehart, Furst, Hill, & Krathwohl, 1956). The intent of Bloom's taxonomy was to provide a set of educational goals that would define terms such as "really understand," "internalize knowledge," "grasp the core essence," and "comprehend" to facilitate discussion among teachers about curriculum and evaluation using a common vocabulary. However, Bloom's taxonomy may be problematic for a number of reasons. Contrary to the idea of hierarchically arranged levels proposed by Bloom, as few as two levels may actually exist (Madaus, Woods, & Nuttall, 1973). Also, because Bloom's taxonomy does not use definitions of the operations at each level, inter-rater agreement may be low among educators attempting to classify items with the framework (Seddon, 1978). Finally, Bloom's system pertains to the manner in which the learner acts on information but tells little about the nature of the information itself (Roid & Haladyna, 1982).

To avoid problems associated with Bloom's taxonomy, we focus on Interactive Learning Tasks (ISC), which are identifiable behaviors that allow us to determine whether or not students are able to manipulate content knowledge in meaningful ways. We use a framework that employs six components: reiteration, summarization, illustration, prediction, evaluation, and application. Our framework is particularly useful in describing the independent variables in the proposed research project because each of the operations represent directly observable behaviors that can be linked to the content and structure of instruction. Table 1 presents a description of the six types of Interactive Learning Tasks.

ships. Illustration can be used only with concepts and principles since it requires an individual to recognize or generate previously unused examples of a piece of information. A fact cannot be illustrated because it consists of a single simple relationship between constructs. (How would you illustrate the fact, "Shakespeare wrote *The Tempest*?). Similarly, none of the other three operations—prediction, evaluation and application—can be employed with facts alone.

Prediction of a concept occurs when a student is given some, but not all, of the defining attributes of a concept, and then uses these attributes to infer other attributes of the concept. Prediction of a principle

Table 1. Six Types of Interactive Learning Tasks

INTERACTIVE LEARNING TASKS	refers to the behavior employed in using or manipulating knowledge forms.
Reiteration	<i>A verbatim reproduction of material that was previously taught.</i> <ul style="list-style-type: none"> The emphasis is on <i>verbatim</i>. The wording in the student's response must be very nearly identical to that presented in instruction.
Summarization	<i>Generation or identification of a paraphrase, rewording or condensation of content presented during instruction.</i> <ul style="list-style-type: none"> The emphasis here is on previous presentation of material. Therefore, summarization involves remembering information to a much greater extent than manipulating it.
Illustration	<i>Generation or identification of a previously unused example of a concept or principle.</i> <ul style="list-style-type: none"> The emphasis here is on use of an example that was not presented in instruction. In this respect, the student is expected to employ information about the attributes of a particular concept or principle rather than to simply remember whether or not an event exemplifies a knowledge form.
Prediction	<i>Description or selection of a likely outcome, given a set of antecedent circumstances or conditions that has not previously been encountered.</i> <ul style="list-style-type: none"> Again, the emphasis is on the <i>use</i> of information in a novel context rather than remembering a response from previous instruction.
Evaluation	<i>Careful analysis of a problem to identify and use appropriate criteria to make a decision in situations that require a judgment.</i> <ul style="list-style-type: none"> Evaluation focuses on decision making. The student must first recognize or generate the options available and then use a set of criteria to choose among them.
Application	<i>Description of the antecedent circumstances or conditions that would be necessary to bring about a given outcome.</i> <ul style="list-style-type: none"> Application is the reverse of prediction. The student must use information about a concept or principle to work backwards from the circumstances presented and tell what happened to create it.

These six learning tasks interact with the knowledge forms described below so that some operations may be more realistic than others with some knowledge forms. Reiteration and summarization can occur with all three knowledge forms—students engage in these forms of behavior when asked to recite facts, recall definitions of concepts, or restate lawful relation-

ships. Illustration can be used only with concepts and principles since it requires an individual to recognize or generate previously unused examples of a piece of information. A fact cannot be illustrated because it consists of a single simple relationship between constructs. (How would you illustrate the fact, "Shakespeare wrote *The Tempest*?). Similarly, none of the other three operations—prediction, evaluation and application—can be employed with facts alone.

Prediction of a concept occurs when a student is given some, but not all, of the defining attributes of a concept, and then uses these attributes to infer other attributes of the concept. Prediction of a principle

initial state and then determines the conditions (causes) required to achieve that outcome. Although it is possible to use application with concepts, application works more readily with principles, due to the complex nature of the intellectual operation involved. The same is true of the process involved in evaluation.

Evaluation consists of both analysis of a problem that requires a decision or judgment to determine factors that should be considered in making the decision, and weighting of each of these factors. It involves anticipating consequences of an act and then judging whether those consequences are acceptable according to certain criteria. Evaluation tasks require three basic steps: (a) select criteria; (b) operationalize criteria; and (c) make a judgment based on these criteria. The judgment needs to be supported by the criteria.

Logical and illogical interactions among Organized Knowledge Forms and Interactive Learning Tasks are shown in Figure 2.

The boxes in the lower left side of the diagram are shaded darkly to show that facts cannot be manipulated in illustration, prediction, evaluation, or application operations. Similarly, the two boxes at the bottom of the concepts column are shaded lightly to show that evaluation and application of concepts depends on the specific attributes of the concepts themselves and the manner they are taught.

These interactions among the knowledge forms and Interactive Learning Tasks are important in the conceptualization of the independent variables in this study. Our classification scheme is based on the work of Miller, Williams and Haladyna (1978), with one

important distinction. Their typology was aimed at development of test items, particularly for criterion referenced tests. Rather than limiting these tasks to testing situations, we have incorporated them into the instruction itself. We incorporate ILTs within actual instructional episodes and use them to prepare students to manage and manipulate information. In essence, students are taught directly how to engage in such learning tasks. Clearly, it becomes important for the format of the information (facts, concepts, principles) to align with the learning task. For example, if teachers expect students to make predictions on authentic tasks, content cannot be formatted as facts during instruction. Therefore, the context in which the Organized Knowledge Forms and Interactive Learning Tasks are observed and manipulated represents an important dimension of the independent variables.

Contexts for Organized Knowledge Forms and Interactive Learning Tasks

The specific contexts in which Organized Knowledge Forms and Interactive Learning Tasks are observed and manipulated are (a) curriculum materials, (b) teacher planning, (c) interactive instruction, (d) student perceptions, (e) learning strategies, and (f) assessment tasks. Figure 3 illustrates the interactions among OKFs and ILTs.

As in the previous diagram, the darker shading in boxes on the front of the cube indicates that illustration, prediction, evaluation, and application are not possible

		<i>Organized Knowledge Forms</i>		
		Facts	Concepts	Principles
<i>Interactive Learning Tasks</i>	Reiteration	Yes	Yes	Yes
	Summarization	Yes	Yes	Yes
	Illustration	No	Yes	Yes
	Prediction	No	Yes	Yes
	Evaluation	No	(Yes)	Yes
	Application	No	(Yes)	Yes

Figure 2. Interactions among organized knowledge forms and interactive learning tasks.

with material formatted as facts. Similarly, the lighter shaded boxes indicate that evaluation and application with concepts is conditional, depending on the characteristics of the concepts themselves. It is clear that serious misalignments may occur when attention is not paid to all relevant classroom contexts for content area instruction. For example, curriculum materials that contain a preponderance of facts may provide few opportunities for students to practice complex learning tasks such as evaluation and application on authentic

these settings. The research base describing the nature of content curriculum materials in content classes has been well established and elaborated in the last two decades. An extensive body of evidence supports the conclusion that such materials tend to be heavily facts-laden across content areas. However, the absolute number of each knowledge form is not as important as their relative occurrence with instruction and student learning. Content teachers typically prioritize key concepts and principles for instruction which can then be

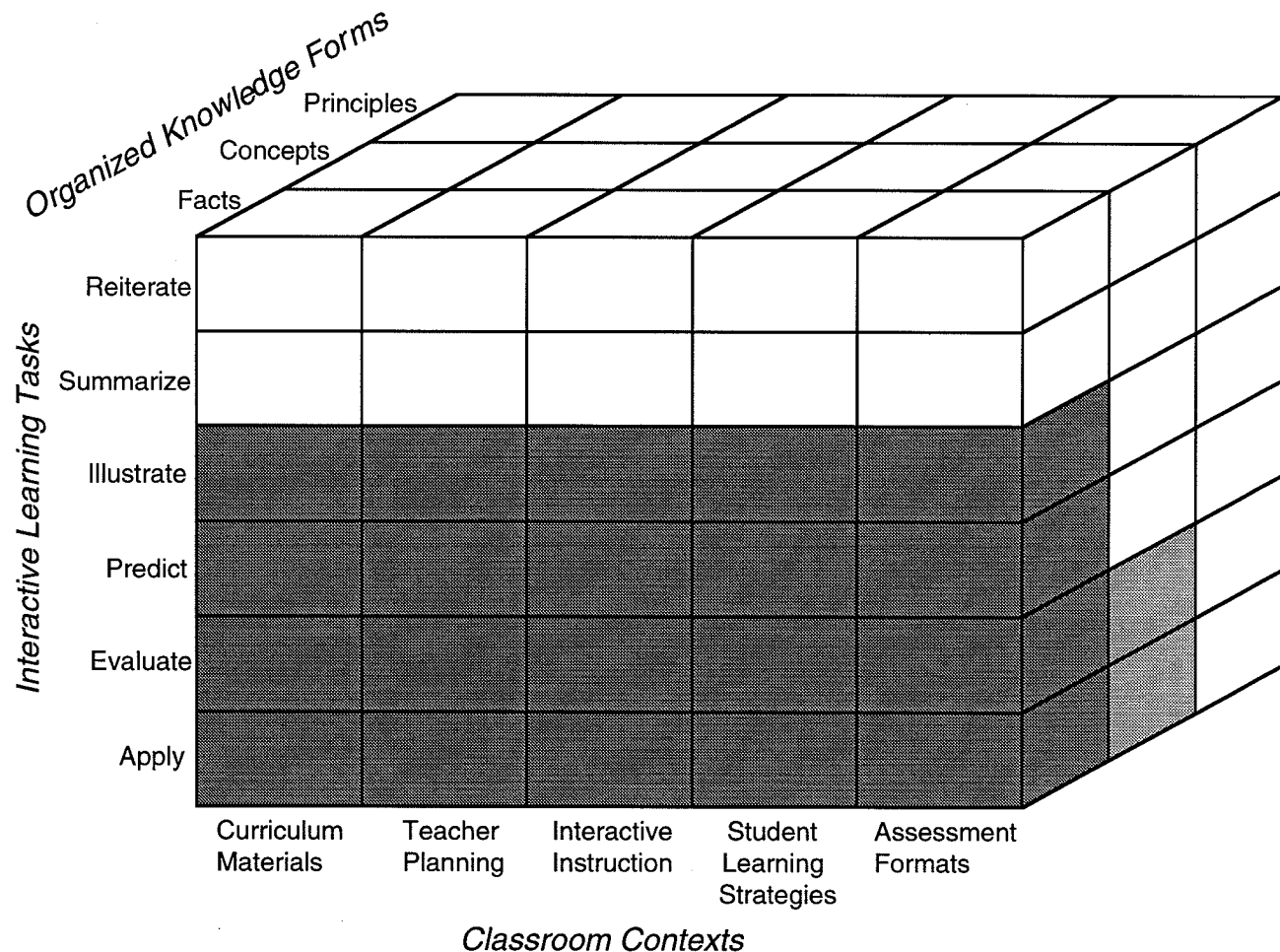


Figure 3. Three dimensions of manipulatable variables.

tasks in subsequent assessment contexts. However, during planning, a content teacher could prioritize certain key concepts and principles in the content, and then, during interactive instruction model use of these with a range of learning tasks. Descriptions of each context are presented above.

Curriculum Materials

Curriculum materials represent the basis for virtually all instruction in content classes. In particular, textbooks are central to the instruction delivered in

used to structure data collection and intervention procedures. Therefore, teacher planning represents a second context in which knowledge forms will be organized.

Teacher Planning

Most of the previous research on teacher planning and decision making has focused on the types of teacher planning (yearly, unit, weekly, daily), the topics around which planning has occurred (objectives, activities,

content, students), teachers' use of specific cognitive processes (use of mental images or problem-solution focus), and classroom outcomes from planning. This research has been conducted with a fairly limited methodology, employing primarily the stimulated recall, in which teachers describe their thoughts in response to viewing a videotape of their teaching. One of the consistent findings from this research is that "teachers seldom followed...formal models of planning and that, during interaction, teachers seldom made logical choices among several different alternatives. Rather, their actions seemed to be largely governed by rules and routines, with decision-making in a studied, deliberative sense taking a minor role in their interactive thinking" (Carter, 1990, p. 297).

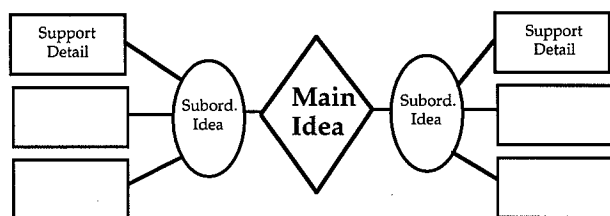
When the dimension of teacher planning includes a novice-expert focus, the concern is largely confined to classroom experience. For example, experts typically draw upon a rich elaborated knowledge of classroom experience to understand teaching tasks and classroom events. They know the common forms of classroom activities (recitation, seatwork, discussions) and assignments; they are familiar with typical behaviors, interactions, and situations that allow them to make highly accurate predictions about what might happen next (Carter 1990). We want to find out how teachers structure content instruction, not from a pedagogical view, but rather with respect to the knowledge forms inherent in the content that they teach. Therefore, we incorporate Carter's (1990) conception of expertise, which is characterized as not only more efficient in

general problem-solving skills, but also is based on "highly specialized knowledge in a particular domain...[which is] organized around interpretive concepts and propositions that reflect the task environments in which they operate." (Carter, 1990, p. 299).

A major variable we used to help organize relationship analysis was teacher-identified concepts and principles. A teacher planning sheet is presented in Appendix A. The first page simply helped us know the gist of the entire unit: the timing for teaching certain materials and content and the overall format they plan to use. The critical information for our case study was the identification of the key concepts and principles listed on page two of this form. These teacher-targeted concepts formed the basis for a number of the interventions and analyses we eventually used in the case study.

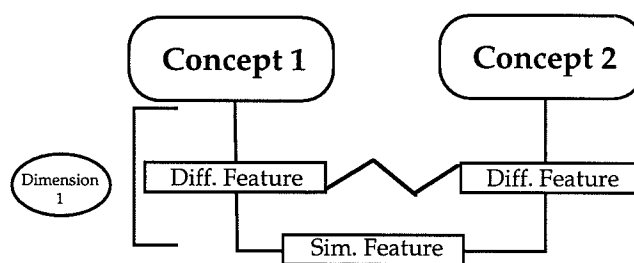
It is unlikely, however, that even a thoughtfully planned and well executed instructional episode can succeed without also incorporating a number of prosthetics to support it. In curriculum modifications, teachers often need to both reduce the amount of information and increase the density and explanatory power of that information. The purpose of this procedure is to instantiate *for teachers* the key relationships among knowledge forms that must be included in their instruction. In this respect, our case differs from previous research in use of graphic organizers (GO) in which they were presented to students. By further clarifying the relationships among key pieces of information, this content can be more directly communicated to students during interactive instruction.

Descriptive or Thematic Map



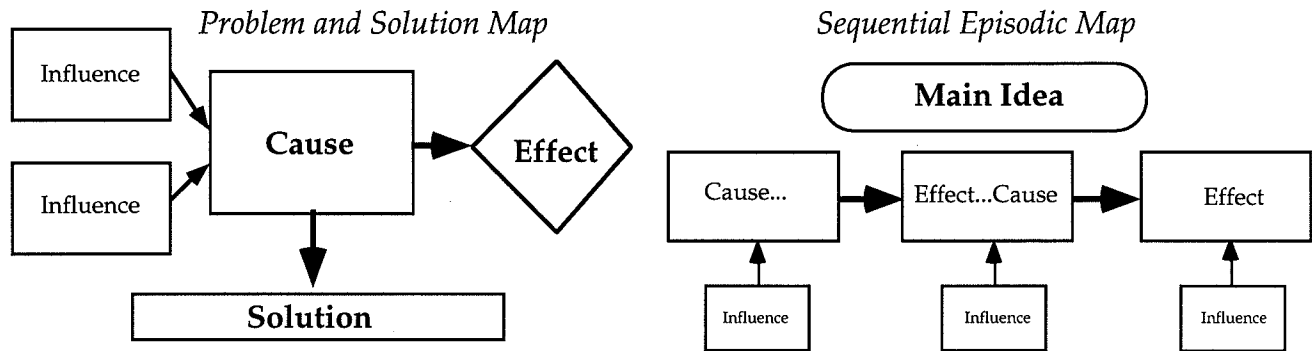
A generic GO, which works well with information that has no underlying cause-effect or time sequence. It works especially well for information that is (or can be) hierarchically organized.

Comparative and Contrastive Map



This type of GO is well suited for organizing concepts by attributes and then comparing them in terms of both similarities and differences.

Figure 4. Graphic organizers for enhancing text structure.



When information can be organized to reflect problems and solutions (often reflecting time or sequenced activities), then this type of GO map is very appropriate.

In this type of GO less emphasis is placed on solutions and more emphasis is given to a sequential chain of causes which turn into effects and further causes.

Figure 4. Graphic organizers for enhancing text structure (continued).

Interactive Instruction

While teaching must encompass expertise and reflections, specific strategies and procedures also must be well defined. These are the nuts and bolts of *how to teach*. More specifically, during instruction, teachers need to manage its delivery by specifying rules, reinforcing these rules, reinforcing, monitoring, and recording responses and monitoring seatwork. They must execute instruction by setting an appropriate pace, providing clear signals, giving students think time, and engaging student's interest by teaching with enthusiasm. Finally, delivery must use accurate corrections, preteaching strategies, and firming cycles. Obviously, we cannot address these components without also addressing the earlier focus on what to teach. We emphasize using the key concepts and principles identified by teachers on the Content Planning Worksheet (see Appendix) to structure communication and teacher self-monitoring.

Activity Structures

We have adopted the notion of activity structure as an important dimension in how instruction is delivered (Doyle & Carter, 1987). The basic premise of this term is that "teachers organize groups of students for work by creating activities... (which) have two major dimensions. First, an activity has organizational properties, including (a) a pattern for arranging participants in the room (e.g., small groups versus whole-class presentations), (b) props and resources used (e.g., books versus films), and (c) duration, the time it takes for the activity to run (typically 10 to 20 minutes of class time). Second, an activity has a program of actions for the teacher and students that guides behavior when an activity is set into motion. The program of action includes (a) roles,

responsibilities, and action sequences for carrying out events, e.g., oral answering or writing workbook entries), and (b) rules of appropriateness that specify the kinds of behaviors that are allowed or disapproved (e.g., talking during snack time or silence during seat work)" (Doyle & Carter, 1987, p. 191). Although classrooms can have a wide range of potential activities and organizational structures, they are generally characterized by only two: (a) whole class lecture or presentation, which typically takes up another 25-30% of the time, and (b) independent seatwork, which comprises approximately 60-70% of the time (Doyle & Carter, 1987).

