

Evaluating Writing in an Academic Setting

Michael Clark
Department of English
The University of Michigan

The current debate about the decline of literacy will probably have no more influence over the use and evolution of language than such debates have had in the past, but it has produced significant changes in the institutional programs responsible for the perpetuation and dissemination of literacy in our culture. Among these changes is the move away from ancient prescriptivisms and mechanistic drills on sentence patterns and towards a more utilitarian or functionalist presentation of reading and writing as useful skills. This shift in pedagogical goals and techniques has resulted in a number of curricular changes, and within the past five years it has also begun to alter both the character of textbooks designed for college students and, to a lesser degree, those designed for elementary and secondary school students as well. Most of the changes involve the incorporation of a much broader range of reading and writing into the "English" curriculum than has been customary, with the aim of introducing students to the many forms of language they will encounter outside the classroom. From a practical perspective, such reforms are eminently sensible, and when the greater variety of texts spans racial, sexual, and class differences, these changes also perform an important cultural function.

It is unlikely that these concrete changes will have much effect on the debate that inspired them, since opinions about the state of literacy in our society are usually motivated by broader political concerns rather than specific evidence or systematic research. Neither teachers nor linguists have been able to identify features of written English that can serve as reliable and consistent measures of quality, and even if they could, the increasing heterogeneity of the stu-

dent population in the United States renders historical comparisons between the writing occurring in today's classrooms and that from fifty or even twenty years ago meaningless and misleading. The growing uncertainty about the possibility of absolute qualitative judgments across a wide range of writing has, however, generated a healthy skepticism about tests that are used to make such comparisons, such as the standardized tests designed by the Educational Testing Service and other organizations. Much of the current concern about the demise of reading and writing has been inspired by the decline in SAT scores in the last ten years (even though ETS has repeatedly warned against drawing such broad conclusions from the scores), and those who propose these tests as evidence for the degeneration of literacy usually cite the accuracy with which the scores can predict a student's grades in college. But teachers interested in a wide variety of writing situations and in equipping students to deal with specific tasks and strategies in their writing have become suspicious about the validity of measuring writing ability according to the abstract, generalized form in which written texts are usually presented in these tests, and that suspicion has led to a number of efforts to design writing tests and develop criteria that reflect the new attitudes towards writing that have emerged in the classroom. Unlike the standardized tests, these locally developed measures of writing skill deliberately reflect the specific programs from which they emerge and focus on the specific interests and needs of the teachers who design them and the students who take them. They usually predict students' success in these programs as well as or better than standardized tests do, but more importantly they constitute an ex-

PLICIT public commitment to evaluate students according to the same principles by which they are taught. As such, these tests serve an important ideological function as well as meeting the bureaucratic needs of assessment, diagnosis, and placement. By connecting the evaluation of written texts to the material conditions in which they are produced and read, these local tests reinforce the truth that real literacy is always a specific response to a concrete situation and never a generalized touchstone for personal development, social respectability, or pedagogical success.

The necessarily decentralized character of efforts to develop these local tests has precluded the formation of a systematic paradigm for their construction. Among the more useful attempts to get beyond the trial-and-error stage of these efforts are the descriptive Guide to Published Tests of Writing Proficiency and A Directory of Writing Assessment Consultants, both published by the Clearinghouse of Applied Performance Testing, and the formation of a National Testing Network in Writing at the City University of New York, which plans a number of forums for the exchange and publication of information regarding the design of different kinds of writing tests.¹ A number of large universities have also combined private and government funds with extensive research facilities to conduct large-scale experiments with testing procedures and curricular designs, and

¹The Clearinghouse for Applied Performance Testing is part of the Northwest Regional Educational Laboratory, 300 S.W. Sixth Avenue, Portland, OR, 97204. The address of the National Testing Network in Writing is CUNY, 535 East 80th Street, New York, NY, 10021.

these programs have begun to yield instructive if not conclusive results.

The Writing Assessment Examination developed at The University of Michigan makes an interesting case study of one such effort. Recent changes in the program of Introductory Composition at The University of Michigan reflect the general movement towards a more eclectic, functionalist approach towards the teaching of writing. However, rather than trying to incorporate a full range of writing tasks like many writing programs in secondary schools and community colleges quite rightly do, the program at Michigan focuses almost exclusively on the sophisticated modes of academic inquiry and expository forms characteristic of reading and writing tasks that the students encounter in their other classes at the university. The program thus resembles more traditional composition programs that focus on academic writing, yet at Michigan that focus is motivated and informed by an emphasis on the connection between the written text and the institutional structure of the functions it must serve. The test used to place students in that program and the criteria used to evaluate that test therefore make a useful illustration of the similarities and differences between tests and criteria proposed to measure writing against a disembodied, generalized standard and those that assess writing as a concrete act within a specific setting.

All students enrolling in The University of Michigan for the first time must visit the campus for an orientation session before they register for classes. During that visit, they take a number of tests to determine their placement in several subjects; among those tests is the Writing Assessment Examination administered

by the English Composition Board. This examination lasts one hour, and students are required to write about a given topic that changes daily. The examination specifies the audience to whom the students are to address the essay, the purpose for which the essay is to be used, and, through the sentences that the students must adopt as the beginning of their essay, the tone, level of diction, and language-conventions to be used. The following assessment examination is one of the many which are used to elicit student writing:

Write a letter to the parents of a young child advocating a particular policy of television watching for their child. Explain to them why you advocate such a policy.

Begin your letter with the following sentence (which you should copy into your bluebook):

By the time the average person in North America graduates from high school, she or he will have seen 18,000 hours of television.

