

/THE NECESSITY OF GOOD WRITING CRITERIA TO THE TESTING OF WRITING/

by

ed

DEBBIE DIETRICH LEASURE

B. A., Kansas State University, 1982

A MASTER'S REPORT

submitted in partial fulfillment of the
requirements for the degree

MASTER OF ARTS

Department of English

KANSAS STATE UNIVERSITY
Manhattan, Kansas

1986

Approved by:

Donald C. Stewart
Major Professor

LD
2668
.R4
1986
L42
c. 2

A11202 663669

In 1976, the Executive Committee of the National Council of Teachers of English (NCTE) established a committee to explore the implications for English and reading programs of minimal competency testing in schools. The reason, as Charles Cooper explains, is that recently there has been a movement toward accountability and that has generated a renewed interest in proficiency examinations. Because society has seen a slip in skills, whether real or not, there is a call for some proof that the situation is being monitored. This need for order is not something peculiar to any one group. It is a part of human nature. In the introduction to The Mismeasure of Man, Stephen Jay Gould explains that "The mystique of science proclaims that numbers are the ultimate test of objectivity" (26). It is this emphasis on statistics that has led to the recent observation that there is a "flaw" in the educational system. Cooper points out that between 1963 and 1980, the average score on the verbal section of the SAT dropped from 478 to 424. While concern may be in order, there are some logical explanations for this phenomenon. During this time, college populations changed with many people entering that never would have gone to college before. In addition, Cooper suggests that changes in curriculum, disrupted social and political customs, increased television viewing, and broken homes also contributed to these figures. Interestingly, during this same time, scores on College Entrance Examination Board Achievement Tests rose slightly in writing proficiency. The result is that, in Cooper's words, "the general public believes that there is a lack of standards in the schools and that the high school diploma no longer has any meaning" (7).

This problem is not new. Albert R. Kitzhaber, in his 1953 doctoral dissertation, "Rhetoric in American Colleges: 1850-1900," said that the colleges in the late 19th century were suffering from an epidemic of students failing entrance exams. Kitzhaber explained that the entrance exams got started because Harvard wished to delegate to the "lower" schools the more "mechanical details" of writing. This delegating of responsibility was inspired by the European system of education, which was based on the notion of higher learning; a student should come to college well grounded in the fundamentals, free to pursue higher learning. The preparatory schools were under pressure to ready students for the entrance exams. As a result, students received a selectively narrow education, one that tried to second guess what would be on the exam. But Freshmen who did get into the university then found that they were unable to fulfill assignments-- their knowledge being so selective. Even though students were ill-prepared for the exams, the concept spread through the country. Kitzhaber related that in the 1890s everything came "to a boil." Because so many freshmen were failing entrance exams or just barely passing them, the freshman composition courses became more and more important. New enrollment had also gone up tremendously during this period, adding an extra burden to the system.

In 1891, the Board of Overseers of Harvard College appointed a committee of Charles Francis Adams, E. L. Godkin, and Josiah Quincy, to study the problems in composition at Harvard. The authors of the report argued that college teachers should not have to concern themselves with "elementary" matters (style and correctness), and it was decided that preparatory schools were to blame. Following this report, two more were issued, one in 1895 and one in 1897. The matter became a public one and in 1894, the Dial ran a long series of articles about the state of English teaching in the colleges. The consequences were both good and bad. "On the credit side, those articles centered attention on the inadequate provisions for English training in the lower schools and no doubt did a good deal to help establish English there in a prominent position" (Kitzhaber 78). On the debit side they contributed significantly to an overemphasis on "mechanical correctness."

Kitzhaber asserted that there is nothing wrong with being reminded from time to time about what we're doing with our students. But the tendency to head for "mechanical correctness" everytime a flaw in the system is seen, works against the good of the profession because it undermines the other aspects of writing that are very important. Edward White: "There is no evidence that writing quality is the result of the accumulation of a series of subskills" (Teaching and Assessing 123). Lee Odell, in "Defining and Assessing Competence in Writing," says that adopting evaluation procedures that reflect "trivial" understanding of writing will do double harm. "They can mislead us about students' skills as writers. . . and they can misdirect the work we do as teachers of writing (112).

Most professionals agree that assessment needs to be done. However, what holds us back is a general lack of agreement about what proficiency is. As White argues, "Our students rightly expect us to evaluate their writing products, and they need such evaluation in order to improve" (Teaching and Assessing 100).

Because students have the right as well as the desire to be evaluated, the solution is to continue to search for better means of assessment. Lee Odell writes, and others agree, that "the surest way to get rid of invalid assessment procedures is to replace them with something better" (134).

Deciding what would be better holds us back. In fact, establishing what it is that we're looking for, in terms of criteria for good writing, causes much disagreement. "Most faculty tend to be unaware of the many unstated criteria they assume the students will understand" (White, Teaching and Assessing, 100). Edward Hirsch, in The Philosophy of Composition, says that "We cannot get reliable, independent agreement in the scoring of writing samples unless we also get widespread agreement about the qualities of good writing. Apparently, that is the sine qua non of a solution to the assessment problem" (180).

This report will examine two well-known and several representative national standardized tests in order to understand what they measure, and more importantly, what they don't. The criteria stated or implied in these tests will be then compared to those in placement tests that have been written by people in the English profession, on a local level, and to some of the criteria for good writing established in college composition textbooks, as well. By bringing to bear some of the field's own advice about the criteria for good writing, this report will aid in future discussion of testing. After all, without these criteria, standards will remain unstated and writing tests will lack the "credibility" needed to take the place of, or be provided in addition to, multiple choice tests.