Teacher and Student Behavior

Teachers can engage in a range of behaviors that structure a lesson. These can include lecturing, demonstrating, and leading students through a series of steps, observing students perform, asking students questions, and answering questions asked by students. Similarly, students can engage in a variety of behaviors, including listening, asking questions, answering questions, performing tasks, and writing/copying. Every classroom utilizes any number of different *tasks* to convey information, which can be interactive instruction (high rates of short bursts of speech by both the teacher and the students), textbooks, media (such as computers and films), tests/quizzes, overhead-chalkboards, worksheets, and activity-labs.

Grouping Arrangements

When instruction occurs, the activity structure can be set with different student grouping arrangements, including whole class, individual seatwork, paired seatwork, and finally, group seatwork. Finally, in all of this (teacher and student behaviors, tasks, and group-

ing arrangements), the *demands* can be set by the task itself, the teacher, or the student. We expect that different teachers use different activity structures; however, it is the relationship between the format of instruction and the other variables that is most important.

Learning Strategies

Students' strategic knowledge has considerable impact on their ability to remember and access content information. Therefore, if we attend only to the Organized Knowledge Forms and Interactive Learning Tasks described above, we are likely to miss a key variable in student learning, particularly those students who come the content classroom strategy deficient. However, we want to avoid simply teaching students traditional study skills. Most study skills interventions focus issues such as locating and selecting information, organizing and remembering it, and acting in appropriate ways to foster learning. For example, Hoover (1989) lists adjusting reading rates, listening, note taking, report writing, oral presentations, use of graphic aids, test taking strategies, using the library (including various types of reference materials), managing time, and self-monitoring. In addition, Smith and Smith (1989) consider study skills to "refer to a wide number of skills necessary for success in any academic setting (e.g., coming to class prepared with pencil and paper, writing down homework assignments on an assignment calendar, turning in completed homework, paying attention in class) as well as those skills necessary to acquire, retain, and recite information" (p. 20).

A critical component of our case is a specific type of learning strategy taught by teachers and used by students to engage in "authentic thinking," i.e., thinking associated with the culture of a discipline. For example, Weinstein and Mayer (1987) have organized learning strategies into eight categories, involving basic and complex rehearsal, basic and complex elaboration, basic and complex organizational, monitoring, and motivating strategies. These categories simply sort the different strategies in which learners review, sort, elaborate, or interact with information. Our interest in this research project is with three strategies used in interacting with complex tasks: (a) rehearsal—"copying, underlining, or shadowing the material presented in class," (b) elaboration—"paraphrasing, summarizing, or describing how new information relates to existing knowledge," and (c) organization—"outlining a passage or creating a hierarchy" (p. 316-317).

We agree with Weinstein and Mayer (1987) that "good teaching includes teaching students how to learn, how to remember, how to think, and how to motivate themselves" (p. 315). Our definition of learning strategies has been borrowed from them: "Learning strategies can be defined as behaviors and thought that a

learner engages in during learning and that are intended to influence the learner's encoding process...[they include] the way in which the learner selects, acquires, organizes, or integrates new knowledge" (p. 315). Similarly, Prawat (1989) describes strategies as a broad range of routines that facilitate knowledge acquisition and utilization. As such strategies are distinguished from specific techniques that are more appropriately characterized as learning tactics (Derry, 1990).

While organization and rehearsal may help students to select and acquire information, such strategies don't help them construct internal connections or integrate the information with prior knowledge (Weinstein & Mayer, 1987). Therefore, elaboration strategies are needed in which learners construct and build internal associations between knowledge forms. Previous research in the area of study skills encompasses this particular intervention in part; however, we believe current conceptions of study skills are too decontextualized to adequately address our focus. Therefore, we have termed the intervention, *learning strategies*, to include both dimensions: (a) specific systems for presenting, organizing, and retrieving information, and (b) manipulation of information using various Interactive Learning Tasks. If we do not consider both dimensions, we are likely to find ourselves in the same awkward position as Purnell and Solman (1991). They used some very sophisticated graphic organizers to teach geography in 11th grade classes and found them to be inconsistently effective.

Rather than using a metacognitive approach which builds on a general problem-solving skill, we argue that use of complex learning tasks such as evaluation and application are content-dependent. In this respect, we view learning strategies as a context in which the interaction between knowledge forms and learning tasks occurs. Learning skills instruction simply cannot be viewed as a substitute for instruction in domain-specific content. This component of our research will involve teaching students the components of knowledge forms and specific learning tasks (along with supporting study skills). This instruction will occur in the context of specific content domains, but will recur using similar formats across content areas.

We use the same strategy investigated by Carrier and Titus (1981), in which students are taught to take efficient notes, defined as containing the greatest amount of information using the fewest number of words. Therefore, a major component of many interventions is an efficient visual format to organize and present information. Students need visual organizers and strategies for sorting and framing information, procedures for managing it, and systems for connecting the concepts

or rule relationships to remember them. As noted earlier, our curriculum modifications are often heavily oriented toward the very organizational strategies that Weinstein and Mayer (1987) review as part of the following research:

- Networking and linking, in which a passage is broken down into parts and the linked relations among the parts are identified (Dansereau, 1978).
- Top level structure to organize relationships among main ideas in expository passages (Meyer, 1975; 1981).
- Prose structures to organize science information so it can be generalized, enumerated, classified, and contrasted (Cook, 1982).

The graphic organizers used with teacher planning serve as the basis for this aspect of our case interventions. Yet, rather than simply using such visual displays irrespective of knowledge forms, they are used as a teacher planning framework in support of a more central dimension: the form of knowledge presented in the content.

Assessment Formats

The final context reflecting manipulation of Organized Knowledge Forms and Interactive Learning Tasks is in the assessment formats used in content classrooms. A range of formats is needed to focus on response demands (production or selection), knowledge forms (facts, concepts, principles), and intellectual operations (reiterate, summarize, illustrate, predict, evaluate, apply). Our premise is that, if all the other contexts are aligned (curriculum, teacher planning, interactive instruction, and student learning strategies), assessment procedures must reflect this. Assessment formats must (a) be representative of the knowledge forms that are prioritized and taught, (b) employ appropriate demands for manipulation of information, and (c) reflect authentic response demands. All previously mentioned variables (e.g., curriculum, planning, interactive instruction, and learning strategies) must be aligned with effective assessment formats. During interventions, then, the assessment formats reflect useful data about student performance. Examples of such assessment formats and scoring systems are included in the case study.

Problem solving tasks should allow us to avoid the difficulty in student responding that Baker, Freeman, and Clayton (1991) describe in their research. Basically, the format (multiple choice) "frames how information is presented, learned, and retained. These tasks assess learning in an artificial, decontextualized manner that is remote from how students learn or will apply knowledge in the future...(and) the problem of showing 'im-

provement' in achievement is difficult and daunting" (p. 134).

We used many of the strategies described by Baker, Freeman, and Clayton (1991) in their efforts to create large-scale history assessment procedures. First, we shifted from measuring broad constructs to important and specific classes of information, knowledge acquisition, "deep understanding, and problem-solving." In this effort, we wanted to capture performance as it was displayed in various tasks and content domains, and as a general construct that could transfer across subject matter domains. Second, we wanted their research to inform ours in the development of qualitative measures for scaling "deep understanding." For example, they began by using two major dimensions in content-quality scoring rubrics: elaborated detail and important, unique material. Furthermore, they had content experts rate the quality of student essays and found that history teachers were no different than English teachers in their scoring and that no clear criteria were being used by the experts. However, when they shifted the rubric development from what experts said they valued (prior to and during their scoring) to how they themselves responded in writing essays to the test prompts (which were identical to the tasks provided to students), they found two significant differences: Inexperienced students (and teachers) often (a) paraphrased or even restated the text in their answers, and (b) tried to cover all the elements presented in the text, failing to distinguish between more- and less-important details. Specifically, they found that experts used the following elements:

- A strong problem or premise that directed a focused answer.
- Use of prior knowledge, including principles as well as facts and events, for elaboration.
- Text references.
- Explicit effort to show interrelationships.

Their third and final finding, therefore, was driven by the need to design assessment situations that reflect and consider the influence of prior knowledge (expertise) and elaboration of information as a means for expressing understanding. Basically, in this last study, they included a more complex test stimulus that incorporated and reflected prior knowledge and utilized a more extended answer which included several complex intellectual operations (summarization and evaluation). They found that deep understanding could be explained with two factors: (a) content quality, use of principles-based prior knowledge, premise-focused writing, and interrelationships, and (b) misconceptions, use of facts, and use of text-based material.

We were not, however, confident that students are familiar with the types of focused writing tasks that we presented. Using a similar task, Baker Freeman, and Clayton (1991) found that "most students reported that...[the] tasks were unusual for them" (p. 149). These authors further stated that "test designers must recognize that the measurement of significant processes takes significant time and consequently tests of many short items and broad content sampling may need to be supplanted or supplemented by fewer, but more complex, assessment situations" (p. 151). In our case study, then, we wanted to allow students to use their own language and yet not place too much load on writing as a tool skill.

Summary of Classroom and Collaboration Components

In summary, the five key features for delivering educational programs to TAG students so they are placed in the appropriate instructional level and learn at an optimal rate involves the following variables: (a) well defined curricula, (b) content experts who can prioritize the information into Organized Knowledge Forms, (c) interactive instructional techniques that facilitate learning, (d) learning strategies that allow students to take control over their learning, and, (e) effective assessment emphasizing both traditional declarative knowledge (what) and conditional and procedural knowledge (how). Of these variables, clearly, the most powerful and most amenable to manipulation are the interactive instructional techniques, learning strategies, and assessment formats. Furthermore, interactive teaching employs use of graphic organizers to highlight the critical concepts and principles either for them to internalize the content or as a strategy with which students interact. Graphic Organizers (GO) both reduce the amount of information and increase the density and explanatory power of that information. Therefore, a major component of the case was development of an efficient visual format to organize and present information. Rather than constructing such visual displays irrespective of knowledge forms, a study-skills framework was used in support of a more central dimension: the form of knowledge presented in the content. Specific knowledge forms were highlighted, followed by identification of an overarching principle that unifies all key concepts and principles, and a number of visual aids were included to help organize the salient knowledge forms.

In all three of the above areas (interactive instruction, student learning strategies, and classroom-based learning outcomes) two major tools are available to teachers for communicating and working with each

other. The first is the Content Planning Sheet (see Appendix A) and the second is actual student protocols on problem-solving tasks, which are conducted frequently over time.

Rather than using the Instructional Planning Form only for content teachers to Organize the Knowledge Forms and develop Interactive Learning Tasks within instruction, we used it as a method for communication between them and the TAG educators. This form requires the content teacher to identify 5-10 critical concepts (including the major attributes and exemplars), 2-3 principles, and a Graphic Organizer that inter-relates all of the OKF with some Interactive Learning Tasks (ITL). The form (shown in Appendix A) is used so the specialist can develop supplemental instruction for students with talents. At the same time, teachers develop alternative measures of performance to track the progress of these students.

The other form, an Individual Student Contract (ISC) had three major components: (a) a project/problem statement, which essentially defined the Organized Knowledge Forms within a content unit, (b) intermediate steps or sub-products, which included teacher models and student practice tasks on using these OKFs in the service of solving problems anchored to specific Interactive Learning Tasks (or intellectual operations), and (c) a final product that specified the form of Interactive Learning Task that could be used in either a criterion- or individual-referenced manner (for assessing instructional level or learning rate). Beyond these educational components, the other components of the ISC were specification of requisite materials and resources, a schedule depicting time and place, and a signature of commitment from the student. The key to this case study, however, was a final product that provided the critical information for ascertaining level and rate.

This case was heavily premised upon the issues noted above but operationalized with the goal of placing students in an appropriate instructional level and monitoring their rate of learning. These two measurement issues formed the basis of the TAG mandate in Oregon, and although they appear to be excellent ideas, they also have been very difficult to develop. More often than not, teachers either have placed students into pull-out TAG programs that have little to do with general education curriculum and instruction, or students are allowed to flexibly participate in the general education program as they (the student) deem appropriate (on the assumption that learning for TAG students should be inherently self-directed). In either program option, level and rate typically have been based on anecdotal reports by either teachers and/or students.

In this case study, we propose a much more formal way of doing business, in which all parties involved (teachers and students) are much more accountable and a data base is developed to vindicate program options or form the basis for making program changes. To reach this level of sophistication, however, Oregon statutes must be clearly understood. Therefore, prior to our description of the case study, we address the model for placement into an appropriate instructional level and monitoring the rate of learning.

Establishing Level and Rate Using Evaluative Standards

Three strategies are available for interpreting student performance and identifying appropriate level and monitoring learning rate: criterion-referenced and norm-referenced, for level, and individual-referenced, for rate. Each strategy provides an interpretive guide that helps teachers understand and appreciate a student's performance. These strategies can be intermixed somewhat, providing us a blend of interpretive guidelines. For the moment, however, consider them as quite distinct.

In a *criterion-referenced* approach, we are not concerned with a student's standing in a group, but with the success of performance on well-defined tasks (skills). In a criterion-referenced approach, we must always define the domain, including how we sampled the items. Generally, we also use some guidelines for determining success, often referred to as mastery or proficiency. In most curricula tests/measures, the cut-off score for defining mastery from non-mastery is somewhere between 75 to 90 percent correct; however, this cutoff can be established anywhere; often this level is tied to grades and teachers' judgments of success.

In a *norm-referenced* approach, the student's performance is compared to other students (who are comparable in age, cultural background, race, sex, etc.). The important index is how the student compares to others—their relative standing. In a norm-referenced approach, a number of different metrics are available; nonetheless, they all interpret performance as relative group standing.

In an *individual-referenced* approach, we interpret performance by comparing a student's score to previous performance levels attained by that student. Rather than compare performance to other students or to some absolute standards of mastery, the important dimension is whether an individual's performance has improved. This strategy is very much like the stock market's Dow Jones Average, which increases or decreases relative to the previous day's performance. All student performance can be interpreted according to these three guides.

Placement in Instructional Level

An appropriate instructional level implies that the student has enough skill and information to interact with the material and/or interactive instruction but is not yet proficient without support. What is needed is a system for ensuring that student's interactions within an instructional episode (with teachers, other students, or materials) are meaningful (successful and extend current performance).

Criterion-Referenced Evaluation

The most obvious procedure to determine if a student is in the appropriate instructional level is to see if the student has the requisite skills to succeed within instruction. Criterion-referenced evaluation closely matches assessment with instruction; it is predominantly used in the classroom and most curricula (i.e. model-lead-test). We refer to it as the "near-sighted" approach: The focus is directed upon objects within a short distance, while objects that are far away are unfocused. To implement this evaluation strategy, three steps have to be completed:

1. A domain of items must be defined. This domain can be included within interactive observations, permanent products, or tests/measures. However, it must be clear in specifying the boundaries of the domain. Which item (task) types are in and which item (task) types are out? For example, the following domains are all clearly specified: addition math facts 1-9; spelling words with a consonant-vowel-consonant (CVC) construction, concepts from the instructional unit on the Civil War (in U.S. History, chapter 9), principles relating velocity and force. In all of these examples, the skill or knowledge content is clearly defined. Not all assessment domains are as clear, as in the following examples: fourth-grade math problems, reading vocabulary words, geography of the Far East, the bio-mechanics of movement. These domains only provide a general picture of the instructional (and assessment) content. Interpretation of performance in a criterion-referenced approach should be as much a part of the domain definition as it is with the score or outcome performance level. When you see the outcome, you should be able to understand exactly what was taught and learned.

2. Once a domain is defined, some strategy needs to be developed for sampling items from that domain. Generally not all items within the domain can be sampled; therefore, we have to come up with a system for selecting only some of them. This system should be sufficient to give us confidence that, although the student did not answer all items, he probably would have answered them correctly.

An obvious example can be constructed from a CVC (consonant-vowel-consonant) domain. Do we

need to present all 100 words to determine whether the student can sound out or spell the different consonants and vowels? Probably not. Instead, we simply can take a random sample, present it to the student, and if performance is adequate, assume that performance on items not presented would have been the same. While several different approaches (too technical to cover here) are possible, you should note that the item selection process can influence the quality of the assessment, particularly with interactive observations and tests/measures; it is uncertain how the sampling system influences permanent product analysis.

3. The last step, determination of mastery or proficiency, is optional, although it is typically included in most criterion-referenced assessments. Some level of performance is set, above which student performance is deemed adequate, acceptable, proficient, etc. and below which it is deemed inadequate, unacceptable, non-proficient, etc. This step, if employed, is fraught with controversy. Many journal articles and entire books have been written on this subject, and given the complexity of this issue. You may want to look at a measurement book to better understand what it means to use a mastery cut-off.

The only two issues that we consider here involve the number of items (tasks) and the certainty of the decision. Generally, more items (tasks) are better than fewer items (tasks), particularly with reference to a mastery measure. Although many tests/measures have as few as one or two items for making mastery decisions, this number is probably too few. Generally, a minimum of 10 items is needed to be certain of mastery. Of course, the number of items (tasks) needed is dependent upon the definition of the domain. With very highly constricted domains, fewer items are needed; with broad domains, many more items are needed to be certain that all those not included on the assessment are mastered. The last issue, certainty of the mastery decision, can best be described as follows. Each mastery measure has three zones: clear mastery, uncertain mastery/non-mastery, and clear non-mastery. The outside two zones (clear mastery and non-mastery) generally present no problem. The dilemma appears with that indifference zone (Shepard, 1984), where there is uncertainty about proficiency. We recommend that you include all three outcomes in summarizing performance.

You should be aware of two serious drawbacks of this system. The most obvious problem, as mentioned earlier, is the difficulty in establishing mastery or proficiency. No clean and proven technology exists yet to eliminate the need for caution. Be sure to include enough items or samples of behavior, and consider all three zones when making decisions. The other problem, though less noticeable, relates to the "near-

sightedness" of criterion-referenced evaluations. Since instruction and assessment are closely matched, and assessment does not occur until instruction has been delivered, two types of error can be made. First, the student is paced no faster than instruction and assessment (no preview performance levels are ascertained). Second, retention is assumed, since each assessment closely focuses only on the material that was taught (no review of performance is ascertained). Both problems can be overcome only by systematically building in preview and review assessments, which are really incorporated into the long range sampling strategy of the individual-referenced evaluations.