Now select one of the following as your second sentence and copy it into your bluebook:

- A. The present generation of pre-schoolers watches an average of 42 hours of television a week.
- B. Since real experience is the primary source of learning, children are growing up addicted to television and ignorant of life.
- C. According to a well-known critic, television is giving the present generation "an extraordinary exposure to standard adult English and opportunities to see many things."

Now complete your letter developing the argument that follows from the first two sentences. Do your best to make your argument convincing to the parents who are your readers.

Each sample of student writing is then evaluated holistically and scored by at least two trained readers, and students are assigned to one of three categories on the basis of that score. Some students are exempted from Introductory Composition, but, like all other students, in their junior or senior years they must enroll in an upper-class composition course taught by faculty members in their areas of concentration. Most incoming students are assigned to Introductory Composition; others are required to enroll in special tutorial classes where they receive individualized instruction to meet their special needs. These tutorials last seven weeks; students who successfully complete the tutorial course may then enroll in Introductory Composition the following semester. In some cases, students are required to take a second tutorial course before going on to Introductory Composition, and, occasionally, a student progresses so rapidly in the tutorial that he or she takes the assessment again and is exempted from the introductory course. All students, however, may appeal their assignment if they feel they have been misjudged and if their work early in the first semester indicates that they have already mastered the writing skills they will need in their future classes.

The readers who evaluate the writing samples are all experienced teachers of composition. Before they begin evaluating the samples written by the new students, they participate in training sessions in which they read a selection of samples from previous examinations, compare the scores they give with scores given by other readers, and discuss their reasons for scoring each essay as they do. Among these experienced teachers, significant disagreement about the visible characteristics of a specific essay is rare. There is, however, legitimate debate about the relative importance of these characteristics; when disagreements do arise in the training sessions, they usually center on this issue. Although complete agreement about the absolute value of any aspect of the composing process is unlikely, readers must agree on

the relative importance of certain characteristics in order to standardize their scores. Therefore, a list of guidelines or criteria has been developed to assign specific weight to those features of the text which best predict students' potential to perform the composing tasks they will confront in their future classes. Training sessions continue until all readers have incorporated these criteria into their responses to the essays and correlated their scores as closely as possible.

Pairs of readers then begin reading essays of new students who take the examination on orientation days throughout the summer; as part of the reading routine, the two people reading together each day compare their scores for the first few essays in order to insure a continuing correlation among the responses of the whole group. If readers disagree about an essay, they review it with reference to the criteria, and if disagreement persists a third reader scores the essay. Such problems are rare, however, and the correlation among the scores of these trained readers usually remains consistently high over the months of reading, even though the combination of readers changes daily.

The criteria used to train these readers take into account two major sources of information about student writing in the university: (1) actual student essays obtained during preliminary administrations of the assessment examination, and (2) a survey of faculty members' attitudes toward student writing as they see it in their classes. As a result, the criteria reflect both what students do when they try to write well and what their teachers expect them to do to write well in class and in other assignments. This balance of student performance with faculty expectation results in criteria that exempt from Introductory Composition only those students who can already meet the expectations of their future teachers, and that require special work only of those students whose needs are unlike those of most incoming students.

Since these criteria were compiled by professional teachers of composition with

years of experience at all levels of the undergraduate curriculum and with the help of an interdisciplinary committee composed of departmental chairman and writing specialists from most of these disciplines in which the students will major, the readers' evaluations should be good predictions of students' ability to perform the writing tasks they will encounter in their classes. Nevertheless, the accuracy of these predictions is limited by the nature of the skill that is being tested and by the undeterminable relations of that skill to the other factors influencing students' performance in college. This latter problem imposes a formidable obstacle to the reliability of predictions based on any test that focuses on a specific skill. The Educational Testing Service has found that students' success in college is better predicted by their high school grades and counselors' recommendations than by the "raw" ability measured by conventional IQ or aptitude tests. And, while some important factors such as emotional stability, the ability to perceive and adapt to the demands of a specific task, and general psychological traits can be measured by combining the results of several different kinds of tests, there is yet no way to assess accurately the whole complex of forces and influences that determine a student's success or failure in college. As a result, even experienced educators using assessments of psychological health, general aptitude or intelligence, or writing ability at times dismiss the results as hopelessly inadequate predictors of performance.

The skepticism behind this attitude has long been prevalent among educators, counselors, and testers alike. This recognition of fallibility has, however, performed a valuable function in the design and use of such examinations because it requires testers to identify functional limits that specify and measure discrete abilities without having to determine their absolute value.

When a student is asked to explain a chemical reaction, for example, or to solve a mathematical equation or stand-

ardize the grammar and syntax of a sentence, the assigned task isolates and tests the discrete ability to remember or recognize an abstract formula or process and then to use that process to manipulate the specific elements of the problem. In such cases when the skill to be assessed can be so isolated and narrowly defined, criteria can be designed that assess with great accuracy a student's ability to solve such problems. In simpler versions of this kind of examination, the criterion determining the ability to solve the problem is quite simple: the students' solution corresponds to the correct answer. When there are no variables allowed in the solution, such as in arithmetic problems, that correspondence can be measured with absolute accuracy. Moreover, the assessment of ability that results from this criterion is also highly accurate since the solution itself can be identified with the process of solution because only one process will produce the correct answer. Thus, if the ability to be measured is defined within such narrow parameters and if the variables in the solution are held to a minimum, criteria can be designed that will accurately measure that ability and, consequently, reliably predict a student's success in situations that are of comparable complexity.

The key to the accuracy of such assessments is the identification of the process of solution with the product, the solution itself. As soon as a solution may be produced by more than one process, however, or as that process entails more than one or two simple steps, the reliability of evaluations based on the identification of the product with the process begins to decrease. If, for example, the answer to a question can be guessed, the correct answer is not necessarily a criterion for mastery of the process of solution; similarly, if the process involves a number of steps, the last of which is a simple arithmetical calculation, a student may master the entire process, make a mistake on the last step, and still get the wrong answer. In both cases, the evaluation of the product will not be a reliable measure of the stu-

dent's ability to perform the task. This dilemma can be resolved to some extent by increasing the variables in the solution so that the failure to complete one aspect of the process does not render the solution totally incorrect; but, just as the accuracy of the criteria decreases when the complexity of the process increases, when the complexity of the solution increases, the correspondence between student answers and the acceptable range of answers can no longer be measured exactly. When the process is as various and ill-defined as the process of writing an essay, and when the product has as many variables as a piece of writing, the number and obscurity of the variables render the identification of the product with the process totally intractable. As a result, if assessments of the composing process are to claim any accuracy at all, that accuracy cannot be based on the same principles that determine the criteria for other kinds of tests. Those who use the assessments as unconditional statements about students' ability would imbue the results with a precision that is not only impossible to obtain but different in kind from the use for which the examinations were originally designed.

The simplest solution to this problem would be to measure the quality of the finished product. But, since the assessment is used to predict a student's future ability to write very different kinds of essays under very different circumstances and constraints, the quality of the specific essay a student writes for the assessment is largely irrelevant to the purpose of the examination if it does not in some way point beyond the limits of the text. Measuring the process of composition itself is equally untenable, for even if the practices of every student could be observed while they wrote their essays, the actual stages in the composing process are virtually impossible to document and judge. It is possible, however, to justify criteria based on those features of the writing sample that are least restricted to the particularity of the finished text, under the assumption that such features--because they are not

Parents, if they allow television to be viewed at all, should be very careful what shows the family sees and furthermore, what shows they themselves, watch and express interest in. Because, as children will believe the television, even more than the television will they believe their parents.

This manifest attention to the cause-effect relationship itself suggests that the writer knows and can use an organizational principle that is not constituted by the specific subject of his essay. This student could presumably organize a variety of subjects in this way. Thus, even though the second example may be no more acceptable than the first to a reader looking for polished texts--indeed, in some ways the first is more interesting--the second indicates a grasp of the principles by which other organized texts may be produced and so would receive a better score by the readers in the assessment program. (In this case, the author of the first essay was required to do tutorial work before enrolling in Introductory Composition, whereas the author of the second was allowed to enroll directly into the introductory course). To be considered for exemption, a writer must be able to integrate such indicators of organizational ability more smoothly with the specific content of the essay than this writer did. So even in the evaluation of the more accomplished essays, criteria are used that measure those features of the text that indicate such an effort rather than measuring the complete success of the attempt to polish the final product. Consequently, these criteria should be reasonably accurate measures of the writer's ability to produce a number of different texts in different situations as well as direct evidence of the ability to complete the specific task of the assessment examination. Evaluations based on these criteria should be judgments of the writers' potential rather than of a product written under the artificial and unique constraints of the examination.

In addition to this theoretical framework, development of criteria used to

evaluate the writing samples was influenced by a more practical consideration. At any large university, completely individualized instruction in composition cannot be offered to all students. If, however, the needs of most incoming students can be determined and are made the basis for the central program in composition, then instruction in that program can be particularized if not individualized. This particularization in turn allows for a more thorough treatment of the problem shared by most students than would be possible in a class designed to meet the needs of all the students. Because the curriculum of Introductory Composition has been designed with this goal in mind, the course is limited to those students who have already mastered three composing skills that are basic to writing in the University: (1) they can organize their essays to fulfill simple argumentative or rhetorical purposes; (2) they can write sentences that present a specific idea and recognize the place of that idea in the structure of the whole essay; and (3) they can maintain consistent patterns of case endings and punctuation marks. Since the purpose of the assessment is to place students where they can receive instruction suited to their needs, the criteria must reflect this purely pragmatic division of skills and distinguish the features that indicate their mastery. In the rare cases when students are unable to perform one of these three kinds of tasks but are able to perform the other two quite easily, they may be allowed into Introductory Composition and advised to attend voluntary individual counseling sessions in the Writing Workshop to bring their weakest skill up to the level of the others. Should a student's essay indicate not only mastery of all three skills but also demonstrate the ability to adjust these skills precisely to the specific task of the assessment, then that student may be exempted from Introductory Composition.

Because the criteria used in the assessment procedure must respond not only to the nature of the examination but also to the role that examination plays in the

entire composition program at The University of Michigan, the features described by those criteria are distinguished by their usefulness to the evaluative procedure rather than by any value or function they may have in a specific text. Thus, the examples used below to help characterize those features are not in themselves necessarily examples of good or bad writing. Their exemplary function exists only within the parameters established by the purpose of evaluation, and that purpose is, in turn, a product of the specific theoretical and practical constraints described above.

DESCRIPTION OF THE CRITERIA USED TO EVALUATE THE WRITING SAMPLES

I. Organizational Features:

These features indicate the students' knowledge of simple organizational principles and their ability to adjust them to the specific constraints of the examination.