Norm-Referenced Evaluation

While the criterion-referenced approach is clear and clean in establishing whether students have the requisite skills to learn within instruction, another procedure that can provide an appropriate backup to any placement decisions is to use a norm-referenced strategy. In this procedure, a task is standardized for administration and scoring and then given to all students in a class or grade level. The essential point of comparison is how any particular student stands in relation to his/her peers. Level is defined as performance within an arbitrary band of one standard deviation within the group mean (which would accommodate about 67% of the group). If a student is above this band, the materials are either too easy or he/she is being instructed within a group of students lacking many of the prerequisite skills to succeed. In contrast, if the student is outside this band on the low end, he/she is lacking the requisite skills. A norm-referenced approach simply adds an element of comparability to any decisions about level.

Individual-Referenced Evaluations

In Carroll's (1963) conception, time to learn is the critical dimension. By tracking engaged time, we have an idea of the amount of time spent learning. However, we also need to know if the material has been learned. The formula states that learning is a function of *time spent actually learning relative to the time needed to learn*. The numerator (time actually spent learning) will be the smallest amount that arises from any of the following three factors: (a) opportunity, (b) perseverance, or (c) aptitude. The denominator (time needed to learn) is determined by (a) aptitude, (b) the quality of instruction, and (c) ability to understand instruction. The ideal outcome would be a ratio of 1:1, that is, the time spent in learning is equal to the time needed to learn.

Yet, assessment of learning rates also must address the content of learning. Such content potentially can influence all of the factors noted above: opportunity, perseverance, aptitude, quality of instruction, and ability to understand instruction. Which specific outcomes should be included? Furthermore, a conception of

learning must include judgments for determining whether learning indeed has taken place. How should performance be judged? When can we be certain that the skills, content, or routines are really learned? Are there any standards we can use to help us make this discrimination?

When considered in conjunction with time spent learning (or time needed to learn), both of the last two issues, performance outcomes and standards for making judgments, provide the necessary components for ascertaining learning rates. However, whether we are using interactive observations, permanent products, or tests/measures, we can only sample a finite amount of learning. We simply cannot assess students on everything we taught them or everything they know. We hope that the performance outcomes on our assessments accurately and adequately reflect all the outcomes that are possible in the domain. In moving Carroll's (1963) conceptual model to actual practice in the classroom, therefore, we must develop a system for sampling specific instructional content (basic skills, content information, or procedural routines) and interpreting performance according to some standards.

This evaluation system, focuses on change over time, with rate of improvement as the primary criterion for determining if programs are working. As mentioned earlier, this system is the "Dow Jones Average" of education. As in the stock market, the two important characteristics involve a historical look at how much improvement has occurred in a certain time period and a prediction about future rates. If you find that stock in which you just invested is beginning to rise, you are likely to hold on to it, at least in the short run. In contrast, if it begins to drop, you may bail out while you can. Although this system is conceptually easy to understand, it has certain technical characteristics that are quite subtle, but very important. In the following material we first describe the critical features, then outline the implementation procedures. Next, we present an analysis that includes both advantages and disadvantages, followed by a system for recording and communicating assessment results.

To compare a student's performance to previous levels and eventually calculate the rate at which learning is occurring requires comparability between all the data values. A criterion-referenced test or measure could not be used in this manner because each measure is a unique opportunity and is not directly comparable to other (previous or subsequent) raw score values. Therefore, the sampling plan needs to be broader than that used for a criterion-referenced evaluation; it has typically been described as long range goal assessment. Rather than matching the content of assessment with that of a *single* instructional episode, the domain for assessment spans *several* instructional episodes.

The major limitation of a criterion-referenced evaluation is "nearsightedness," and the results from any one assessment have little generalizability over time or across items (and is therefore most appropriate for level). In contrast, the sampling plan of an individually-referenced evaluation has built into it both a preview and review component. This feature, however, makes it unlikely for content information assessment. It is best employed with either basic skills or procedural knowledge. Individual-referenced evaluations imply a certain kind of uniformity in the knowledge or skill being assessed, so that both near examples (closely related to instruction) and far examples (appropriate generalizations not directly taught) can be included within any given assessment task. Because content information is so specific, and probably has few far examples that represent generalizations not directly taught, this form of evaluation is restricted.

The following examples all represent an appropriate focus on generalized learning of basic skills and procedural knowledge. In spelling, many words follow phonetic rules quite consistently (not perfectly). We can get a picture of the student's proficiency by including representative items from these various word families; we certainly don't have to include all possible items. Reading decoding words also provides many examples in which various structural characteristics can be considered, with a measure including a representative sample of them. Math computation problems are probably the most lawful in terms of generalizations. In fact, they can be considered as lawful with no exceptions to the rules. Most basic skills, as we have defined them, lend themselves well to this form of assessment. Procedural knowledge also can utilize an individual-referenced evaluation. Here, the generalizations involve rules across different problem types. Most math story problems and science problems include routines that generalize past the immediate items included on an assessment. Although the wording may change and the contexts may vary, the routines nevertheless remain constant. Continuous measurement is possible, since comparability is ensured in the items.

The following steps can be followed in conducting an individual-referenced evaluation:

1. A domain needs to be established that includes all material and activities from the entire instructional series, from the beginning to the end of instruction (i.e., a long-range goal for the end of the year). For example, a teacher may indicate that the students, who are now reading at the beginning of a certain book, should have finished it by the end of the school year. This material represents the long range goal. Likewise, a math teacher may decide that one-step story problems with addition and subtraction represent the long-range goal for a student.

2. Alternate assessment samples are then devised by selecting representative items. For example, a teacher may randomly select passages that are typical of most others in the book and have students read from them, asking them to retell the story after each reading. Or a series of problems may be collected that typify the types of problems in which the student will come in contact during instruction. This step precludes the use of content information for individual-referenced evaluations.

3. The assessments then are collected on a schedule—every week, biweekly, or monthly. In order to develop an adequate database for viewing change in performance, a certain number of data values (7 to 12) should be generated.

In summary, an individual-referenced evaluation uses time-series information to ascertain not only how much improvement is occurring, but also the rate of this improvement. The student's performance is interpreted by looking at previous levels (over time). To evaluate program effects, the general trend of improvement (referred to as slope) and the amount of variation or fluctuation (referred to as variability) should be used. If a program is working, you should see a general trend of increased performance; if the program is not working, such an increase is not discernible.

An individual-referenced evaluation employs the ultimate criterion in education: Are students retaining the necessary information to solve a wider range of problems, some of which may be generalizations of strategies, and with greater automaticity or fluency? Indeed, the programs we devise for students are not really meant to increase a student's relative standing, which is the outcome of a norm-referenced evaluation. An ill-defined state referred to as "mastery," reflected on a list of specific skills, developed from a criterion-referenced evaluation, is not necessarily the goal of education. However, we can be certain that, with an individual-referenced evaluation having a positive trend, improvement has occurred, the program is having an effect, and that such improvement is meaningful. The student's performance is not being compared to that attained by others or interpreted using a list of required skills. Rather, the student is being compared to himself or herself.

Because assessment proceeds from a goal-oriented view, which is "far-sighted," the items represent both preview and review. The content of instruction and the content of assessment are not tightly linked, as in a criterion-referenced evaluation. This strategy, therefore, provides a valuable adjunct to the immediate outcomes from instruction, reflecting appropriate generalizations that have been acquired by the student. For example, if a student is taught the rule for doubling a

consonant when adding a suffix, and then is immediately tested, we might not be very confident that even 100 percent accuracy reflects substantive learning. However, if we have taught the rule, along with other spelling rules, and include items that incorporate many different rules intermittently over many different occasions (alternate forms), we can ascertain whether or not the rule has been learned.

An important part of this far-sighted approach is that any one data value is quite limited in making interpretations. Because the domain for item sampling is broad and the strategy for actually sampling items is often random (within some strata), any one assessment may have a disproportionate number of items of a certain type. For example, in the reading example above, an easy story may have been selected; or the story may have just been covered in class. In either case, performance for that day may be very high. The next measure may sample a more difficult passage or one that has not been taught. Consequently, performance levels are much lower. Therefore, to appropriately utilize this approach, the general trend of data must be considered, not any specific data value in isolation. If the student is becoming a better reader, the number of words read correctly will generally improve over time.

A word of caution is in order, as this approach is not without its drawbacks. Such individual-referenced evaluations must be interpreted in relation to the material used for assessment. A very steep slope may indicate that learning is rapid and there is a ceiling effect, in which performance improvement will no longer be visible (the maximum level has been attained). In contrast, a very low slope, or no slope, may actually reflect assessment materials that are too difficult and are insensitive to change. A floor effect has been attained and a considerable amount of time is needed before any discernible changes appear.

Several other limitations also should be noted. The definition of the domain from which items are sampled can be a major potential problem. Often it is difficult to project into the future. Few standards exist for determining an appropriate amount of material to sample for assessment. Short of applying this strategy and becoming familiar with it, our only sage advice is to pay attention to realistic long-range goals. As discussed below, guidelines for evaluating instruction in this system are quite flexible. In fact, few firm procedures have been identified, except to increase learning rates as much as possible. Finally, domains from content information are probably not appropriate; the system is appropriate for either basic skills or procedural knowledge, where alternate forms reflecting generalizations across item types are embedded within the assessments.

Summary of Measurement Options

In summary, three evaluative systems help delimit specific performance outcomes and allow us to devise standards for interpreting level and rate. These evaluative systems represent different approaches to sampling behavior and interpreting whether learning occurred. When we add in time, we also can ascertain the pace at which learning has occurred.

In a *criterion-referenced evaluation*, the focus of assessment is on *what* the student actually can or cannot do on specific skills and knowledge tasks; performance is interpreted in terms of *mastery* (synonyms may be *proficiency*, *fluency*, *facility*). Importantly, the term implies an absolute standard of acceptance: Interpretations of performance are based on a scale that is non continuous (above this cut-off is considered *mastered* and below it is considered *non-mastered*). The differences between scores within either side of this cut-off are less critical than between those across the cut-off. To help vindicate this judgment, we can use a *norm-referenced* evaluation in which a group of comparable students' performance is compared; this standard often helps define the level of success (mastery, fluency) that is needed. Finally, an *individual-referenced* evaluation reflects change in performance over time. No specific (absolute) performance level is identified; rather, change is noted on a continuous scale, from less to more, noting the direction of change over time and the rate at which it is changing. The materials are not qualitatively different and alternate forms are used to generate comparability across tasks.

Case Overview

The subject of this case study was a 12-year-old seventh grade male student named Josh. He received one period of Talented and Gifted (TAG) services daily; the majority of these initial activities involved independently using a computer to create Hypercard™ "stacks (interactive programs)." The student also participated in a regular social studies class daily (the subject area for this study) during the class period immediately following his TAG period. His textbook for this class was *The World Past and Present* (Bacon, 1991).

The focus of this case study was to operationalize a model for placement of TAG students into an appropriate instructional level and develop a progress monitoring system for ascertaining rate of learning. Both issues were addressed by using an Individual Student Contract (ISC) in which supplemental instruction was implemented, using alternative assessments to generate student work samples, which were then analyzed for level and rate of learning.

In this case, we present the model first (using the ISCs) and then the range of student work samples, all of which were problem-solving essays generated either as a sub-product (during instruction) or as a final product (following instruction). With these work samples, three references are provided for understanding level and rate: a criterion-referenced view that focuses on specific skills, a norm-referenced perspective in which performance is compared to peers (both of these serving to make appropriate placements in instructional level), and finally, an individual-referenced orientation that is used to monitor rate of learning.

The case provides an overall explication and vindication of the issues raised in the first section of this report in which five components were identified for manipulation in any model of collaboration or consultation: (a) teacher expertise, (b) curriculum, (c) interactive instruction, (d) student perceptions and values, and (e) assessment of learning outcomes, including declarative, conditional, and procedural knowledge (the latter two of which set the stage for critical thinking and problem-solving). The most important premise of this case is that students who have been labeled as talented and gifted cannot and should not be served in separate rooms, with different resources, and by specialized teachers, i.e., a pull-out program. Rather, this case reflects a model in which TAG students participate in a "pull-away" program, participating in the content class which is supplemented with extended activities in a range of settings, with a variety of resources, and by various support staff. The model is based on flexible scheduling of all these components with instruction primarily implemented by content teachers.

Development of Individual Student Contracts

To generate student work samples, an Individual Student Contract (ISC) was developed to highlight specific knowledge forms and structure both the Interactive Learning Tasks as well as the instructional arrangements (where and when pull-away sessions would be conducted). A sample ISC used for this case study is depicted in Figure 5. Following is a description of its main components:

Project/Problem Statement. This part contains the organized knowledge forms (primarily concepts) that the teacher wants the student to learn. A problem vignette is created that sets up the context and content for learning these concepts. For example, in this case study, a number of essay prompts are depicted with specific concepts that orient the student to a problem that is to be solved.

Intermediate Steps (Checks). A series of student activities and interim work samples are specified that

provide the student the opportunity to use the organized knowledge forms with specific Interactive Learning Tasks. These steps include instructional activities providing appropriate practice in applying the concepts and the mode of response (in this case, essay writing).

Final Product. This section focuses on a specific intellectual operation or learning task that frames and

focuses the student's response; it also implicitly describes the scoring criteria for judging the quality of the response. In this case, evaluation responses are emphasized.

Materials and Resources Needed. A number of issues are addressed in this section, including a list of the needed materials and books, specification of where (and when) instruction takes place, identification of

<div style="border: 1px solid black; padding: 2px; display: inline-block;"> Individual Student Contract </div>	Beginning Date _____ Completing Date _____	Project/Problem Statement (see attachments)																
Intermediate Steps (Checks)	Materials/Resources Needed																	
Final Product	Location(s) _____ _____ _____	Support Staff _____ _____ _____	Mini-Schedule (next 2 weeks) <table border="1" style="width: 100%; border-collapse: collapse; text-align: center;"> <tr> <td style="width: 12.5%;">MON</td> <td style="width: 12.5%;">TUE</td> <td style="width: 12.5%;">WED</td> <td style="width: 12.5%;">THU</td> <td style="width: 12.5%;">FRI</td> </tr> <tr> <td style="height: 40px;"></td> <td style="height: 40px;"></td> <td style="height: 40px;"></td> <td style="height: 40px;"></td> <td style="height: 40px;"></td> </tr> <tr> <td style="height: 40px;"></td> <td style="height: 40px;"></td> <td style="height: 40px;"></td> <td style="height: 40px;"></td> <td style="height: 40px;"></td> </tr> </table>	MON	TUE	WED	THU	FRI										
MON	TUE	WED	THU	FRI														
<div style="display: flex; justify-content: space-between; align-items: center;"> <div style="text-align: center;"> Student Signature </div> <div style="text-align: center;"> Student Signature </div> </div>																		

Figure 5. Example of an individual student contract.

who is responsible for helping the student (e.g., teachers, parents, non-educational mentors, etc.).

The basic premise of this ISC is that the core content be delivered by the general education content teacher but that supplemental instruction be organized through a specific and structured negotiation with the student, TAG coordinator, content teacher, and any appropriate support staff. The contract is designed to be flexible so content instruction can be adjusted (compressed and expanded) as needed; once the core content is covered (mastered), teachers can structure a number of activities that keep students connected and in pace with the rest of the class while learning new skills and knowledge (i.e., placement in the appropriate instructional level). At the same time, the system is designed to allow teachers and TAG coordinators to ascertain the student's progress over time (moving across final products) to document rate of learning. No assumptions are made about where instruction takes place; on occasion, students may be kept in the content classroom; other times, (supplemental) instruction may take place in the media center, the community library, a business location, a resource room, etc. Likewise, specification of appropriate support staff is left open and may change across units of study. For example, various science units may involve support staff from the local community, cross age peers, or cross-grade content teachers. The critical feature of the ISC is that a problem-project be centered on organized knowledge forms and result in well-structured intellectual operations or learning tasks. Then instructional level and rate of learning may be determined. Figure 5 depicts all the main components.

Overview of Study Phases

During the year, this study underwent three different phases, with each one structured by an ISC (see Figures 9-11). Carmen, a student teacher, served as the primary support staff to work with the content teacher (in social studies) and Josh, the student. Most of the ISCs were structured around specific units in the content area, each of which lasted approximately 2-3 weeks. On occasion, the same ISC was used across several units.

In the first phase of this study, a two-week project was developed in which the student completed several essays based on the use of an evaluation (intellectual) operation. Josh used a computer to gather information through a CD ROM encyclopedia, and to compose essays during instruction and assessment. The final two phases of the study involved pulling the student out of his social studies class 2 days each week and administering essay writing tasks. These sessions aver-

aged 25 minutes in length, with essay scoring criteria briefly reviewed, the social studies key concepts were reviewed if necessary, and the student independently wrote an essay without any prompting from the student teacher. During this phase, Carmen sat next to Josh and was available to answer questions; however, he never asked for any kind of assistance and, when queried, he always indicated no help was needed. Although Josh was allowed to work throughout the 47-minute period, he usually finished within 20 minutes.

Following is a general description of the phases of the intervention, from the 2-week activity described above to the short essay-based projects. Each phase is depicted graphically with a timeline specifying when intermediate activities (sub-products) and final products were completed. After this overview, the student's performance relative to his class is discussed in depth. Finally, several samples of the student's work, drawn from the various projects, are presented from all three reference standards: criterion-, norm-, and individual-.

Placement in Instructional Level: Criterion-Referenced Perspectives

Placement in an appropriate instructional level can be completed by either a criterion perspective, which defines the range of skills, or a norm-referenced perspective, which defines the comparability of skills. As we have defined instructional level, the critical component revolves around success: Can the student engage successfully in the work? How successful is the student compared to age-grade appropriate peers? These two questions are answered only by referencing criterion and normative standards, respectively.

In this case study, the criterion-referenced standards were anchored to the Interactive Learning Tasks or intellectual operations that were noted in the first section of the report. Specifically, six different tasks or operations were considered: (a) reiteration, (b) summarization, (c) illustration, (d) prediction, (e) evaluation, and (f) application. Student performance was assessed by analyzing work products and ascertaining their presence. Once these intellectual operations were defined within the assessments (i.e., a system was established for evaluating performance quantitatively and qualitatively), the appropriateness of instructional level was determined by assessing the student's performance on an absolute, fixed scale or by finding the student's relative standing in a distribution of peers.

The first phase/project occurred while the class was studying Greece and Rome. The project began with a take-home concept identification worksheet in which the student was asked to write several sentences about each of the concepts listed (which were drawn

from a content planning worksheet). Most supplemental instruction, however, was completed later in the unit using one period pull-away supplemental instruction twice per week. The timeline for the entire unit is depicted in Figure 6.