A. The specific topic of the examination appears as the central focus of all parts of the essay. The topic must be presented in a form general enough to comprehend all subjects in the essay and yet not exceed the boundaries of the topic as it is presented in the examination and limited by the combination of sentences that the student chooses from the given options. The essay, in turn, must explore a range of subjects encompassed by the central topic while recognizing the limits it imposes.

The range of subjects treated by each of the better essays usually reflects the breadth of possibilities encompassed by the topic given in the examination as well as its limits. Thus, the following example deals only with the relation of TV to a child's other experiences, but it explores several aspects of that relation.

EXAMPLE TWO

(1) By the time the average person in North America graduates from high school,

she or he will have seen 18,000 hours of television. (2) Dr. Edward Palmer, head of research at Sesame Street, writes: "I think that watching television is a rather remarkable act in itself. All the while kids are watching they're...actively relating what they're seeing to their own lives." (3) Because early childhood is a crucial developmental stage in which the individual forms the skills of comprehending reality and living in the real world, it is consequently crucial that the television programs viewed by the young, impressionable mind give it the kind of experiences which are wholesomely relevant to its own real life experiences.

(4) This is particularly true because recent experimental work has shown that young children cannot distinguish between the commercial and the program, nor do most regard the characters seen on the television screen as electronic reproductions of fictional characters. (5) Rather, most young children actually perceive even cartoon characters as living beings who exist within the television. (6) The very young apparently cannot distinguish between the fantasy world and the real world. (7) It can thus be seen that virtually anything viewed on the television could be interpreted by children to be something very real.

(8) Additionally, it has long been observed that children seem to have a need for role models. (9) That is, children have a need to emulate people and their actions for the purposes of forming their own identity. (10) For those unfamiliar with psychological theories, most people in the field agree that human beings have a need to define themselves in terms of other people and other experiences. (11) This process of forming an identity, for example, "I am a boy," and then later in life "I am a Catholic," or "I am a doctor," begins very early in childhood.

(12) In the past, parents and close relatives served as almost exclusive as the role models. (13) Today, parents as well as relatives are often absent, so that television characters become surro-

gate role models for the children. (14) Since children perceive television characters as real, and since children will emulate whatever is available, it is crucial that children view proper television material, particularly during their early childhood.

(15) Caution and vigilance must be exercised by the parents in view of all the violence and sexually inclined actions on the air. (16) For example, sexually oriented actions occur at the rate of two per hour on television, mostly between unmarried persons. (17) Unless parents are present to give at least equal time for a counter-example, the child is in danger of forming his values and concepts around the fantasy-type world of television programming which only the adults can perceive as fantasy, fiction, or entertainment.

(18) Although children can profit immensely from television, they can do so only if society is willing to insure that the content of these programs is positive and relevant. (19) The child's programs should be oriented towards reality as it is, but should also directly and indirectly emphasize values other than instant gratification of one's own need for excitement and pleasure. (20) Values such as concern for others, honesty, diligence, etc. are present in some television programming. (21) Parents should consequently view a television series before the child does, or at least with him, so that positive things can be reinforced while negative aspects can be interpreted to the child so that he does not develop inaccurate concepts or emulate values which are contrary or undesirable to his parents and society at large. (22) It is essential to realize that the adults of tomorrow's world may very well be practicing the values which they have derived from the childhood experiences. (23) If that experience contains only the assertions of violence, sex, and self-gratification seen on so much of television today, many people will be living lives devoid of those values which distinguish man as an intelligent, moral being.

Most of the essays lack the sharply defined limits of Example Two, and they also fail to recognize adequately the range of subjects relevant to their topic. In the following example, the topic of a child learning from TV is restricted without justification to the topic of televised information about birth control.

EXAMPLE THREE

(1) The average home in North America has a television set on for more than 6 hours a day. (2) Since all experience contributes to learning, children in North American homes learn a great deal from television. (3) The topic of birth control and the use of contraceptives is one important subject that should be expressed in television.

(4) In today's complex society, little is done to teach young children of the vital issues dealing with birth control. (5) On the other hand, television allows its viewers to watch commercials publicizing the consumption of alcoholic beverages and other such invaluable products. (6) North American children should approach sexual activities in a mature manner. (7) Since most parents explain very little to their children about this topic, it is up to the television networks to use their control with the utmost concern.

(8) A program in this form would eliminate many problems between the young adults of America. (9) It could decrease the percentage of unwed mothers, abortions, forced marriages, and other social disorders. (10) Problems such as these, are only a menace to society and should be treated with great care. (11) Publicizing birth control and the use of contraceptives could only help children, as well as the parents.

(12) Television in America can be a tool used to support American youth throughout society. (13) Birth control is one of many topics that should be accepted in a civilization of the future. (14) The use of this topic through television will give all Americans a better understanding of the proper care in their future lives.

Here, even though the essay is clearly focused on something the child can learn from TV, the writer's failure to broaden the focus to the full range of the topic described by the first two sentences of the essay suggests that she or he would have difficulty adapting this skill to other tasks and subjects.

In the poorer essays, the range of subjects usually is similarly limited, but here even the restricted range is poorly focused and at times ignores the limits of the topic.

In the following example, the writer begins by ostensibly limiting the essay to the waste of time involved in watching TV, but then addresses the problems of value vs. entertainment, irresponsible procedures, commercialism, and exploitive sex:

EXAMPLE FOUR

(1) The average home in North America has a television set on for more than 6 hours at a time. (2) Every member of the average North American family watches at least 3 hours of T.V. a day. (3) It's very important that the programs viewed are of worth and value, as 3 hours is a long time to waste on useless trash. (4) If a child is left alone, with his eyeballs staring motionless into a T.V. he/she will surely pick the most entertaining programs. (5) These, however, are not necessarily the most beneficial to the child, yet it is difficult to constantly police the T.V. set. (6) So the root of the problem lies in the lack of concern of T.V. producers with the quality of T.V. (7) They're just trying to attract attention for their sponsors, and they do so by kramming their 1/2 hour situation comedies with sex. (8) But for the time being, parents, just keep an eye on what your kids watch, and try not to act like the culprits on T.V. themselves.

B. The essay consists of several discrete units distinguished by their subjects and/or functions. One of the units must clearly function as an introduction, another as a conclusion. Intermediate

sections should perform a variety of clearly distinguished functions such as definition, exemplification, qualification, restriction, or elaboration, and the boundaries between sections should be distinct and firm. Where the boundaries are unclear, markers such as "in conclusion" or "for example" should reinforce divisions in the text (see II.D, below).

In the better essays, subjects are clearly grouped according to their function, and that function is usually described implicitly by sentence modifiers when it is not obvious. A wide variety of functions is demonstrated, and the subjects are appropriate to the function they perform. Thus, in Example Two, the subjects of the introductory paragraph are broad enough to include all of the specific subjects of the essay: the introduction enumerates a "developmental stage," "skills of comprehending reality and living in the real world," "the kind of experiences which are wholesomely relevant" to "real life experiences." All of the general subjects are then made more specific and concrete in a section devoted exclusively to that function (sentences 4-7). Following this section, another specific subject of the topic appears (role models), its function in relation to the preceding subjects is described ("Additionally"), and the subject is then defined (sentence 9), its relevance to a developmental stage explained (sentence 10), and concrete examples are offered (sentence 11). After several other sections, which demonstrate other functions, the essay concludes with a section that moves back to more general subjects (sentences 18 ff.) by referring to the specific subjects of the essay in terms that give them the broadest scope possible within the limits of the topic (see especially sentence 22). Within each section, the subjects can easily perform their function; in sentence 11, for instance, the examples offered to explain the process of forming an identity are concrete enough to serve as clear examples and specific enough to identify stages in that process as the nature of that identity moves from the general ("I am a boy") to the more particular ("I am a Catholic").

Most of the essays differ from Example Two in that they demonstrate fewer functions, the differences between sections are not as clear, and the boundaries between sections are less sharply defined. In the poorest essays, such as Example Four, only a few simple functions are present, they are not confined to discrete units, and their connections to one another are ambiguous. In Example Four, the central topic is difficult to identify because most of the subjects exist at the same level of generality; hence the difficulty of finding a clearly introductory unit (though sentence 8 provides a clear conclusion). There are no examples offered as such, little explanation, and no elaboration of the absolute and apparently arbitrary statements. Consequently, this essay offers little to indicate even an awareness of these basic principles of exposition.

C. The essay recognizes and responds to the extra-textual constraints of audience and purpose as defined by the examination. The explicit references to these constraints should be appropriate, and the diction of the essay and the extent of elaboration and definition should fulfill the specific demands of the purpose and the specific needs and expectations of the given audience. Thus, whereas the writer of Example Two has chosen an objective, declarative voice, she or he also explicitly recognizes the breadth of the given audience and responds to it (sentence 10). Furthermore, since the specific subject chosen to elaborate role models is especially appropriate to an audience of parents, the central topic of the relation of TV to a child's life is brought to bear directly on the concerns of the audience. In the conclusion, the application of the topic to its ultimate purpose is made explicit (sentence 21) without completely changing the focus of the essay to describe the details of the policy, a gesture that indicates the author's full incorporation of the precise task described in the exam.

This precise adjustment of broad organizational features to the constraints of a given audience and purpose characterizes

only the better samples. The majority of the essays do not incorporate the demands of the audience and purpose into the text so thoroughly, although many address the audience directly and explicitly explain a policy for controlling a child's viewing habits. In the poorer essays, these features are often internally inconsistent as well as isolated from the extra-textual constraints. For instance, in Example Four the formal syntax and careful diction of the first clause in sentence 3 is quite different from the strident, colloquial phrases that end the sentence, and sentence 4 reflects a similar contrast between the formal features of the last clause ("he/she," "surely," "the most entertaining") and the informal phrasing of the interjected prepositional phrase "with his eyeballs staring." Similar inconsistencies characterize the rest of the essay, such as the abrupt shift from the indicative to the imperative mood and the shift into the informal phrasing and diction of sentences 7 and 8. When they are controlled, such contrasts can be very effective and can indicate exceptional skills. Here, however, there is no context in which the shifts can function, so they suggest that the writer is both unsure of the expectations of the audience and unable to control diction and syntax. Furthermore, the recognition of the assigned purpose in sentence 8 is irrelevant to the rest of the essay, and its connection to the essay through the introductory phrase implies some sort of long-range plan that is not mentioned in the preceding sentences. Not only is this writer unable to identify and manipulate those features of the text that would enable a response to extra-textual constraints, she or he is also unable to control them to meet the purely textual demands of consistency and internal coherence within the essay itself.

II. Local Features:

These features indicate the student's ability to portray words, phrases, sentences, and even groups of sentences as part of larger units up to and including the essay. These features are not, of

course, independent of the organizational features; indeed, those global features emerge only from the local features, and the propriety of the local features, in turn, is determined by the global characteristics of the essay they constitute. Consequently, distinctions between these two groups of features cannot reflect separate stages of composition or even clearly discernible levels of the text. These distinctions, like the features they categorize, are only articulations of the evaluative process, and the skills they indicate are similarly only convenient predictors of a student's ability to meet the demands of future teachers, not characterizations of the composing process.