Josh gave minimal explanations for each of the concepts and principles, and his student teacher was dissatisfied. Because the planned teaching focused specifically on the concepts of *democracy* and *republic*,

5. Creating a dialogue, or short play, between two hypothetical characters (1/25).

The student spent six class periods (over 6 school days) working on these sub product essays. He was pulled away from his regular social studies class and received initial instructions from Carmen, the student teacher. He usually worked independently, but was supervised.

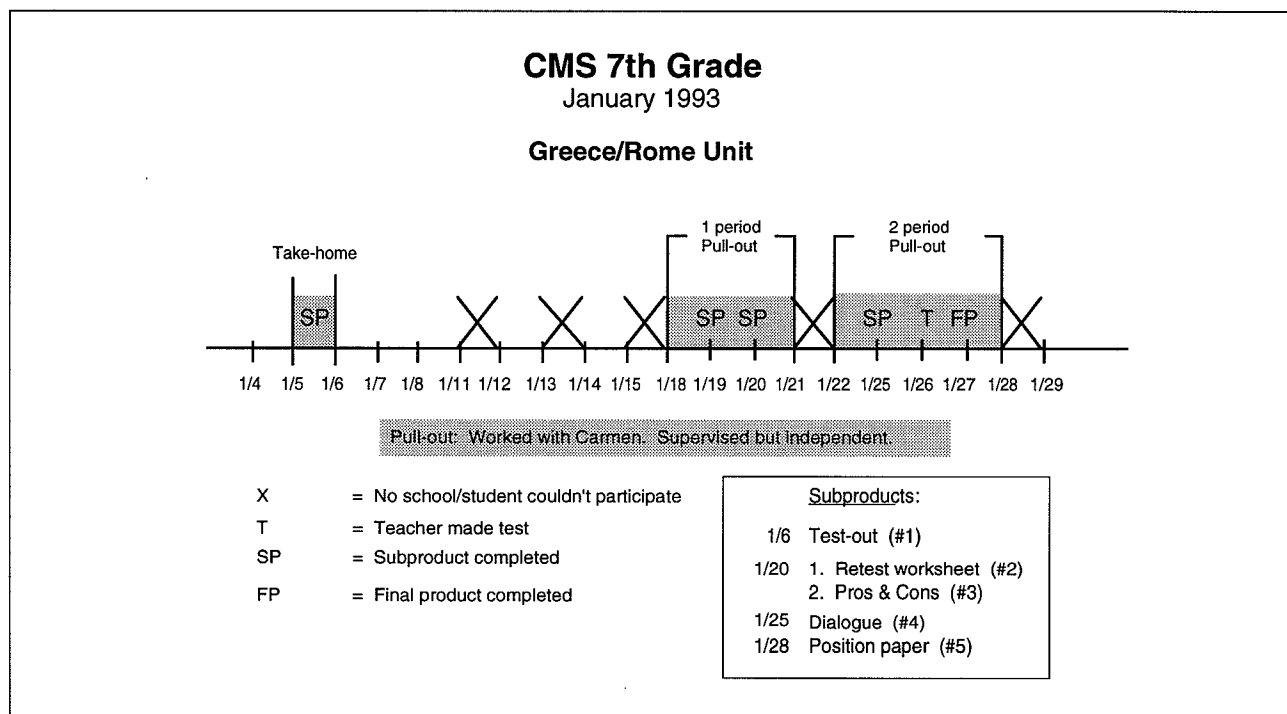


Figure 6. Timeline for Phase I.

Carmen and Josh reviewed and completed portions of the worksheet together. The ISC was developed during this process and was accompanied by a discussion of how the plan would be carried out. This first project had two parts: a *summarization* of the key concepts of *democracy* and *republic*, and then a choice between them (*evaluation* of *democracy* and *republic*). The ISC was supplemented with a written task analysis of what Josh was assigned to do and yielded daily activities, or subproducts, in which the learning task of *summarization* was used. This task included:

1. Defining the key concepts and principles as designated by the teacher (1/5-1/6).
2. Redoing the test-out for key concepts to be targeted during the project (1/20).
3. Listing the pros and cons of both a *democracy* and a *republic* (1/20).
4. Taking the regular social studies test with the rest of the class.

An essay was designed to reflect *evaluation* responding, although the student had not yet been presented with the criteria for judging any of the intellectual operations. While Josh met all of the criteria and scored a 5, the design of the project seemed unfeasible. That is, he spent two full class periods researching the topic and adding factual details, but could have discussed the issue and scored a 5 with his arguments in much less time. Furthermore, his focus was centered on facts, rather than the concepts Carmen and his social studies teacher were trying to promote. From this standpoint, a new intervention was developed.

The second phase of this project drew from concepts being taught about the *Middle Ages*. For this phase the student was pulled away from the content classroom 2 days each week for 2 weeks so he could complete a short essay that utilized a specific intellectual operation. For this project, Josh completed the following:

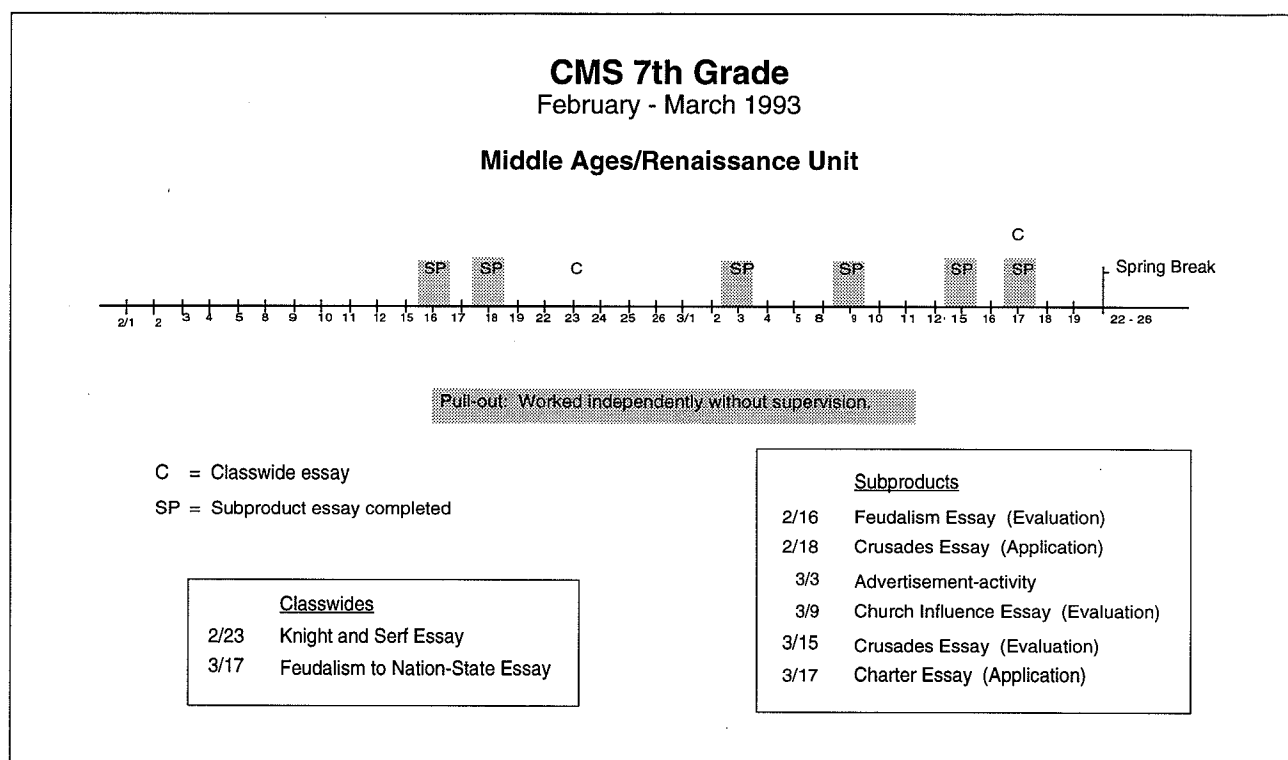


Figure 7. Timeline for Phase 2.

1. *Evaluation* essay on the concept of *feudalism* on (2/16).
2. *Application* essay on the concept of the *crusades* on (2/18).
3. *Evaluation* essay on the concept of *feudalism* (2/23).

Josh was not taught the scoring criteria for these sub product essays, which took him about 20 minutes each to complete. Again, he worked independently.

Pulling the student away 2 days a week and measuring his use of concepts and intellectual operations in discrete essays was more effective than the previous phase of the intervention. As the class continued to study the Middle Ages, Josh was assigned to another 2-week project from the previous ISC; the sub- and final products, therefore, changed to new concepts, including *feudalism*, *crusades*, and *nation states*. As in the first ISC, a number of sub-product essays were generated, with the intellectual operations reflecting both *evaluation* and *application* responses. Throughout the seven weeks, Josh was ostensibly taught by the content teacher, but also he was pulled away weekly to work with Carmen. In this unit, Josh completed two essays that had been administered to the entire class, providing a classwide (normative) reference to double check whether he was in the appropriate instructional level.

This ISC lasted most of 2 months during which time he completed four sub-product essays (taking 20 minutes each), as noted below:

1. Creative activity (3/3).
2. *Evaluation* essay on *church influence* (3/9).
3. *Evaluation* essay on *crusades* (3/15).
4. *Application* essay on *charters* (3/17), *feudalism*, and *nation-states*.

Finally, the class began preparation for a school project fair, during which time the social studies teacher did not introduce new curriculum material. Josh's student teacher, therefore, decided to use this time to review previously taught concepts and to administer an essay every two weeks to the class. Furthermore, another intervention style was implemented. For this phase, Josh was pulled out of his content class and taught particular knowledge forms through 5-10 minutes of explanation, modeling, and testing before each essay administration.

His essay writing was independent but supervised and included the following:

1. Review of the concept of *dictatorship*, an *evaluation* essay (4/22).
2. *Evaluation* essay on *natural resources* was completed with his class (4/26).

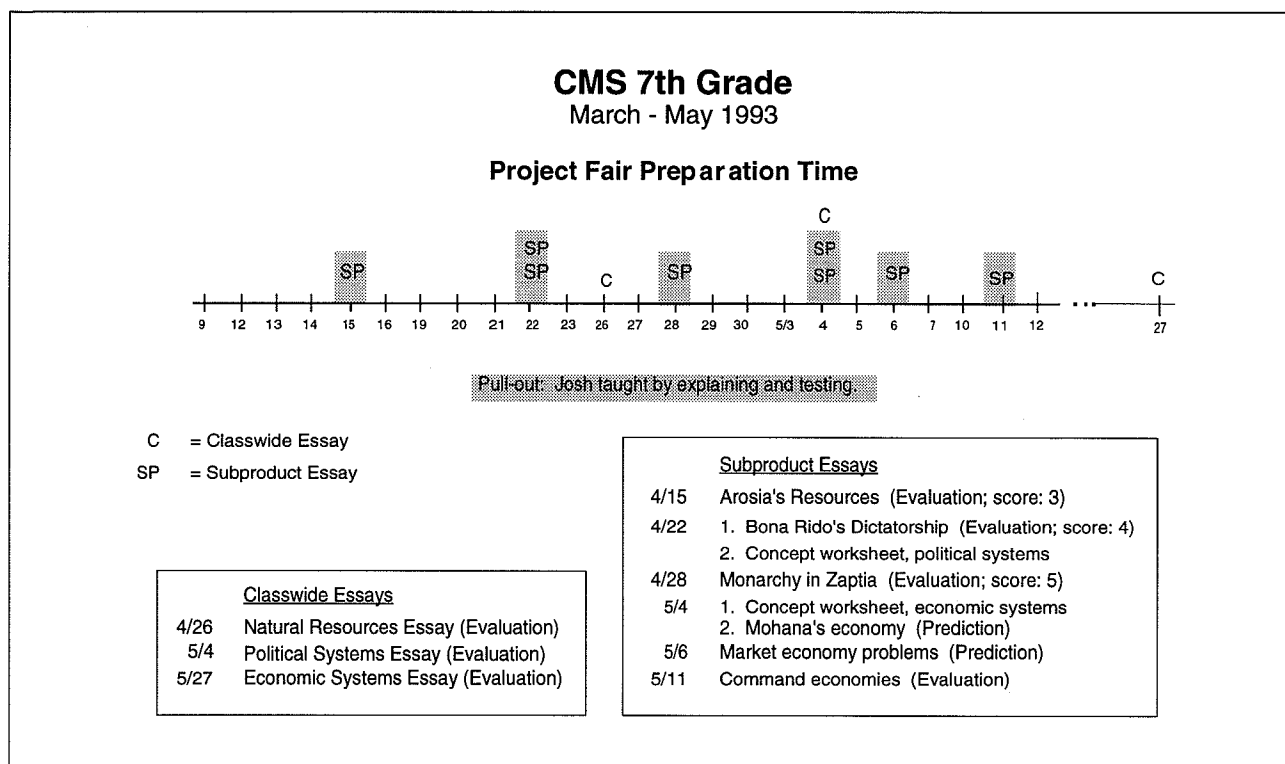


Figure 8. Timeline for Phase 3.

3. *Evaluation* essay on *monarchy* (4/28).
4. *Prediction* essay on *traditional economies* (5/4).
5. *Evaluation* essay on *political systems* was completed with his class (5/4).
6. *Prediction* essay on *market economies* (5/6).
7. *Evaluation* essay on *command economies* (5/11).
8. *Evaluation* essay on *economic systems* was done with his class (5/27).

Most of these sub-product essays were implemented at relatively equal intervals throughout the 6 weeks. The concepts, of course changed to reflect the new content; nevertheless, the intellectual operations remained oriented to *evaluation*. One prediction essay was developed near the end of the unit. Again, as in the previous unit, classwide essays were implemented to ascertain the comparability of Josh's performance to those of his peers. Because TAG services demand that instructional level be appropriate, such a comparison can determine whether the TAG student's response is so far superior that s/he really belongs in different content or whether performance is comparable enough to maintain pace with the rest of the class, varying only the manner in which this content is interactively learned (used within specific intellectual operations).

Criterion-Referenced Analysis of Essays for Evaluation Responses

Given the timelines noted and the content covered, Josh completed a number essays, some of them viewed

as sub-products (allowing him to engage in Interactive Learning Tasks in which the teacher wanted him to become proficient) and some of them viewed as final products (produced at the end of a content unit and essentially representing performance on the terminal task). From a criterion-referenced perspective, all these products can be analyzed in reference to two specific dimensions: the domain of learning and the standard for performance.

The domain represents the types of skill and knowledge the student is learning. In our model, the domain is response defined: It is the intellectual operation in which the student expresses a response. We use mostly *evaluation* operations or learning tasks, but also include an *application* and *prediction* response. In our model, the organized knowledge forms don't represent the domain (which is counter to most views in which the domain is defined by the content being learned). Because we are focused on rate of learning, comparability in the domain has to be attained, otherwise change over time cannot be ascertained. Therefore, in the student samples that follow, the criteria for judgment are oriented around the manner in which intellectual operations are incorporated (i.e., how well developed is the argument within an *evaluation* response—making a decision and then supporting it?). The content (concepts) simply provide the medium for expressing that intellectual operation and, at best, delimit or qualify the domain.

Name <u>Josh</u>	Mrs. Content Teacher 1/6 - 1/28	Individual Student Contract	Knowledge Forms: Summary Evaluation	Beginning Date <u>1/6/92</u> Completing Date <u>1/28/92</u>
Project/Problem Statement (see attachments)				
<p>This is a two-part project. The attached pages will give you many more details. The first part will be a written argument between a Greek and a Roman who believe their own government is better than the other. You will have to think about and write the arguments each would make. In the second part of the project you will choose which form of government you think is better, and explain why.</p>				
Intermediate Steps (Checks) <div style="background-color: #f0f0f0; padding: 10px; border: 1px solid black;"> <p style="text-align: center;">SEE OTHER PAGES</p> <ol style="list-style-type: none"> 1) Completed worksheet (1/19). 2) Notes on strengths and weaknesses of both the democracy and republic systems. 3) Outline of each man's argument. 4) Completion of play. 5) One or two page essay defending the system you think is better. </div>	Materials/Resources Needed <div style="background-color: #f0f0f0; padding: 10px; border: 1px solid black;"> Textbook </div> <div style="display: flex; justify-content: space-between; margin-top: 20px;"> <div style="width: 45%;"> Location(s) TAG Room </div> <div style="width: 45%;"> Support Staff Carmen </div> </div>			
Mini-Schedule (next 2 weeks)				
MON	1/19 TUE	1/20 WED	1/21 THU	1/22 FRI
<div style="background-color: #f0f0f0; padding: 5px; border: 1px solid black;"> <div style="text-align: center; font-size: 2em;">X</div> </div>	4th period in AG room - worksheet - start reading	notes on Athens - read about Home	notes on Rome - start arguments	finish arguments - start play
<div style="display: flex; justify-content: space-between; width: 100%;"> 1/25 1/26 </div>				
- finish play - do opinion paper	- finish opinion paper - artwork			
<div style="display: flex; justify-content: space-between;"> <div style="width: 45%;"> Final Product <div style="background-color: #f0f0f0; padding: 10px; border: 1px solid black;"> 1 - 2 page dialogue between Imagines and CIPHERUS and 1 - 2 page essay on your opinion of which is better. </div> </div> <div style="width: 45%; text-align: right;"> <div style="text-align: center; margin-bottom: 10px;"> Student Signature _____ </div> <div style="text-align: center;"> I understand that I am responsible for completing this project by the date noted above. </div> </div> </div>				

Figure 9. The first Individual Student Contract (ISC).

The standards for evaluating response are widely debated in the field, with no absolute agreement among measurement professionals. Therefore, we have approached this issue from two perspectives. Because the content teacher is viewed as the expert, initial standards can be forged from their "clinical judgments." By simply transcribing the student's work and analyzing it for specific strengths and weaknesses, we can develop "bootstrap" standards within a criterion-referenced perspective. However, because the criteria for this approach are so ill-defined (and difficult to com-

Cipherus was seated in an atrium, waiting for his friend Imagines. Although the men have very different ideas about how a government should be run, they enjoy debating together about it, measuring the pros and cons of each. Ah! Imagines has arrived. He sits down on the bench next to Cipherus, and the two men greet each other.

IMAGINES: Welcome, my friend. What shall we discuss today?

CIPHERUS: We have not yet shared our views on our very different governments.

Choose the form of government [democracy or republic] you think is best and write a one to two page paper defending it. Take the pros and cons you outlined before for that system, and explain why they pros are strong--and how the weak points, or cons, can be fixed or minimized.

I think democracy is better than most forms of governments because the people of the democracy get to choose what they want. They get to choose what laws, presidents, vice presidents, and other government figures they want to help control the country. Some other forms of governments have total control over Land, buildings, schools, and newspapers. This kind of government is called a totalitarianism this means the government has total control. In democratic systems they have constitutions that guarantee certain rights. Their government can't deny them these rights.