A. The semantic information conveyed by the unit (i.e., a phrase, clause, group of sentences, etc.) is appropriate to the local boundaries of the unit.

Semantic information is simply the meaning of the text, its propositional content. In an essay about the debilitating effects of watching TV, for example, all comments should in some way convey information about that problem, and in a section of the essay devoted to explaining how TV stifles the imagination, all the statements should convey information about that specific debilitating process. The importance of this feature as a unifying device and its consequent relation to I.A is obvious enough; its properties, however, are not always easy to identify. Generally, the semantic range in any unit must be wide enough to elaborate the relationships that the writer wishes to convey and narrow enough to remain within the boundaries of the unit. Thus, the semantic information of a unit can indicate a student's ability to recognize the limits of that unit and explore the range they delimit. This feature thus indicates the same skills indicated by the features in I.A and differs from the global features it resembles only in its extension down to units smaller than the essay. Semantic information thus furnishes a criterion for identifying those students whose difficulty in organizing their essays stems from their inability

to decide how much and what kind of information to put in their sentences rather than from an inability to handle larger units.

In the poorer essays, the semantic information of sentences and divisions within paragraphs tends to be completely restricted to the boundaries of the unit itself. Thus, in Example One, the semantic information conveyed by sentence 3 is unrelated to that conveyed by sentence 4, and the information conveyed by sentence 5 bears little relevance to the information in sentences 4 and 6. At other times, the limits of the unit are ignored, and totally unrelated kinds of information are enclosed by a single unit. In sentence 6 of Example One, the children's isolation and their inability to learn in school are both confined by the same part of the same sentence, though the following sentence makes no effort to connect them. Occasionally, as in Example Four, the poorer essays will organize semantic information into larger units, e.g., the pair (sentences 6 and 7), but those units are then relatively isolated from surrounding units. In the better essays, on the other hand, the semantic information of one unit often appears as a subordinate part of the surrounding units, creating a coherent passage that develops an idea by gradually increasing the range of semantic information used to express it. In the following example, the semantic information of one sentence appears in a slightly varied form in the next sentence, which in turn goes on to focus on the variation, and so on. In this case, by systematically repeating this technique the author is able to develop the purely statistical information given by the examination into a powerful rhetorical force that justifies the purpose of the essay.

EXAMPLE FIVE

(1) The average home in North America has a television set on for more than 6 hours a day. (2) Every member of the average North American family watches at last 3 hours of television a day. (3) This means that 21 hours a week are spent

in front of a television by an average family member. (4) When viewed in such short time spans, the hours are not necessarily seen as many. (5) When you consider that this also means 84 hours a month and 1008 hours a year, the statistics become a bit more frightening. (6) If you consider that your child is spending 42 days out of one year in front of a television set, you might begin to feel strongly about establishing a policy for television watching.

B. The grammatical relationships among elements of a proposition accurately reflect the logical or affective relationships determined by other features of the essay.

The grammatical information conveyed by a unit enables us to determine causal, circumstantial, or other relations between the pieces of semantic information it conveys. For example, in the sentence "The police arrested my roommate yesterday," the action of the event is done by "the police" and done to "my roommate." The surface subject of the sentence is "police" and the surface object is "roommate." In the following sentence, the surface roles have been reversed but the grammatical information is unchanged: "My roommate was arrested by the police." Here, the action is still performed by "the police" and still done to "my roommate." Thus, if a student is discussing the ways that TV affects a child's relationship to his parents and then without explanation writes a sentence in which the familial relationships affect the child's viewing habits, that grammatical information would contradict the information given by the surrounding sentences even though the semantic information would remain within the boundaries of the unit. (If such boundary violations are frequent, then the problem may be organizational rather than grammatical; i.e., the sentence may correctly represent the student's idea but simply be in the wrong place in the essay). Without specific diagnostic testing, the exact nature of the problem can only be assumed to stem from a generally tenuous grasp of the information being conveyed by the sen-

tence and so always indicates potential problems.

The basic relationships conveyed by the grammatical information are quite simple, and only the poorest essays contain serious inconsistencies in this feature. Perhaps the most common problem occurs when a student is unsure of the technique by which modifying clauses and phrases may be related to the independent clause of a sentence and so tends to use dangling modifiers ("Flying over the city, my house looked great") or to confuse the agent role in passive constructions ("My roommate was arrested by not being able to run fast enough"). Such problems are usually accompanied by many others. The assessment process penalizes no student for occasional slips if the context of the unit clearly compensates for the error.

C. The thematic information determining relative importance of elements within a unit is clear and consistent, and it is appropriate to the role of the elements in the context of the unit and the essay.

This feature, which is sometimes called "foregrounding" or "weighting," indicates the ability to manipulate the grammatical and syntactical roles of elements in the sentence so that the sentence emphasizes certain pieces of information and therefore reflects priorities determined by such extra-textual factors as the purpose of the essay and the expectations of the audience as well as by the function of a sentence in the context of the essay. Information about the relative importance of these pieces is carried by syntactic structures and by the order of words in clauses and phrases, as well as by explicit description such as "And now for the most important factor in TV programming: commercials." For example, in the noun phrase "A white powdery snow" the most important word is the substantive, "snow"; in the following phrase, however, departure from conventional word order gives special emphasis to the adjectives: "A snow, white and powdery." Moreover, a number of other syntactical constructions can reverse the usual dominance of sub-

ject over modifiers: "The snow was white and powdery." More complex relationships can be further described by subordination and coordination. The relative importance of the elements in the sentence "The white powdery snow blew against the house" changes when the semantic information is rearranged: "The snow was white and powdery, and it blew against the house." Depending on the specific connectors that are introduced, both the relative importance of the pieces and their relationship to one another can change: The causal connection in "The snow blew because it was white and powdery" is different from that in "The snow blew although it was white and powdery," and so forth.