The word democracy comes from the Greek words demos means "the people" and kratos which means "rule". People in America have the right to vote, speak out, and join political groups. In a big democratic government there might be smaller governments called state governments. State governments control their own schools, highways, and law enforcements. The United States government is divided into three separate branches executive, legislative, and the judicial. Each of these branches have different responsibilities. The executive branch enforces the nation's laws. The legislative branch is responsible for lawmaking. The judicial decides whether laws are fair or not.

Anyone over 18 is allowed to vote in democracies long ago only male citizens were allowed to vote. Unlike a dictatorship a democracy won't let one person get too much power.

Criterion-Referenced Evaluation: Democracy

Number of Words Written	218
Number of Concepts Used	3
Evaluation Rating	5

municate to non-experts), we also include a subsequent, norm-referenced perspective, in which classwide performance from other students is used to interpret a student's response. In this latter approach, a range of responses is judged with range-finder essays used to develop and calibrate a (norm-referenced) scale.

The First ISC

In this first essay, the content material from the classroom focused on Greece and Rome. The teacher used a standard curriculum text but had reviewed it for specific concepts. Of course, with this content, the concepts of *democracy* and *republic* appeared quite prominently, so the content teacher advised the student teacher to focus the project on these concepts. A number of sub-products were developed to help Josh become proficient in evaluating these concepts; the final product was in response to the following prompt:

IMAGINES: Well, I think that my government is better because....

YOU FINISH WRITING THE DIALOGUE BETWEEN THE TWO MEN.

Save all of your notes; you will be asked for them later.

The entire ISC is depicted in the figure on the previous page. The major components described earlier reflect a problem statement centered around two concepts, a series of activities (and sub-product essays), a final product for the student to complete so he can obtain practice on the intellectual operation of *evaluation*, and finally, the requisite materials and schedule needed to complete the contract.

At the end of the unit, a final product essay was administered solely to Josh, not the entire class (see above). It was used to sample his early use of the

intellectual operations and was a *summarization* essay on the unit. The prompt also included an *evaluation* question, as it required the student to make a choice and defend it

Instead of creating a rationale, as his student teacher had hoped, Josh reiterated some reasons and supporting facts from his book almost verbatim. His arguments were implicit within his listed facts, rather than explicit statements. Although the essay achieved a score of 5 due to its use of multiple arguments for the choice and against the non choice, his reasoning was neither original nor well-developed.

The Second ISC

In the second ISC, the approach was the same: (a) a problem statement was developed, (b) activities were planned and implemented, and (c) final products were completed, some of which also were given to the entire class. Josh was not pulled away from class very often (only 3 times); most instruction came from the content teacher, with the student teacher administering the sub- and final-product essays.

The second essay, which was classwide, focused on concepts the teacher identified from a unit on the *Middle Ages*. In this assessment, all students completed an

The Knight and the Serf

This is a story about a small kingdom in Central Europe during the Middle Ages. The vassal who rules this small kingdom is Lord Martin. Lord Martin is a mean man. Everyone has to pay him very high taxes. All the serfs in the kingdom work very hard tending Lord Martin's crops, but Lord Martin doesn't take good care of them when they need help. He sends his knights off to fight battles, but he doesn't give them enough swords. Nobody likes Lord Martin the vassal.

Two people who live in the kingdom have decided they can't take it any more. One of these people is a knight named Sir John. The other person is a serf named Harold. Sir John and Harold have decided they do not want to be loyal to their vassal any longer. They will not pay Lord Martin anymore high taxes. Harold the serf will not work hard in the fields and Sir John the knight will not fight anymore battles.

Here are two things that could happen if Sir John and Harold stop being loyal to Lord Martin. Place an X beside the statement you think is most accurate.

___ Life will be easier for Sir John the knight than it will be for Harold the serf.

___ Life will be easier for Harold the serf than it will be for Sir John the knight.

Tell why you made your decision. Tell what information you used to make your decision. If you think things will be easier for the serf than for the knight, tell why. If you think things will be easier for the knight than for the serf, tell why.

I think things would be easier for the knight instead of the serf because the knight is in a higher class than the serf. I made my decision because I think the serf won't be listened to because he is at the lowest point in society and the knight might be respected for helping win a battle saving the town or village from villains. The knight is helping the king with his services so I think the king will lower the taxes for knights and higher the taxes on serfs and take some of the money and buy the knights better weapons and armor.

Criterion-Referenced Evaluation: Knight and Serf

Number of Words Written	104
Number of Concepts Used	3
Evaluation Rating	5

The ISC for this unit on the Middle Ages is depicted in Figure 10 (next page).

Name <u>Josh</u>	Individual Student Contract	Beginning Date <u>2/16</u> Completing Date <u>3/17</u>
Mrs. Content Teacher 2/16 - 3/17	Project/Problem Statement (see attachments)	
Josh will be writing a collection of short essays covering various concepts and principles, and using a variety of intellectual operations. Part of project completed in tandem with another student. Content: Middle Ages concepts of feudalism, crusades, and charters-nation states.		
Intermediate Steps (Checks) 1. Evaluation essay on concept of feudalism. II. Application essay on concept of crusades. III. Church/crusades advertisement. IV. Evaluation/prediction essay on concept of church of church influence. V. Evaluation essay on concept of crusades VI. Application essay on concept of charters.	Materials/Resources Needed <div style="border: 1px solid black; padding: 5px; margin: 5px 0;"> Location(s) TAG Room </div> <div style="border: 1px solid black; padding: 5px; margin: 5px 0;"> Support Staff Carmen </div>	
Mini-Schedule (next 2 weeks)		
MON	TUE	WED
2/16	2/18	2/18
FRI		
Pull-out Feudalism Essay	Pull-out Crusades Essay	Pull-out Cathedral Essay
2/23	3/4	
Final Product Classwide assessments administered. 1) Knight & Serf (Evaluation) 2) March 17 Feudalism to Nation-States (Application)		
I understand that I am responsible for completing this project by the date noted above. <div style="display: flex; justify-content: space-between;"> <div style="width: 45%;"> Student Signature _____ </div> <div style="width: 45%; text-align: right;"> </div> </div>		

Figure 10. The second Individual Student Contract (ISC).

evaluation essay focused on the concept of *feudalism*. Although Josh had not been taught the scoring criteria, he performed well. He applied the concepts of *feudalism* and *chivalry*, and also referred to the *social class system*, a concept that had been presented earlier.

Because the class studied the Middle Ages for more than one month, another classwide assessment was administered, which utilized an *application* essay incorporating the concept of *feudalism*. The prompt and Josh's response follow:

the intellectual operations in the sub- and final products. The last set of essays, therefore, reflected the teacher's interest in consolidating her students' understanding of concepts learned earlier with *governmental* and *economic systems*. On the page 31, the final products are mostly *evaluation* responses, with one *prediction* response.

The remainder of this section presents two essays that were given to Josh during the last part of the year; they also were classwide essays and, as *evaluation* tasks,

Feudalism to Nation-States

Imagine you are a middle-class merchant in the late Middle Ages. You support the government of your nation-state because you need peace and a strong government so you and other merchants can safely carry on your business of buying and selling. You belong to a group or guild that represents other merchants and trades-people like tailors, potters, weavers and carpenters. You are very interested in protecting the rights of the townspeople where you do business.

Your nation-state has a strong government where the ruler has to share power with groups that represent the people. This is very different from the system of government of the early Middle Ages, which was known as Feudalism. You are explaining to your young apprentice how people lived and worked under the feudal system. He wants to know what happened to make things the way they are now in your town. Explain what you think happened to cause the growth of nation-states in the late Middle Ages.

I think when there was a feudal system there wasn't much opportunity for people to start their own businesses. So that's when I think guilds started forming in order for people to work independently. I also think that after the guilds were formed they grew into small towns to eliminate other competition but after more people started

Criterion-Referenced Evaluation: Feudalism

Number of Words Written	56
Number of Concepts Used	1
Evaluation Rating	2

Josh linked his answer to one of the concepts taught in class (*guilds*) while others in his class had not done so. He clearly stated an antecedent condition and then offered a progression of logical events. However, he did not develop the reasoning behind the events and stopped the essay short, leaving him with a lower score than initially indicated by his strong start.

The Third ISC

The third and final ISC was really an extension of the materials that had been taught earlier throughout the year. The major focus at this time was expanding

offered a choice between a *monarchy* and a *democracy*. At this point, Josh was familiar with the criteria and had practiced them during the pull away interventions. He carefully satisfied all of the scoring criteria, discussing two reasons for having a *monarchy* and two reasons against having a *democracy*.

Research, Consultation, & Teaching Program

Figure 11. The third Individual Student Contract (ISC).

Threvnia's New Political System

Your country, Threvnia, has been under the rule of a harsh dictator for many years. No one has had freedom of speech or the ability to vote for a new leader. Nor has the economy been good; few people have jobs and the quality of life is very low. Now, the United States has intervened in Threvnia, and has removed the dictator from office. The U.S. president wishes to give Threvnians a choice in a new kind of government.

A monarchy can be established, where a well-liked ruler shares some power with government branches. Because the king or queen would rule for life, stability and probably peace will be guaranteed. However, most of the decision-making power would lie with the king instead of the people.

Another option is democracy, in which everyone votes for a new president every few years and there is freedom of speech and political party membership. This might potentially cause instability because no one president could be in office long enough to carry through long-term changes.

Pretend you are a Threvnian citizen and can cast a ballot for the kind of government you think will be best for your country. Place an X next to your choice below.

☐ Monarchy

☐ Democracy

Now, write an essay which tells why you made your decision. When answering this, consider these aspects of a democracy: people vote to enact laws, people have their freedom protected by law, and people choose representatives who promote their point of view. Tell what information you used, and why you chose a monarchy or a democracy. If you chose a monarchy, explain why. If you chose a democracy, explain why.

I think a monarchy would be best for the country because the monarchy would have time to carry out long term plans and ideas. If there was a democracy only small things would get done and probably no long term arrangements would be made. A monarchy wouldn't have to worry about people liking him so he would have the most time to do his work unlike a president who might waste a lot of time trying to make everyone happy and not much gets done. With a monarchy things would go through a steady change but with a democracy after a few years everything is changed and only a few things remain the same.

Criterion-Referenced Evaluation: Threvnia.

Number of Words Written	112
Number of Concepts Used	2
Evaluation Rating	5

The final classwide assessment was an *evaluation* essay concerning *socialist economies* and *capitalist economies* (next page).

Yin and Yang

The countries of Yin and Yang both have a democratic political system where people help rule by voting on laws and electing representatives. However, Yin is a socialist society while Yang is a capitalistic society.

Yin's socialist society is one where all the people share in the work and products made. Everyone lives comfortably, but no one individual has any great wealth.

On the other hand, in Yang's capitalistic society, businesses are privately owned and operated for individual profit. Some people are very wealthy, but many are also poor; there are few in the middle class.

Both countries are developing and some citizens are dissatisfied and hope for changes in the economy.

Think about which aspects of socialism make it better or worse than capitalism. Then, decide which kind of country - with a socialist democracy or capitalistic democracy - you would choose to live in. Mark an X next to your choice below.

_____ Democracy with a capitalistic society

_____ Democracy with a socialist society

Now, explain your choice in the space below. If you chose a democracy with a capitalistic society, explain why. If you chose a democracy with a socialist society, explain why.

I chose a Democracy with a socialist society because in a socialist society everyone would be equal. There would probably be less fights and arguments in a socialistic society because no one could say they were better than someone else. In a capitalism society many people die because they can't get enough healthcare, food, or shelter. There is usually more violence in a capitalism because if some people don't have enough things to survive they turn to drug dealing, prostitution, stealing, and other ways to make money. In a capitalism society inflation might occur and cause a depression unlike in a socialism society. Being treated equally is one big reason to live in a socialism.

Criterion-Referenced Evaluation: Yin & Yang

Number of Words Written	113
Number of Concepts Used	3
Evaluation Rating	4

Josh received a score of 4 because, while he made several succinct arguments against the non-choice, he did not make more than one differentiated argument for his choice. As prompted, he mentions 3 concepts from this unit.

Criterion-Referenced Analysis of Essays for Different Intellectual Operations

While the essays depicted above were primarily classwide, a number of other essays had been given to Josh as part of the sub-products within the ISCs. As mentioned earlier, a criterion-referenced perspective implies specific intellectual operations; most of the earlier essays had been *evaluation* responses. In this section, we present some essays given only to Josh and reflecting different intellectual operations. The central question is whether Josh was learning broad thinking skills (transfer across the intellectual operations) or

whether his performance is highly specific to certain intellectual operations (*evaluation* in this case).

During the second phase of the intervention, Josh had been given one essay 2 days a week. These essays included *prediction*, *evaluation*, and *application* prompts. After each administration, he received immediate feedback; his score improved when the same intellectual operation (in this case, *evaluation*) was re-administered. Although his first evaluation essay did not include multiple arguments both for the choice made and against the choice not made, these later essays were broadly based, reflecting his learning the criteria, and he achieved a score of 5.

Application Essay

The next essay assignment given to Josh while his class was immersed in the Middle Ages material used the *application* prompt below, which is followed by his response:

Directions: Read the question below; write a two or three paragraph answer which describes your reasoning. You may use your book.

Question: You are a merchant with a spice store in the township of Chestwick in the late Middle Ages. Many of the merchants in Chestwick have seen how well a charter has helped the people in a neighboring township, and want to try this system of self-governing. What happened to cause this movement toward writing charters in the late Middle Ages?

If I were a merchant I would like a charter because it allowed the merchants to control their own land and town. In the beginning when the lords owned the land they could force the merchants to do what they wanted them to do. With a charter the team was more organized and more got done. The charters started to come in place when the merchants started to grow more powerful because masters of wealth. When the merchants become more powerful they forced the lords to give them their own land.

Criterion-Referenced Evaluation: Chestwick

Number of Words Written	91
Number of Concepts Used	2
Evaluation Rating	2

While this assignment was given to Josh before he was taught the scoring criteria, he did outline (in the second half of the paragraph) a chain of events that brought about the use of *charters*. He was given a score of 2 because, while clearly stating an antecedent event or condition, he did not support it with more than one reason.

Three Evaluation Essays

Following the *Middle Ages* unit, the class launched into preparation for a project fair. Therefore, a series of *evaluation* essays was administered utilizing earlier concepts. The prompt and response were as follows:

You are a government official for Arosia, an agricultural country. While Arosia has plenty of food and water to feed people, it is also facing some big problems.

One problem is that Arosia is quickly running out of timber. This is causing trouble because the population is growing quickly, and timber is needed to build new homes. Another problem facing Arosia is the lack of oil. Without oil, the country cannot fuel factories. Arosia wants more factories so that they can manufacture products and make more money in the world market than they can by just selling farm produce.

Arosia has enough farm produce to trade with other countries for either timber or oil, but not both. Should Arosia trade for timber or oil? Place an X beside your choice:

_____ Arosia should trade its farm produce for timber.

_____ Arosia should trade its farm produce for oil.

Now, write an essay which tells why you made your decision. Tell what information you used, and how you chose between trading for timber or trading for oil. If you chose to trade for timber, explain why. If you chose to trade for oil, explain why.

I think the trade for oil is more important instead of a trade for timber because the oil can be used to run the factories and this would probably help the economy. If they traded for timber they could build more houses and after awhile the city would have too many people and an overcrowded city where no one would want to live or work. With the abundant oil more companies would open helping the city. With the goods produced by factories the government can trade them for timber. With people moving to Arosia who want to buy houses the price will go up and the company would be making more profit.

Criterion-Referenced Evaluation: Arosia

Number of Words Written	112
Number of Concepts Used	2
Evaluation Rating	3

Because, at this point, Josh was becoming familiar with formally writing an *evaluation* response, he made clear attempts to fulfill the criteria asked of him. He included an explicitly stated choice, a rationale, and both an argument for the choice and against the non-choice. However, he then began to refute his own argument against the non choice and added some ideas that were unrelated to the prompt.

The prompt for the second of the three consecutive *evaluation* essays utilized the concept of *dictatorship*; the essay, along with Josh's response, follow:

Josh's score improved over the previous *evaluation* essay. This essay was more complex because there were two non-choices against which arguments had to be made. Again, the choice was clearly stated, and several arguments were made, both for the choice and against one of the two non-choices. The other non-choice had only one argument made against it; because he had included multiple arguments both for the choice and against the others, however, he earned a score of 5. Feedback was given about providing multiple reasons for each choice/non-choice available.

You live in Bona Rido, a poor country in South America. The economy there is very bad and many people do not have jobs. The president is greedy, and takes money and land from the Indians and poor people. The government does not allow people to oppose the president. People are angry with the government because of these problems, and there have been many uprisings.

Suddenly, a man organizes an army and overthrows the government, setting up a military dictatorship. Which of the following do you think is the most important reason for the rise of the dictatorship? Place an X next to the reason you think is most important.

_____ The economy was very bad and most people did not have jobs.

_____ The old president was corrupt and took money and land from the poor people and Indians.

_____ The old government censored the people by not allowing them to publicly oppose the president.

Now, write an essay that tells why you made your decision. Tell what information you used to make your decision. If you think the bad economy was the most important reason for the dictator's takeover, tell why. If you think the president's corruption was the reason, tell why. If you think the government's censorship was the reason, tell why.

I think the reason for the dictatorship takeover was because the president was taking money and land from the poor people and Indians because if the economy was bad the people would probably move but if it was land they owned they would probably stay and fight for it. If the President took all of the land he could kick the people off it and they would have to wander around looking for a new home. If it was because the government wouldn't let the public oppose the president then that would not make sense because the public are the people who are having their stuff taken away and are already aware of the president. If the economy was bad then they would probably move but they didn't have money. So they probably formed the group to take over the country give the land and money back which would strengthen the economy and allow people to have freedom.

Criterion-Referenced Evaluation: Bona Rido.

Number of Words Written	158
Number of Concepts Used	1
Evaluation Rating	5

The final *evaluation* prompt, concerning *monarchy*, and Josh's response are shown below:

You are a famous lawyer living in Zaptia, a small country which is ruled by the monarchy of Queen Bella III. Queen Bella is consulting you for help because many government officials are angry with her. She has taken an active role in creating new laws with the Parliament, and the officials think she is just a figurehead who shouldn't be able to make new laws.

Queen Bella wants your advice about which government she should use to defend her actions. She could tell the officials that because she will be in power for a lifetime, her long range plans will have time to work. She could also tell the officials that it is important for them to let her create laws so that Zaptia appears to be a strong unified country to the rest of the world. Put an X next to the argument you think is best for the Queen to use:

_____ Queen Bella's long-range plans will have time to work.

_____ Zaptia needs to maintain an image of unification.

Now, write an essay that tells why you made your decision. Tell what information you used to make your decision. If you think having enough time to carry out long-range plans is more important, tell why. If you think Zaptia's world image is more important, tell why.

I made my decision the way I did because I think Queen Bella's long range plans are more important than to maintain a good image because if Queen Bella's people here that she is making long term plans the younger people will probably support her. It isn't that important to have a good image if you have a strong army because if other countries think your weak and you aren't it doesn't matter what you look like on the outside. Queen Bella's long term plans would strengthen the country and help it. Showing that it is strong on the outside so no one attacks it doesn't do much for the country. Queen Bella's long term plans would probably attract people to the country in the future.

Criterion-Referenced Evaluation: Zapita.

Number of Words Written	126
Number of Concepts Used	1
Evaluation Rating	5

This essay showed a balance of coverage, with multiple arguments both for the choice and against the non choice.

Prediction Essay

Given that Josh's scores were improving with respect to *evaluation* responses, he was told the basic

criteria for a *prediction* essay and asked to write one. The prompt and his essay are shown below.

Josh implicitly applied several concepts in this essay: *Democracy* was used to explain why the Sirrites have a choice in what happens; a relationship was noted between *gas* and a country's *power* relative to

The nation of Sirris has a market economy; its people have a choice of goods and services, and their demand for particular products increases the supply of those items.

One popular product, car gasoline, has been outlawed by the government because it has been found to be hazardous to people's health. Cars are important to the citizens of Sirris, and there is no convenient replacement for the gas.

Predict what the Sirrites will do with this new development. Remember to include the concept of market economy.

I think the Sirrites will be able to buy gas because if there is a market economy there is probably be a democracy. If there is no gas to be used it would make motion weak because it could be attacked and would not have no way to escape. If the Nation doesn't have gas the whole economy would fall apart and could easily be turned into a traditional economy dictatorship, or Oligarchy. I don't think the people will agree with the government because the gas is so important the people will probably rebel. I also think that if the Nation doesn't change the law the whole nation would probably move to some other country, and that would totally ruin the economy.

Criterion-Referenced Evaluation: Sirris

Number of Words Written	122
Number of Concepts Used	5
Evaluation Rating	1

international affairs; reference also was made to the concepts of *traditional economy*, *dictatorship*, and *oligarchy*. The essay, however, needed one clear outcome and two arguments to be stated.

Placement in Instructional Level: Norm-Referenced Perspectives

At this point, it is only possible to broadly describe Josh's performance, noting specific strengths and weaknesses in the intellectual operations reflected in his essays. If his teacher had to defend his instructional placement, arguments could be generated about how well he does on some things (e.g., uses specific concepts in providing competing arguments) and how much he needs to work on other things (e.g., being more broadly focused on linking multiple reasons that move well beyond just facts). Certainly, Josh appears to be succeeding and may be considered appropriately placed within an instructional level. Within his pull-away sessions, he is learning new information, with the content being adjusted both vertically (i.e., new material is being presented) and horizontally (i.e., he is revisiting old information from new vantages). And we have his actual work samples to review.

Even with this assessment information, however, we cannot be certain that he is definitively placed at an appropriate instructional level. Although he is succeeding, it may be possible that the pace of instruction could be much greater in both depth and intensity. Maybe Josh could be covering even more vertical and horizontal movement within the content area. Without some supplemental information, we cannot resolve this issue. Therefore, in this section, we present the logic and data for justifying and double-checking placement in the appropriate instructional level from a norm-referenced perspective.

We base our view on the categorical nature of the label, Talented and Gifted. If a student has been correctly identified as TAG, we generally infer that "intellectual ability" is greater than most peers (the student is performing within a very small percentage of the distribution of age-grade peers). This label does not imply that all learning in all content areas will be any different than these peers (i.e., faster, deeper, broader, etc., however any of these dimensions are measured). It only implies that educators must be on guard and ensure that the potential for "enhanced learning" may be present. And, when it occurs, instruction needs to be adjusted. This comparison to peers, therefore, is an important and essential component in the definition of TAG. By building on this concept, we can extend the logic to also address instructional level, in addition to its use for eligibility and categorical labeling.

In this section, we compare Josh's performance to a sample of his classwide peers. On several occasions throughout the year, essays were generated and administered to all the students in the class. Recall that initially the pull-away program for Josh was oriented around learning new information in the content class from a self-discovery view. Although the content teacher taught all the students (including Josh), the core concepts from successive units (on Greece and Rome, the Middle Ages, and governmental and economic systems), Josh was provided additional time and opportunity to express these concepts in a problem-solving task. Later, he also was provided more structured instruction in how to accomplish this outcome.

If this program is effective in maintaining him at an appropriate instructional level, his performance should be comparable to his peers. If the program is excelling him toward a faster trajectory (he is learning more information in greater depth and with greater breadth), then his performance should be considerably greater (or at least qualitatively different) than his peers. If he is being moved too fast in this trajectory, his performance should be equally different but in a negative manner.

In these essays, Josh's performance is compared to a sub-sample of his peers. Although not all the students took all the essays, they were taken by the full range of students from the classroom: some students in the middle of the group and some students near the bottom (who were receiving special education support). Of the following essay responses, can you tell which one is by Josh?

The task of identifying Josh's essay from the mix of other students in his class is clearly difficult. Two questions immediately come to mind:

1. On what basis (or with what criteria) should identification be made?
2. How different does performance have to be to warrant different instructional placement options?

To address these two questions, both of which rest upon a norm-referenced perspective, we have established a rating scale to qualitatively judge performance on the same yard stick and compare Josh's performance with others in the class. Following are some of the guidelines we used to develop the scale; the actual judgment flowchart is then presented.

In these criteria, we assume that the prompt used to elicit student writing directly affects the quality of students' responses. For example, in making *evaluations*, students must choose from among at least two options (although more may be presented) and then present a rationale for making the choice. If the choices are clearly presented in the prompt, the task is drastically different than if the student is expected to both

The Knight and the Serf

This is a story about a small kingdom in Central Europe during the Middle Ages. The vassal who rules this small kingdom is Lord Martin. Lord Martin is a mean man. Everyone has to pay him very high taxes. All the serfs in the kingdom work very hard tending Lord Martin's crops, but Lord Martin doesn't take good care of them when they need help. He sends his knights off to fight battles, but he doesn't give them enough swords. Nobody likes Lord Martin the vassal.

Two people who live in the kingdom have decided they can't take it any more. One of these people is a knight named Sir John. The other person is a serf named Harold. Sir John and Harold have decided they do not want to be loyal to their vassal any longer. They will not pay Lord Martin anymore high taxes. Harold the serf will not work hard in the fields and Sir John the knight will not fight anymore battles.

Here are two things that could happen if Sir John and Harold stop being loyal to Lord Martin. Place an X beside the statement you think is most accurate.

_____ Life will be easier for Sir John the knight than it will be for Harold the serf.

_____ Life will be easier for Harold the serf than it will be for Sir John the knight.

Tell why you made your decision. Tell what information you used to make your decision. If you think things will be easier for the serf than for the knight, tell why. If you think things will be easier for the knight than for the serf, tell why.

4003

I think Life will be easier for Harold the John because the lord must have a lot of serfd and not notice one working. John is a night and MMartin probaly doesn't have alot of knights and will notice one not working. So my conclusion is Harold will not be noticed working like, John will be noticedreallyquick.

4009

Things will be easier for Sir John the knight than it will be for Harold the serf. To explain this we have to go back a ways. Vassals started having rule over others when the vikings came. Viking's attcked villages without mercy. To remedy this problem, serfs were givin protection from the vassles in return for labor. Knights were recruited and swore allegiance to the vassles; they protected the comon people and fought in wars with the vikings.

The knights therefore knew how to handle battle. In this way, if Sir John were to go out alone, he would be safer than the serf. Serfs knew no warfare, and Harold would probably not know warfare since he was one.

4012

I think things would be easier for the knight instead of the serf because the knight is in a higher class than the serf. I made my decision because I think the serf won't be listened to because he is at the lowest point in society and the knight might be respected for helping win a battle saving the town or village from villains. The knight is helping the king with his services so I think the king will lower the taxes for knights and higher the taxes on serfs and take some of the money and buy the knights better weapons and armor.

4013

I think sir Johns life will be easier Because people respect him. And he has his own land. And Harold the serf is just training so the king could keep him as a serf and never let him become a knite and the king will fire him if he wouldn't serve the king any more. And sir John will have it easyier because theres not that many battle's he would have to fight but Harold the serf would have to keep serving the king every day.

4014

I think that life will be easier for Harold because he knows how to plant food so he can also take care of himself. He will have an easer life than Sir John will becasue John can't fight to take care of himself if he is weak. John needs food that he can't get.

4018

I think that life will be easier for sr. John, because he might have friends that will protect him if the Lord Martin trys to hrut him but the serf Harod will probably get punished or even kill because he dose not have any oher could probly stick up for him.

I think that the punishmet for sef harold will probably be getting dunk into a pool of water with leaches. He will be dunck 100 time and then go with out food for a weeeek, thethe people will finaly help him and lord martin will be lord nomore.

Feudalism to Nation-States

Imagine you are a middle-class merchant in the late Middle Ages. You support the government of your nation-state because you need peace and a strong government so you and other merchants can safely carry on your business of buying and selling. You belong to a group or guild that represents other merchants and trades-people like tailors, potters, weavers and carpenters. You are very interested in protecting the rights of the townspeople where you do business.

Your nation-state has a strong government where the ruler has to share power with groups that represent the people. This is very different from the system of government of the early Middle Ages, which was known as Feudalism. You are explaining to your young apprentice how people lived and worked under the feudal system. He wants to know what happened to make things the way they are now in your town. Explain what you think happened to cause the growth of nation-states in the late Middle Ages.

4003

The feudal system was about every one owning their loyalty to the king. In return the king gave protection. The king would make up the rules. People began to get angry with the strong protecting the weak the feudal system. Some people decided not to obey by the laws. More and more people began to do nothing. The king said. Then when there was a lot of people began to do nothing the way they wanted. The king was not allowed to make decisions anymore. The people would vote on what to do and what not to do.

4009

Towards the end of the middle ages, the king was the sole ruler. Whatever he said was obeyed without complaint. Under the king, there were noblemen who gave their loyalty to the king. These noblemen were powerful and wealthy. Noblemen gave land and protection to lesser nobles, or vassals. People in the middle ages were afraid of vikings, bandits, and warriors. In return for protection and land, vassals gave loyalty and military aid to the powerful nobles. Vassals usually occupied a manor house and lorded over serfs in a village. Sometimes vassals would build castles for protection when enemies came. Serfs labored for the vassals, or noblemen, and in return the vassals would give protection joined villages that did. The villages over time turned into town, and the towns into cities. This was the end of the middle ages.

4012

I think when there was a feudal system there wasn't much opportunity for people to start their own businesses. So that's when I think guilds started forming in order for people to work independently. I also think that after the guilds were formed they grew into small towns to eliminate other competition but after more people started

4013

The feudal system involved a king noble vassal and serfs. The noble gave protection to the weak in return for loyalty and labor. The feudal system stopped when people wanted to be treated equal and not have to do anything the king told them.

4014

The feudal system is the strong (rich) take care of the weak (poor). The protected them. In return the poor gave loyalty, labor and military service. People traveled in groups to trade outside of the manor because bandits and vikings worked outside the manor. People traveled in groups because they didn't want to get robbed. They traveled this way so that they could trade with other manors. When they were done trading there some people decided to stay in that village of the manor. So those villages grew bigger. Then grew bigger. Then the town was soon a city.

4015

Well you could say this is what happened there is a triangle at the top was the king. The king owed nothing he worked about nothing. After the king was the nobles the nobles are relatives of the king they owe their loyalty to the king and control the vassals and serfs. Next is the vassals the vassals are the knights the knights owe their loyalty to the nobles the vassals protect the serfs from vikings and bandits and warriors. The serfs are the slaves and townspeople. That's the triangle of the feudal system the feudal system is where the strong protect the weak the weak are the people with no money, and pretty soon the vikings and people who stole from the villages became more interested in other things like travel so they didn't have to worry about that much but they still did some so the people started to travel in groups so they could trade and some people wanted to stay in the towns so they grew larger and turned to towns and then cities.

4018

We went from a city or a feudal system where the king ruled everything to a self government because the people did not like listening to the king telling them what to do so the people got together and decided that everyone was equal and that they wanted to make their own government with their own rules that even the king would have to follow by. The people even started their own army that if a person did not have to be in it he would not have to he could do something else instead like make clothes, or pots. The people also decided they could do whatever job that they wanted to. I think that they also start a school show that more people could get a better education besides the church. So that's why they made a government.

Threvnia's New Political System

Your country, Threvnia, has been under the rule of a harsh dictator for many years. No one has had freedom of speech or the ability to vote for a new leader. Nor has the economy been good; few people have jobs and the quality of life is very low. Now, the United States has intervened in Threvnia, and has removed the dictator from office. The U.S. president wishes to give Threvnians a choice in a new kind of government.

A monarchy can be established, where a well-liked ruler shares some power with government branches. Because the king or queen would rule for life, stability and probably peace will be guaranteed. However, most of the decision-making power would lie with the king instead of the people.

Another option is democracy, in which everyone votes for a new president every few years and there is freedom of speech and political party membership. This might potentially cause instability because no one president could be in office long enough to carry through long-term changes.

Pretend you are a Threvnian citizen and can cast a ballot for the kind of government you think will be best for your country. Place an X next to your choice below.

_____ Monarchy

_____ Democracy

Now, write an essay which tells why you made your decision. When answering this, consider these aspects of a democracy: people vote to enact laws, people have their freedom protected by law, and people choose representatives who promote their point of view. Tell what information you used, and why you chose a monarchy or a democracy. If you chose a monarchy, explain why. If you chose a democracy, explain why.

4003

I think that we would be better off with a Democracy Because, If we have a bad leader we will vote him out. We have the right to decide on what we want. There will still be rules but, we have the right's for freedom of speech, do what we want and, to chew watermellen gum. If we had a Monarchy we would not have those things. We would have to do it his or her way. they would make decions that not all of us agree on. We do have rights to do what we want and when want. We are not slaves. We do have rights!

4009

I chose democracy as the political system my country, Threvnia should have. Not only will a democracy give us freedom, which we have never expeireinced, but also give us self-confidence in the decisions we make. A democracy will also help the economy, because with the freedom of press and speech, more people can get political jobs that will arise. Democracy will also help the country of Threvnia by dispersing with class systems. The people of Threvnia will have a chance to succeed and move up in the class system. Monarchy, however, would keep the people at the same class they have been in. The people of Threvnia will never have a choice as to what leader they will have, and in this light, Threvnia will not benefit any more than their last political system.

4012

I think a monarchy would be best for the country because the monarchy would have time to carry out long term plans and ideas. If there was a democracy only small things would get done and probally no long term arrangements would be made. A monarchy wouldn't have to worry about people liking him so he would have the most time to do his work unlike a president who might waste a lot of time trying to make everyone happy and not much gets done. With a monarchy things would go through a steady change but with a democracy after a few years everything is changed and only a few things remain the same.

4013

I have chosen democracy because I think we should be able to work for who we want. if we had a monarchy if we got bad leaders we couldn't do anything to get rid of him. With a democracy we'll have diffrent leaders to make desisions and we could get rid of them if we realy hated them. if we got democracy we could have a chance to speak about what we want and we would have much more freedom that way than if we had a democracy. So I say we should have a democracy.

4014

I chose a dimorcarasey because we can have more freedom. I would like to be free inseed of a rat locked up with out any reason. I would like to be able too choose something, and be able to somewhat make my life the way I want it to be. I would also want my children to not have to go through not being free. I feel that If I was ever given any kind of freedom and then have it taken from me in a haert beet I would die. I want to be able to choose how I want my life to be, how to raise my children the way I want them raised, and to be able to choose the things that I think will be best for my country with other people. I could listen to what other people have to say. If you listen close enough you may even change your mind. And then I could change my vote and vote for what I want then too.

4018

I chose a democracy for my form of gorment because I want freedom of speech. I also want to be able to vote for I think is right, and I also whant to vote agist what I think is rong. I a monarcy you don't get to vote for what you think is right and wroge plus you would get less freedom of speech which I think is very important to help me pick who I want. Besides If I would of chosen a monacy It would have been to much like the past because you would have very little say about what happens around her. So that is why I am chossing a democracy.

Yin and Yang

The countries of Yin and Yang both have a democratic political system where people help rule by voting on laws and electing representatives. However, Yin is a socialist society while Yang is a capitalistic society.

Yin's socialist society is one where all the people share in the work and products made. Everyone lives comfortably, but no one individual has any great wealth.

On the other hand, in Yang's capitalistic society, businesses are privately owned and operated for individual profit. Some people are very wealthy, but many are also poor; there are few in the middle class.

Both countries are developing and some citizens are dissatisfied and hope for changes in the economy.

Think about which aspects of socialism make it better or worse than capitalism. Then, decide which kind of country - with a socialist democracy or capitalistic democracy - you would choose to live in. Mark an X next to your choice below.

_____ Democracy with a capitalistic society

_____ Democracy with a socialist society

Now, explain your choice in the space below. If you chose a democracy with a capitalistic society, explain why. If you chose a democracy with a socialist society, explain why.

4009

I would choose the Yin's society to live in, since the country is Socialist. In socialism, much stress and worry is lifted from the peoples' brow. Socialism provides all people with work, health care, and living space. ALL people in a country with socialism also have equal shares of money and food. If the government and people do not think the price of a product is fair, they can change it. Products are distributed evenly to all people.

Every single person in a socialism has the same amount of wealth. Therefore, socialist people have more freedom than those in a capitalism. Socialist people can go to countries with capitalism if they want to, since all the people have enough money. This is unlike capitalism, in which numerous people are poor and can not go anywhere.

Socialism is a prime option rather than capitalism, which has many disadvantages.

4012

I chose a Democracy with a socialist society because in a socialist society everyone would be equal. There would probably be less fights and arguments in a socialistic society because no one could say they were better than someone else. In a capitalism society many people die because they can't get enough healthcare, food, or shelter. There is usually more violence in a capitalism because if some people don't have enough things to survive they turn to drug dealing, prostitution, stealing, and other ways to make money. In a capitalism society inflation might occur and cause a depression unlike in a socialism society. Being treated equally is one big reason to live in a socialism.

4015

I chose a democracy with a capitalistic society because you should be able to choose your own business and get rich because if you are poor you probably got mixed up in drugs or something like that or even got pregnant at an early age and if you messed up your life then I think every one else shouldn't have to pay for it. I didn't choose socialism because you should be able to have your own business and get rich if you can and that is why I choose capitalism instead of socialism.

make the options explicit and then choose/defend a decision.