The criteria for determining a student's grasp of these principles are the number of different kinds of constructions that appear in the essay and the correspondence of those constructions to the context in which they appear. In Example One, sentence 6, two different grammatical relationships are contained in the same sentence and joined by the coordinating conjunction "and." Their relationship is therefore merely additive, and they are roughly of equal importance. Since the sentence is serving as the topic sentence of a paragraph about the latter relationship, it should indicate that the latter relationship is the true focus of the unit. The function of the sentence would have been more precise and its relationship to the sentences that surround it would have been clearer had the student changed the grammatical information to emphasize the true actor in the group and subordinated the first action to the second: "Not only can children not play with others in a normal fashion after they have watched a lot of TV, they can't even learn at school." This rewrite assumes, of course, that the parents' comprehension is not the true subject of the unit and that the paragraph's real subject is the child's learning difficulties. As the student has written the passage, these assumptions are tenuous at best, based on an independent interpretation of the semantic information rather than a reading of

the essay. While the semantic information carried by the words of the unit is not contradicted by other information, the context of the unit does contradict both the grammatical information concerning the role of the parents and the thematic information concerning the relative importance of the two clauses about the children's difficulties. Naturally, when two or more such sentences are strung together, the degree to which each sentence is vague cannot be determined because the context created by the sentences is vague too. In the following passage, the grammatical relationships between the participants in the proposition are unclear, and no focus emerges because emphases determined by the thematic information are inconsistent.

EXAMPLE SIX

- (1) The human imagination, which is so necessary to achievement, becomes worsened after exposure to television.
- (2) Television shows everything in detail, leaving little to the imagination.
- (3) Problems are almost always solved correctly, making only one solution to a problem seem possible.
- (4) The element of uncertainty rarely becomes visible.
- (5) Although it shows bright, imaginative people as winners, children who watch television are assumed to be dull and unimaginative by television programmers.

When the grammatical and thematic information is made more consistent, however, a focus emerges and the relationships among the participants are clarified:

TV worsens the human imagination, which is so necessary to achievement, by showing everything in detail and leaving little for the imagination to do. Because problems are almost always solved correctly, TV suggests that only one solution to a problem is possible so the element of uncertainty rarely becomes visible. Although TV shows bright, imaginative people as winners, TV programmers assume that children who watch television are dull and unimaginative.

The difference between these two passages indicates the difference between the way these features appear in the poorer essays and the level of proficiency that a student must demonstrate before enrolling in Introductory Composition. The essays of those exempted from the introductory course seldom contain evidence of difficulty with the grammatical and thematic properties of text.

D. The function of discrete units, their boundaries, and the relationships between the units are clearly marked where necessary.

This purely structural information is usually carried by adverbs and sentence modifiers such as "consequently," "thus," "secondly," and so on, though occasionally a sample will contain transitional sentences or even paragraphs that explain the relationship of two sections of the essay. In general, because structural information is simply predictive in nature and signals that a particular unit of the discourse is to follow or has just ended, a structural division can be marked by almost any change in the syntax, mood, or person of the text. In the better essays, such as Example One, these markers are unobtrusive and appear only when the structural function they mark might otherwise be unclear. In the poorer essays, the markers might be absent altogether or they might not correspond to the textual property they ostensibly mark. For instance, in Example Four, sentence 8, the introductory phrase alludes to an antecedent plan that is not in the text, and more obvious discrepancies occur when a student uses logical markers such as "thus" and "consequently" without arranging the essay to reflect the logical sequence they suggest.

These markers are present in most essays, but if students are unsure of their writing or unaware of the techniques by which the markers can be incorporated into the rest of the text, the markers may be redundant and obtrusive:

EXAMPLE SEVEN

(1) The average home in North America has a television set on for more than six

hours a day. (2) Every member of the average North American family watches at least three hours of television a day.

(3) What does this mean? (4) Well, on an average it means that every member of the average North American family spends one-fifth of their time awake watching television. (5) Assuming then that that person either works or goes to school for approximately eight hours, he then spends about one-third of his 'spare time' watching television.

(6) Is this then good or bad? (7) First of all it depends upon what type of programs are watched, and secondly, what is done in their 'spare time' while they are not watching television. (8) We first of all have to ask ourselves, or the other person, "Why am I, or you, watching television?" (9) Are we watching television because we enjoy it, because we hope to get something out of it, or are we just watching television for lack of something better to do. (10) This is where I feel television can be harmful, when someone turns on the television and watches whatever is on for lack of something better to do.

The better essays, therefore, will be marked neither less nor more than is needed to indicate the movement of the argument. Most essays are marked excessively, but this redundancy is not in itself a serious problem unless the lack of confidence it suggests becomes troublesome. Only if the markers are totally absent, or if they do not correspond to the text, does the essay indicate serious problems that usually require work in addition to that available in the introductory course.