The flow chart scoring system shown in Figure 12 (pages 42-44) is designed to be used with two types of essay prompts that elicit the intellectual operation *evaluation*: those in which the choices are explicit and those in which the options are not clearly delineated. This scoring system is from the analytic scoring anchors described by Nolet, Tindal, & Blake (1993). The flow chart takes into account students' effectiveness in presenting both sides of an argument as well as providing a rationale for this decision, with several general rubrics.

In applying these evaluation criteria, another judge, other than Josh's student teacher, scored the entire class within the same evaluative session. Basically, the criteria

were applied using the rubrics in Table 2 and then adjustments were made on the basis of the range of arguments presented in the essays depicted earlier.

In the earlier description of Josh's performance, the student teacher (Carmen) had evaluated them; the graph in Figure 13 (following the judgment flow charts) depicts another student teacher's independent evaluation of them. The student teachers agreed exactly on two of the four and were off by a point on two essays, as noted in Table 3.

Given these criteria, Josh's performance can be compared directly to those of his classmates. In Figure 14, the actual ratings have been plotted for this group (Josh is the student on the far left, labeled J12). As indicated by the comparability of the bar heights (across students), Josh is certainly performing well near the

Table 2. Criteria for Judging Evaluation Essays

Score Points and Description of Criteria for Judgement

- 5 This essay presents a cogent argument that uses information and presents multiple arguments pro and con. The logic is consistent and compelling and is supported by specific content information.
- 4 Both sides of the argument in support of the choice are made; only one argument is given. The logic is generally consistent but not compelling; content information is present but not dense.
- 3 The essay reflects a clear choice and information is presented that supports the choice with a logical rationale. However, it fails to present both sides of the argument.
- 2 The essay makes a clear choice but only presents one rationale. It fails to present reasons for and against alternative choices and does not provide more than one reason for the choice that is made. This essay barely moves beyond the realm of opinion.
- 1 The essay makes a clear choice but completely fails to present a rationale for the choice or against other options. We know the student's position but we don't know why (within the context of the prompt).

end of the year (two essays are rated 5), but with two important qualifying considerations: (a) others in the class (non-TAG students) also performed at this level on these essays and (b) his performance on two other essays (the first two) is not actually the best in this group (C15's performance is rated the same on one and better on the other and she is a student receiving Chapter 1 services!).

In summary, placement in the appropriate instructional level implies that the student is successful within a domain: Performance is at some level of proficiency or mastery (either defined qualitatively or normatively). In this case, Josh is placed at a level of instruction that is appropriate and defensible. But what is his rate of learning? This topic is addressed next with the information presented so far.

Rate of Learning

The premise behind rate of learning is relatively simple: A certain amount of material or coverage needs to be accomplished within a given amount of time. And more is better. These two dimensions have actually been addressed in the earlier sections; only the format for presenting the information needs to be changed to depict learning rate. Because our model is based on response-defined criterion-referencing (rather than content defined), we can simply stack the performance indicators into the appropriate time sequence and determine the rate of change.

For example, if we re-plot the data from the entire class, taking Josh only and look at his performance on the classwide essays, we would get the following picture of his learning rate.

Table 3. Qualitative Ratings of Classwide Essays-Reliability

Essay Topic	Carmen's Rating	Other Teacher's Rating
Knight and Serf	5	4
Nation State	2	2
Threwnia	5	5
Yin & Yang	4	5

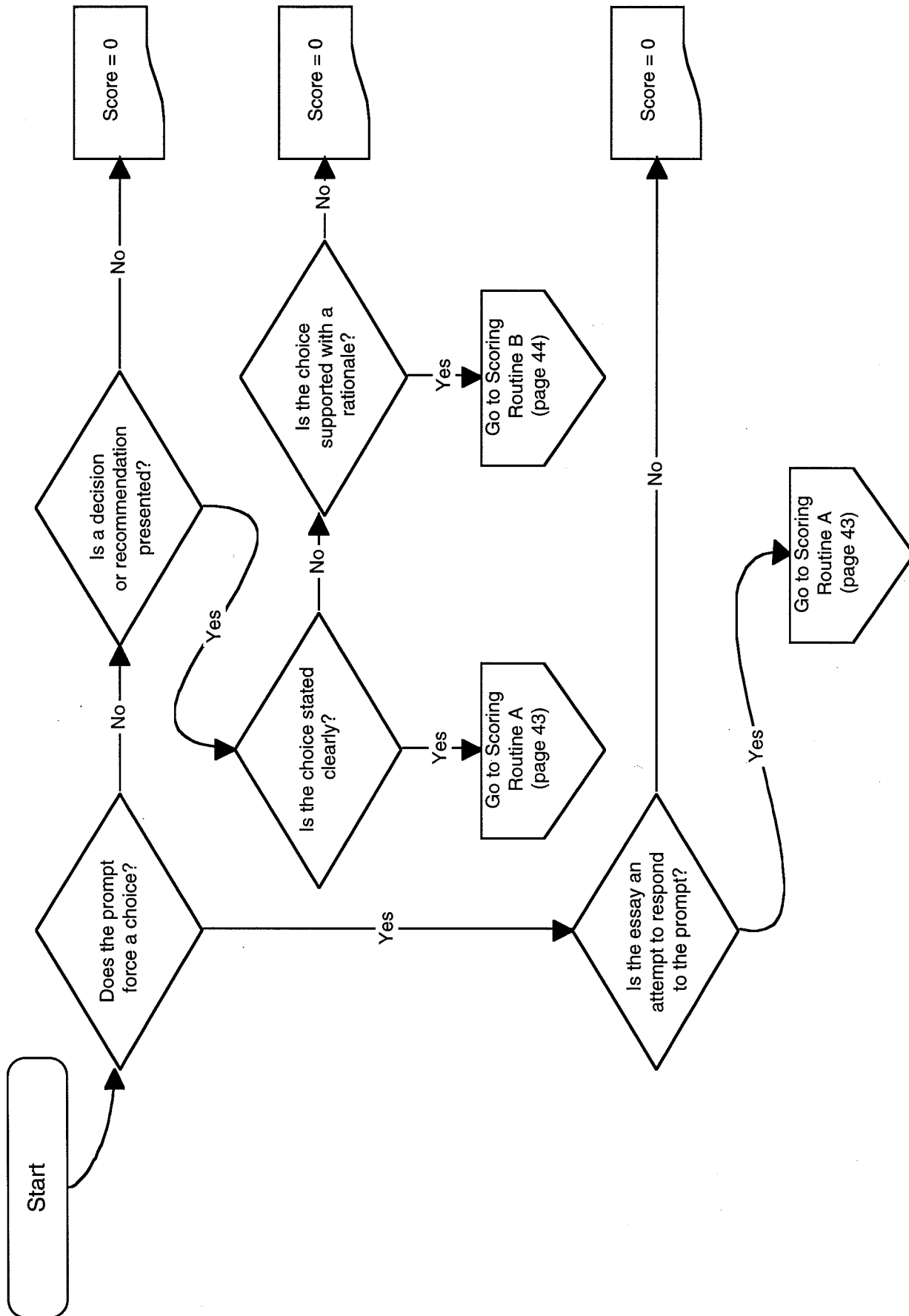


Figure 12. Flow chart scoring system.

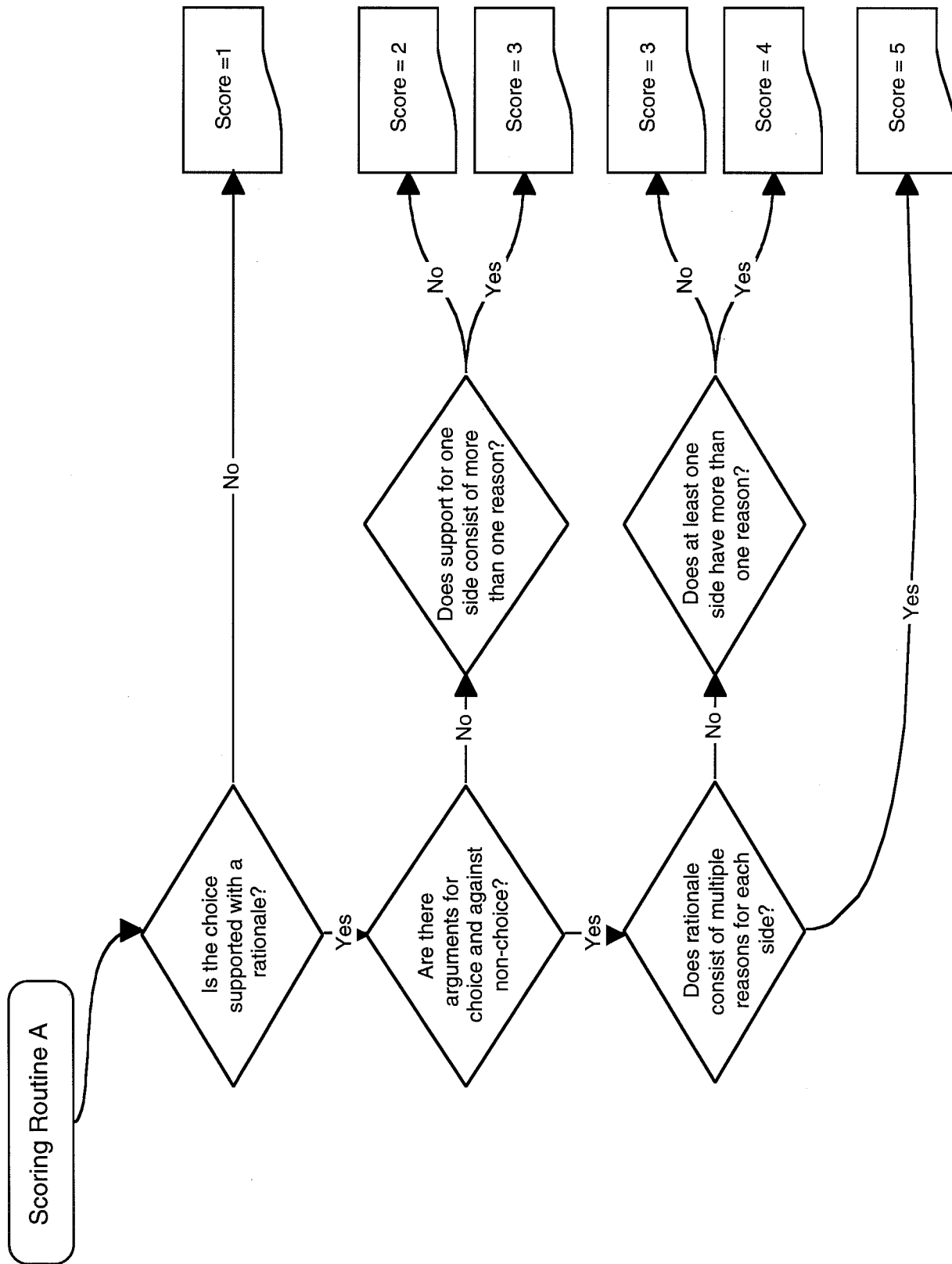


Figure 12. Flow chart scoring system (continued).

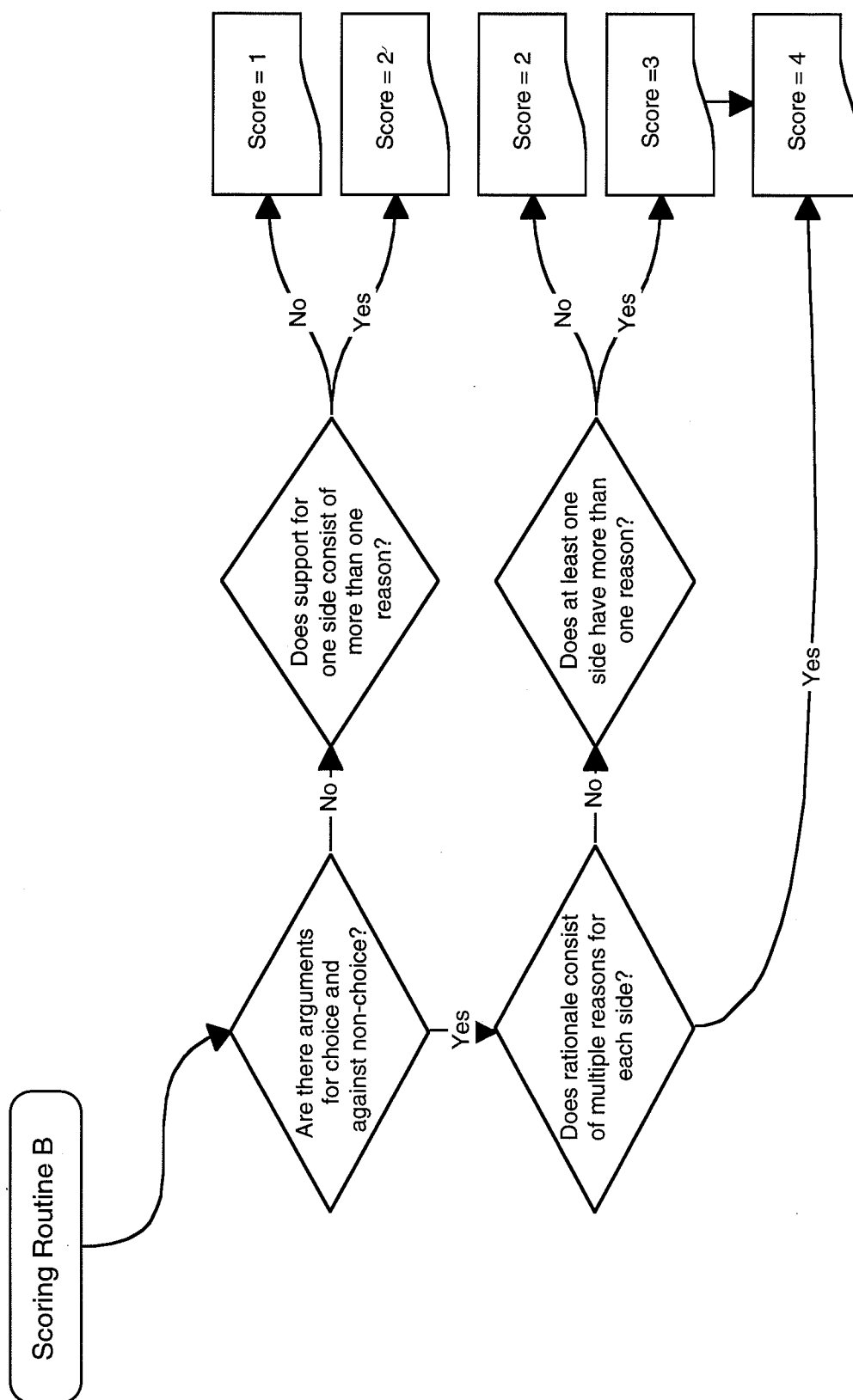


Figure 12. Flow chart scoring system (continued).

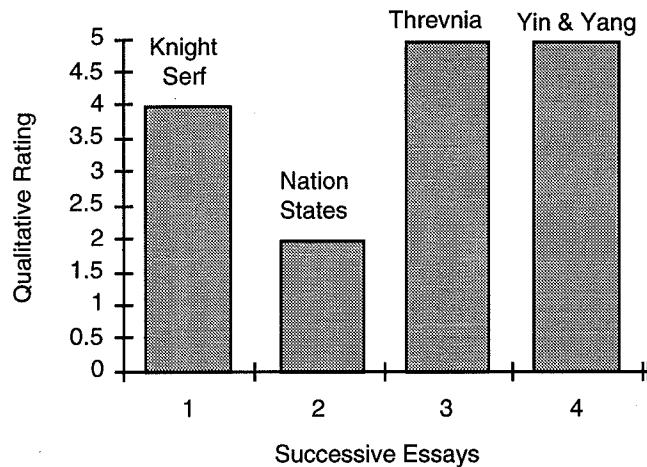


Figure 13. Rate of learning on classwide essays rated (relatively) normatively.

These data indicate that, although Josh may be learning *evaluative* responses, improvement is slight; his critical thinking was well-established when he began the unit. The low performance on the second essay may be an anomaly (a function of the prompt, essay administration, or Josh's personal functioning on that day); then again, it may be content linked, indicating the need for more work in this area. Clearly, at year's end, Josh can make *evaluation* decisions with relative facility.

The next two graphs (Figure 15) depict performance on the final and sub-product essays that Josh took in which his student teacher, Carmen, was the judge.

Josh's performance is similar in the weak signs of improvement that had been obtained when the judgment flowchart criteria had been applied normatively to his classwide performance. In the first graph, it appears that his performance on the classwide essays (this time judged by his student teacher using absolute standards) had a strong beginning, stalled, and then improved slightly. However, in looking at his sub-product essays (the bottom graph), when they were judged in the same (absolute) way by the student teacher, the data are contradictory: (a) considerable improvement in overall quality of argument is apparent until the last essay, but (b) little to no improvement is apparent in his use of concepts until the last essay.

In summary, Josh exhibits a moderate rate of learning that reflects some take in the instructional program. Two reasons may account for these data:

1. Because the whole model for delivering supplemental instruction was new to all parties involved (the content teacher, the student, and the student teacher), the model was never clearly implemented. Indeed, the very focus on intellectual operations was new and a certain amount of practice in the model was needed to

overcome the novelty (or Hawthorne) effects that appeared to be present.

2. It is likely that the basic content instruction provided by the content teacher was sufficient to keep Josh plugged into the units but not enough to structure the pull away programs run by the student teacher. This supplemental program was probably not implemented with sufficient vigor or clarity to actually enhance the content instruction implemented by the general education content teacher.

Nevertheless, the data from this case study support a model of instruction for TAG students that follows the dictates of Oregon state law for providing instruction in the appropriate level and monitoring the student's rate of learning. And it does so without employing a pull-out program, but is based on the content teacher as the primary instructional expert for TAG students.

References

- Alexander, P. A., & Hare, V. C. (1989). Cognitive training: Implications for reading instruction. In J. N. Hughes & R. J. Hall (Eds.), *Handbook of Cognitive Behavioral Approaches in Educational Settings* (pp. 220-246). New York: Guilford.
- Alexander, P. A., Schallert, D. L., & Hare, V. C. (1991). Coming to Terms: How researchers in learning and literacy talk about knowledge. *Review of Educational Research*, 61 (3), 315-343.
- Ambruster, B. B. (1984). The problem of 'inconsiderate text'. In L. R. Duffy (Ed.) *Comprehension instruction: Perspective and suggestions* (202-217). New York: Longman.
- Bacon, P. (1991). *The world past and present*. Orlando, FL: Harcourt Brace Jovanovich.
- Baker, E. L., Freeman, M., & Clayton, S. (1991). Cognitive assessment of history for large scale testing. In M. C. Wittrock and E. L. Baker (Eds.), *Testing and Cognition* (pp. 131-153). Englewood, NJ: Prentice Hall.
- Ball, D. L., & McDiarmid, G. W. (1990). The subject-matter preparation of teachers. In W. R. Houston (Ed.), *Handbook of research on teacher education* (pp. 437-449). New York: Macmillan.
- Banks, J. A., & Parker, W. C. (1990). Social studies teacher education. In W. R. Houston (Ed.), *Handbook of research on teacher education* (pp. 674-686). New York: Macmillan.
- Beck, I. L., McKeown, M. G., & Gromoll, E. W. (1989). Learning from social studies texts. *Cognition and Instruction*, 6(2), 99-158.
- Bloom, B. S., Englehart, M. D., Furst, E. J., Hill, W. H., & Krathwohl, D. R. (1956). *Taxonomy of educational objectives. The classification of educational goals - Handbook: The cognitive domain*. New York: David McKay Co.

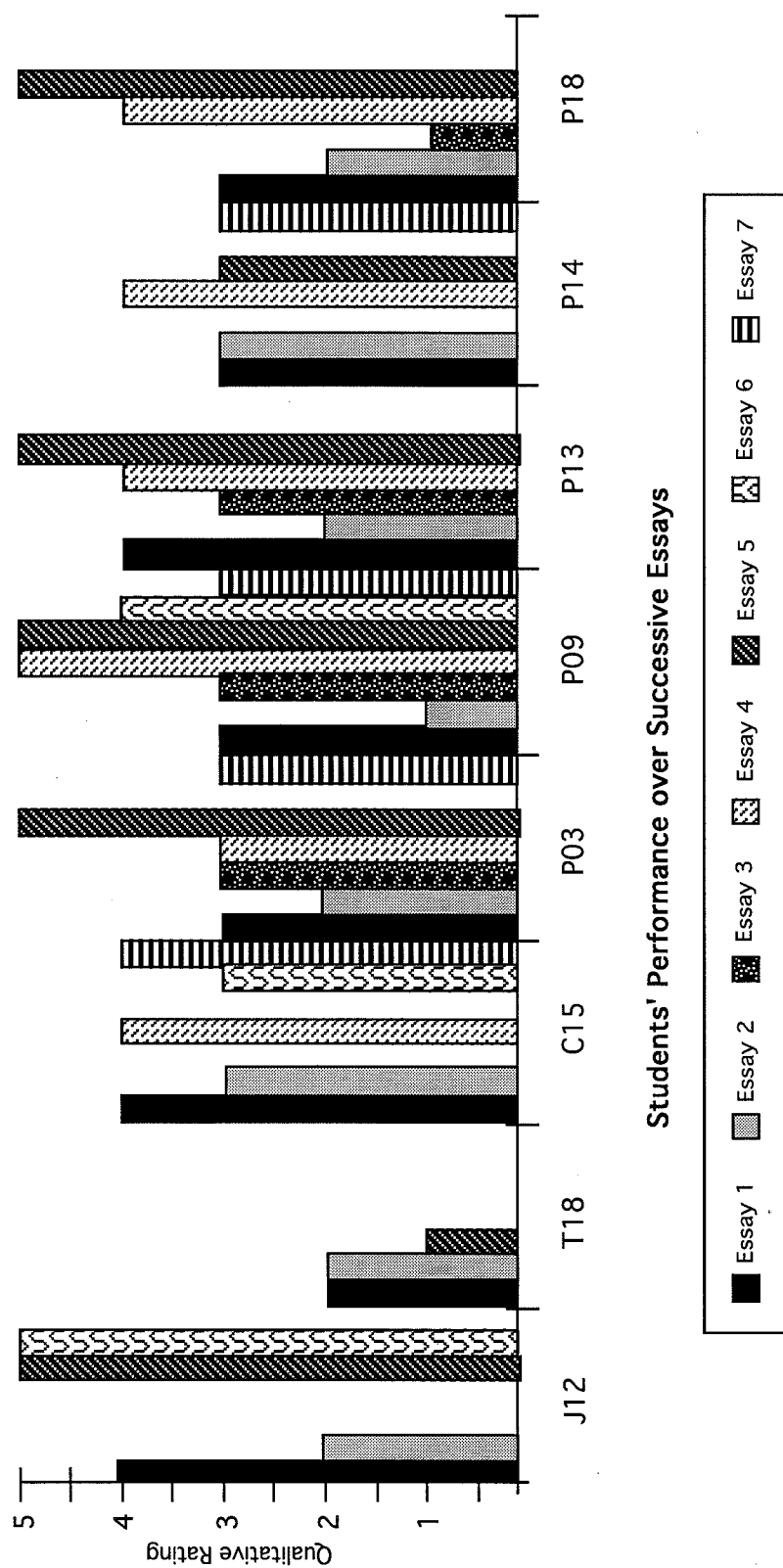


Figure 14. Normative standing in classwide essays.

- Brown, J. S., Collins, A., & Duguid, P. (1989). Situated cognition and the culture of learning. *Educational Researcher*, 18(1), 32-41.
- Carrier, C. A., & Titus, A. (1981). Effects of note-taking pertaining and test mode expectations on learning from lectures. *American Educational Research Journal*, 18 (4), 385-397.
- Carroll, J. B. (1963). A model of school learning. *Teachers College Record*, 64(8), 723-733.
- Carter, K. (1990). Teachers' knowledge and learning to teach. In W. R. Houston (Ed.), *Handbook of research on teacher education* (pp. 291-310). New York: Macmillan.
- Cheney, L. V. (1987). *The American memory: A report on the humanities in the nation's public schools*. Washington, DC: National Endowment for the Humanities.
- Clark, C. M., & Peterson, P. L. (1987). Teachers' thought processes. In M. C. Wittrock (Ed.), *Handbook of research on teacher education* (pp. 255-296). New York: Macmillan.
- Cook, L. K. (1982). *The effects of text structure on the comprehension of scientific prose*. Unpublished doctoral dissertation, University of California, Santa Barbara.
- Crabtree, C. (1989). Improving history in the schools. *Educational Leadership*, 47(3), 25-28.
- Dansereau, D. F. (1978). The development of a learning strategies curriculum. In H. F. O'Neil, Jr. (Ed.), *Learning strategies*. New York: Academic Press.
- Derry, S. J. (1990). Learning strategies for acquiring useful knowledge. In B. F. Jone and L. Idol (Eds.). *Dimensions of thinking and cognitive instruction* (pp. 347-379). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Doyle, W., & Carter, K. (1987). Choosing the means of instruction. In V. Richardson-Koehler (Ed.). *Educator's handbook: A research perspective* (pp 188-206). New York: Longman.
- EPIE Institute (1977). *Report on a national study of the nature and quality of instructional materials most used by teachers and learners* (Report No. 76). Author.
- Fitzgerald, F. (1979). *America revised*. Boston: Beacon Press.
- Grossman, P. (1987). *A tale of two teachers: The role of subject matter orientation in teaching*. Paper presented at the meeting of the American Educational Research Association, Washington, DC.
- Gudmundsdottir, S., Carey, N., & Wilson, S. (1985). *Role of prior subject matter knowledge in learning to teach social studies*. Paper presented at the annual meeting of the American Educational Research Association, Chicago.
- Harper, R. (1966). The new school geography: A critique. *Journal of Geography*, 19, 27-30.
- Harper, R. (1990). Geography's role in general education. *Journal of Geography* (April), 177-185.
- Hashweh, M. (1987). Effects of subject matter knowledge in the teaching of biology and physics. *Teaching and Teacher Education*, 3, 109-120.

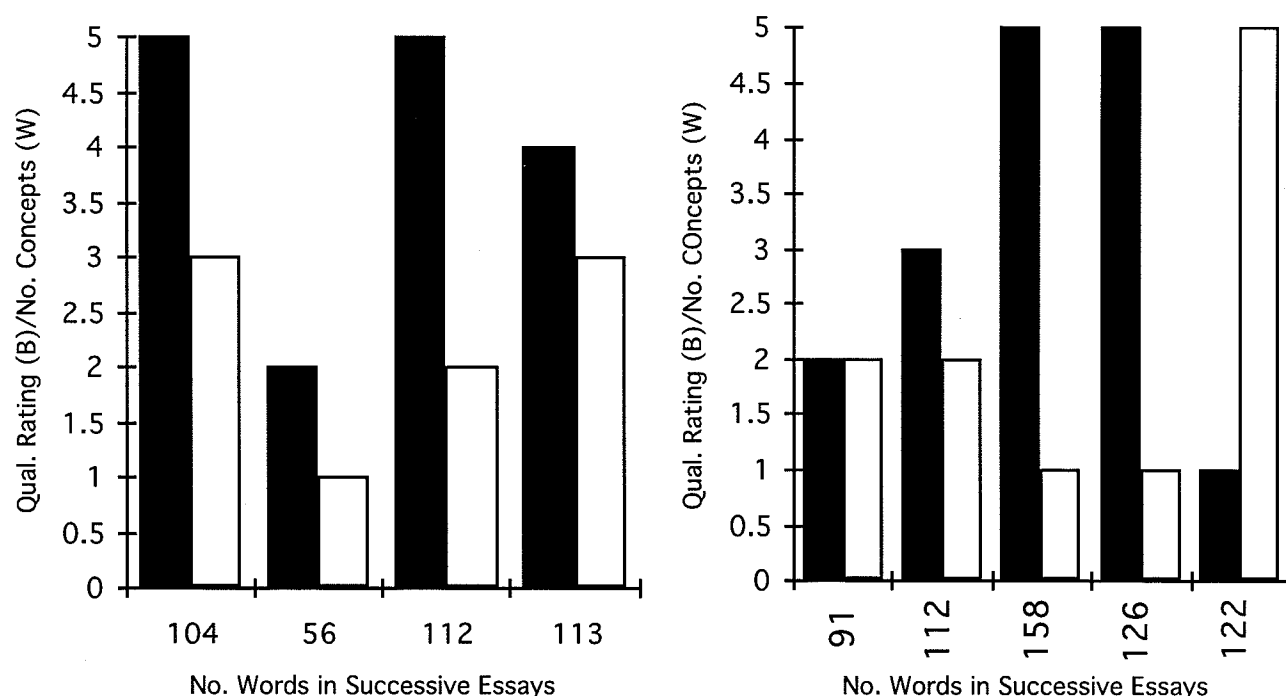


Figure 15. Rate of learning on classwide final and individual sub-product essays.

- Hoover, J. J. (1989). Implementing a study skills program in the classroom. *Academic Therapy*, 24, 471-476.
- Hunt, K. (1964). *Grammatical structures written at three grade levels*. (NCTE Research Report No. 3). Urbana, Illinois: National Council of Teachers of English.
- Kantor, R., Anderson, T., & Armbruster, B. (1983). How Inconsiderate Are Children's Textbooks? *Journal of Curriculum Studies*, 15, 61-72.
- Madaus, G. F., Woods, F., & Nuttall, R. L. (1973). A causal model of Bloom's taxonomy. *American Educational Research Journal*, 10, 253-262.
- Martorella, P. (1972). *Concept learning: Designs for instruction*. Scranton, PA: Intext Educational Publishers.
- McEwan, H., & Bull, B. (1991). The pedagogic nature of subject matter knowledge. *American Educational Research Journal*, 28 (2), 316-334.
- Meyer, B. J. F. (1975). *The organization of prose and its effect on memory*. Amsterdam: North-Holland.
- Meyer, B. J. F. (1981). Basic research on prose comprehension: A critical review. In D. F. Fisher & C. W. Peters (Eds.), *Comprehension and the competent reader*. New York: Praeger.
- Miller, H. G., Williams, R. G., & Haladyna, T. M. (1978). *Beyond facts: Objective ways to measure thinking*. Englewood Cliffs, New Jersey: Educational Technology Publications.
- Nolet, V. W. & Tindal, G. R. (1990). *An analysis of instruction for higher order thinking in middle school geography classes* (Research Report 10). Eugene, OR: Resource Consulting Training Program. College of Education, University of Oregon.
- Paris, S. G., Lipson, M. Y., & Wixson, K. K. (1983). Becoming a strategic reader. *Contemporary Educational Psychology*, 8, 293-316.
- Patton, C. P., Rengert, A. C., Saveland, R. N., Cooper, K. S., & Caro, P. T. (1985). *The world and its people: A world view*. Morristown, NJ: Silver Burdett.
- Prawat, R. (1989). Promoting access to knowledge, strategy, and disposition in students: A research synthesis. *Review of Educational Research*, 59, 1-41.
- Purnell, K. N., & Solman, R. T. (1991). The influence of technical illustrations on students' comprehension in geography. *Reading Research Quarterly*, 26, 277-299.
- Roid, G. H., & Haladyna, T. M. (1982). *A technology for test-item writing*. New York: Academic Press.
- Roid, G. H., & Haladyna, T. M. (1982). The emergence of an item-writing technology. *Review of Educational Research*, 50, 293-314.
- Ryle, G. (1949). *The Concept of the Mind*. London: Hutchinson.
- Schoenfeld, A. H. (1990). On mathematics as sense making: An informal attack on the unfortunate divorce of formal and informal mathematics. In D. N. Perkins, J. Segal, & J. Voss (Eds) *Informal reasoning and education*. Hillsdale, NJ: Lawrence Erlbaum Associates
- Seddon, G. M. (1978). Properties of Bloom's taxonomy of educational objectives for the cognitive domain. *Review of Educational Research*, 48, 303-323.
- Shepard, L. A. (1984). Setting performance standards. In R. A. Berk (Ed.), *A guide to criterion-referenced test construction* (pp. 169-198). Baltimore: Johns Hopkins University Press.
- Shulman, L. (1986). Those Who Understand: Knowledge growth in teaching. *Educational Researcher*, 15 (2), 4-14.
- Shulman, L. S. (1987). Knowledge and teaching: Foundations of the new reform. *Harvard Educational Review*, 57 (1), 1-22.
- Skemp, R. R. (1978). Relational understanding and instrumental understanding. *Arithmetic Teacher*, 26, 9-15.
- Smith, G., & Smith, D. (1989). Schoolwide study skills program: The key to mainstreaming. *Teaching Exceptional Children*, 21, 20-23.
- Tindal, G., Nolet, V., & Blake, G. (1992). *Focus on teaching and learning in content classes*. Training Module No. 3. Eugene, OR: University of Oregon Resource Consultant Training Program.
- Tindal, G., Nolet, V., & Blake, G. (1993). *Focus on assessment and learning in content classes*. Training Module No. 4. Eugene, OR: University of Oregon Resource Consultant Training Program.
- Weinstein, C. E., & Mayer, R. E. (1987). The teaching of learning strategies. *Handbook of research on teaching* (pp. 315-327). New York: Macmillan.
- Weiss, I. R. (1978). *National survey of science, mathematics, and social studies education*. Washington, DC: National Science Foundation.
- White, J. J. (1988). Searching for substantial knowledge in social studies texts. *Theory and Research in Social Education*, 16(2), 155-140.
- Williams, R. G. (1977). A behavioral typology of educational objectives for the cognitive domain, *Educational Technology*, 17, 39-46.
- Williams, R. G., & Haladyna, T. M. (1982). Logical operations for generating intended questions (LOGIQ): A typology for higher order level test items. In G. H. Roid and T. M. Haladyna (Eds.), *A Technology for Test-item Writing*. New York: Academic Press, 161-186.

- Wilson, S. M. (1988). *Understanding historical understanding: Subject matter knowledge and the teaching of U. S. History*. Unpublished doctoral dissertation, Stanford University, CA.
- Wilson, S. M., Shulman, L. S., & Richert, A. E. (1987). "150 different ways" of knowing: Representations of knowledge in teaching. In J. Calderhead (Ed.), *Exploring teacher thinking* (pp. 104-124). London, England: Cassell.
- Wilson, S., & Wineburg, S. (1988). Peering at history through different lenses: The role of disciplinary perspectives in teaching history. *Teachers College Record*, 89 (4), 525-539.

Appendix: Content Planning Worksheet

Directions for Completing Content Planning Worksheet

Complete this planning worksheet for each 2- to 3-week segment of content you plan to teach. This segment probably would correspond to a chapter in the textbook you normally use in the class specified, but it could correspond a entire unit in the textbook, or a few chapters taught together as a short unit, or selected parts of a chapter. However, please refer to a complete segment rather than a specific lesson or set of lessons. For example, if you generally give a test about every two or three weeks (or three or four times a quarter), think of all the material you teach between each test.

Please provide three types of information:

- Key Concepts
- Important Ideas
- Graphic organizer

CONCEPTS

Please use this definition of concept:

- *Concepts are specific words or short phrases that refer to classes of objects or events that share some common defining attributes.*
 - *Concepts involve three parts: a label, key attributes, and a range of examples*
1. Please identify the key concepts that you consider critical for understanding the content you plan to teach during the 3-week interval indicated. Learning these concepts would, in your opinion, mark the difference between mastery and nonmastery of the material you will cover.
List as many concepts as you feel are important, up to ten. Concepts you might target could include terms such as "molecule," "fossil fuel," "holy war," or "vassal." However, specific examples of concepts would not be applicable. For example, the concept "epoch" might be exemplified by "ancient Greece," "ancient Rome," or "the middle ages." These examples would not qualify as concepts according to the definition used here.
 2. List one or two key defining attributes for each concept. These attributes would enable discrimination of what is and is not an example of the concept.
 3. Provide two or three examples of each concept AND, when possible or applicable, also include non-examples that further aid in discrimination of the critical features of the concept.

IMPORTANT IDEAS

Please list *up to three* ideas that you believe are critical to mastery of the content you will teach. Ideas are more general than specific concepts in that they represent unifying themes or topics. Please focus on ideas contained within the context of a single unit rather than global themes or topics that cut across the entire course. For example, in a unit on fossil fuels, you might want students to understand the idea that "Use of fossil fuels results in environmental damage in the form of increased greenhouse gasses and acid precipitation." This idea would be more context-specific than the global theme, "Humans interact with their environment in a variety of ways, with both positive and negative effects," which could apply to a wide range of applications across a science curriculum.

Please frame the important ideas you want students to learn as complete sentences, rather than a few words or phrases.

GRAPHIC ORGANIZER

Sketch a graphic organizer of the content you will teach that shows the key relationships among concepts and ideas.

Appendix (continued)**Content Planning Worksheet**

Date: _____

Teacher: _____

Class: _____

Textbook: _____

Other Curriculum Materials: _____

Approximate Schedule of Content to be Delivered

Week	Dates		Textbook Unit Chapters		Quiz Dates	Test Dates
	From:	To:				
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. _____ |
| _____ |

Appendix (continued)

Concept Descriptions

Concept	Attributes			Examples / Non Examples		
1		Page	1		Page	1
2						
3						
4						
5						

Notes:

Graphic Organizer