III. Graphemic Features:

These features, which include spelling, capitalization, paragraph indentation, grammatical case endings, and punctuation marks, indicate two very different kinds of skills. A consistent marking of the grammatical and syntactical properties of a text, regardless of the markers themselves, is considered an indication of

the writer's recognition of those properties. Thus, if the inflection of the verbs in an essay is consistent though non-standard, such as with the unorthodox person-number concord of certain present indicatives resembling "I does," or "he do," the regularity of the marking suggests that the student recognizes the person-number distinction despite the non-standard nature of the actual marker used. Similarly, if an essay contains a number of sentence fragments, all of which are relative clauses following a sentence ending with a substantive, then the consistency of the fragment suggests that the writer knows what a relative clause is and how to use it (though she or he may not, of course, know what the name is) and is simply marking it with a period and capital letter instead of a comma. Since in both cases the graphemic features indicate a thorough understanding of the priorities of the sentence, no student is prevented from enrolling in Introductory Composition by the use of non-standard marking alone.

The criteria for exempting a student from the introductory course do, however, recognize a priority of standard over non-standard systems of graphemic features. To be exempted, students must submit a sample that indicates they will be able to complete successfully writing tasks assigned in a wide variety of classes, and the survey of the faculty's attitudes towards their students' writing indicated that most features of non-standard graphemic systems were associated with poor writing even by those faculty members who theoretically relegated such features to a trivial status. Also, since the standard systems of grammar and punctuation were determined by the sentences of the examination that began the students' essays, failure to continue using them indicates the students' inability to perceive and adjust to the specific graphemic systems required by the examination and favored by the faculty. Consequently, no students are exempted if their essays do not exhibit the graphemic features considered standard for written English.

The application of criteria such as these is not, of course, a guarantee of consistent and accurate evaluation. The complex social context in which writing always functions complicates its evaluation, and that evaluation, in turn, must take into account the very special configuration of the testing situation itself. Examples Two and Three are eloquent comments on the effects of the examination situation on the writing that it produces. The text of Example Two is followed by this note: "please forgive the handwriting as I have a slight physical disability." Long before the last page, this essay exhibits all of the features necessary to exempt the author from Introductory Composition, but this direct address to the actual readers makes the author's precise response to the given tasks even more exceptional. This student clearly understood that the essay, as part of the assessment procedure, was to be written for two different audiences and for two purposes: the ostensible audience and purpose prescribed by the examination, and the actual audience and purpose of the assessment. As a result of conscious separation of the two kinds of constraints, the author was able to distinguish between the necessary specific response to the given tasks and a more comprehensive but unfocused demonstration of isolated, technical skills. Writers who fail to make the distinction between the two kinds of contexts in which their writing functions tend either to treat the examination simply as an occasion to display every technical skill they have or to limit their essay cautiously to the one or two tasks they know they can perform. Such efforts to perform for the assessment readers cannot, of course, bear much relevance to the purposeful essay that the examination asks for, and often these efforts to use the essay as a showcase result in texts that are internally inconsistent as well.

Example Three exhibits another feature peculiar to the situation of the examination. Because the assessment examination is offered as one of a battery of tests facing the student during a visit to the campus, most of the tests conforming to

the first type of examination described above, it is perhaps inevitable for students to assume that their writing will be evaluated according to the kind of criteria usually associated with more objective examinations. This assumption, in turn, leads them to prepare for the assessment just as they might for the other examinations. By the middle of the summer many students have thoroughly researched the topics used in earlier examinations described by their friends, and they are often determined to display their knowledge regardless of the topic they are given. (The author of Example Three apparently had prepared to write about the questions of making contraceptives available to teenagers without their parents' consent, one of the topics assigned earlier in the summer.) A similar and even more common case is the essays of students eager not only to write a good essay but to give the "right" answer. They try to incorporate into their essays all three of the statements offered in the examination as choices, even when the attitudes and perspectives of the statements are completely different. Fortunately, many of these students are able to create enough coherent units within the essay to demonstrate at least a rudimentary grasp of several basic skills, but it is virtually impossible for them to create the unified, precisely focused essay necessary for exemption from such a congeries of purposes, audiences, and topics.

Both kinds of problems are endemic to any examination which calls for a response as complicated as a written essay, and they are especially troublesome in examinations which propose to assess the students' potential to respond to tasks in situations quite different from that of the examination itself. Since in this instance, however, the problems generated by the unusual situation of the examination can be identified and distinguished from those features which are more indicative of the potential quality of stu-

dents' writing, they do not pose a substantial obstacle to the accurate evaluation of that potential.

Yet the absurdity that results from the students' misguided efforts is often so obvious that the confusion it connotes about the nature and function of writing must have serious implications for the students' university careers. Most of the samples suggest that the writers are simply unsure of what is being asked of them by an examination that requires them to use their own experience and ideas to communicate a point of view, attitude, or idea. In part, their indecision reflects their limited but genuine difficulty in distinguishing between criteria used to evaluate more simple skills associated with objective tests and those used to evaluate their essays. If this description of those criteria serves only to alleviate some of that perplexity, it will be an important step toward helping students adjust to the new tasks they will encounter at the university.

Because the criteria also testify to a general conformity among the beliefs of faculty from a wide range of disciplines about the nature of writing and its importance to all careers, this description should also provide a clearer idea of those properties of communication that are esteemed in the university as a whole and so form part of the ground on which the separate disciplines meet and share their common values. For as long as the pedagogical convenience of the traditional separation of hard sciences from the humanities stands as a justification for relegating the teaching and even existence of writing exclusively to the latter--and often to only one area, English--higher education will be merely a collection of curricular fragments rather than a unifying and liberating experience, and the humane growth generated by the exchange of ideas among the various disciplines will remain an elusive and abstract ideal.