National Standardized Tests

Because national tests like the SAT and the ACT are taken by so many students and the results made so obvious in the press, it is not difficult to understand why the public becomes overly "test conscious." According to Charles Cooper, the primary goal of the tests is to predict a student's grades in the first year of college. Study books produced by the testing services echo Cooper's assertion that the tests are designed to predict student progress. Yet, it is the tendency of the public, who are not always aware of the goals of testing, to make out of these tests much more than is intended. This is where the harm is done because multiple choice tests are only able to measure a limited number of English skills. A pamphlet called Common Sense and Testing in English, written by an NCTE committee during the summer of 1974, says that "Standardized tests have traditionally dealt with the more easily measureable aspects of English instruction, to the neglect of the full range of activities involving thoughts, feelings, and attitudes" (1). The committee, made up of English university faculty, Educational Testing service professionals, and high school teachers, suggests that "tests have a way of determining educational priorities, and the increased use of standardized tests has regrettably shrunk the list of priorities available for people to consider" (1). In addition to forming the committee on testing, NCTE drafted two resolutions at its 1977 annual meeting. These resolutions, though they deal directly with competency testing, can also be useful in a general discussion of assessment. One of these regards "excessive focus on sub-skills." The delegation was "Resolved that NCTE condemn the transformation of the English Language Arts Curriculum from a holistic concern for language development to sequenced but isolated and often unrelated sets of reading and writing skills . . ." (Cooper 19) Though these comments by professionals at the NCTE annual meeting do not directly address themselves to multiple choice tests, one can see that these are the types of tests that isolate skills. It would appear from these statements that the multiple choice test, without the benefit of other forms of testing, is an inefficient measurement of students' English skills. In fact, in its February, 1979, issue

of "Support For Learning and Teaching of English" (SLATE) newsletter, NCTE indicated that the direct method of assessing writing is most appropriate because it allows adequate examination of student skills. This committee says that: "1) We have to ask students to write; 2) We have to ask readers to make judgments about the quality of students' writing; and 3) We must demonstrate that these judgments are reliable" (Meredith and Williams 11).

Despite the advice being given by professionals in the field of English, most national tests are still multiple choice. It's not difficult to see why the multiple choice test is popular. For test experts, especially, this test is cost-worthy and produces credible results, even if the results don't include all that they could. Because the National Testing Service at Princeton makes available samples of the SAT and the College Board Achievement Tests, and because these two series of tests are some of the best known of the national tests, this report will begin by focussing on these.

The SAT (Scholastic Aptitude Test)

The SAT tests examined in this report originated in a booklet, printed in 1983, of actual tests that have been administered by the College Board in years past. The introduction to the publication explains that the motive for publishing the booklet is to make the public more aware of the tests. Certainly, the Board is to be given credit for this effort.

Unfortunately, when copies of the tests fall into students' hands, the students often study them to the exclusion of broader, more inclusive kinds of things. The test then becomes, as NCTE warns, a determiner of education when used in this way, even though the study booklets do warn against this. As important as the mechanics of writing are, this tendency of students to gear themselves towards "success" on national tests, makes them forget crucial aspects of writing. As Lee Odell warns, an over-concentration on the mechanical tasks of writing will encourage the student to ignore the fact that "in order to write effectively, one must have something to be communicated and one must be able to discover what he or she wishes to say" (102). While the SAT is designed specifically to determine a student's success in college, it is not, as the booklet says, the "measure of other factors and abilities-- such as creativity, special talents, and motivation . . ." (Odell 102). It is these "special talents," however, that Odell and others are concerned with.

In the SAT study booklets are a series of sample tests, each divided into six sections, representing various subject areas. One section among the six is called the "Test of Standard Written English," an ironic title considering that it includes no writing. The questions on the Test of Standard Written English are said to measure three skill areas: use of **basic grammar** - - subject- verb agreement, pronoun agreement, and verb tense; **sentence structures** - - complete and incomplete sentences, connections between parts, etc.; and **choice of words** - - recognizing appropriateness of words used.

These skill areas are condensed into two types of questions called "**usage** questions" and "**sentence correction** questions." The **usage** questions are designed so that students identify usage errors in the sample sentences. For example, if the sample sentence reads, "Most people listen to the weather forecast every day, but they know hardly nothing about the forces that influence weather," (SAT 56) then the student is supposed to recognize that the phrase "hardly nothing" is incorrect. The **usage** questions are said to measure the "ability to recognize and use language that is clear, effective and correct according to standards of written English, the kind of English found in Composition textbooks" (SAT 29). In contrast, the **sentence correction questions** ask students to go beyond the recognition of incorrect usage, asking them to correct sentences by choosing an appropriate phrase. The instructions for this exercise read: "Select the answer that produces the most effective sentence, one that is clear and exact, without awkwardness or ambiguity . . ." (SAT 31).

A **sentence correction** question might read as follows:

Young people are not rejecting marriage, but some postponing it.

- (A) some postponing it
- (B) some are postponing it
- (C) it is postponed by some of them
- (D) it is being postponed
- (E) some having postponed it (SAT 59).

In this case, the correct answer would be (B).

Following are a few more examples of first, **usage** questions and then **sentence correction** questions:

Usage:

Until recently, Americans drank five cups of coffee for every cup
A B C
cup of tea but now they are drinking more of it. No error (SAT 57).
D E

The student would answer that (D) is incorrect, and though the student does not have to identify the reason why, this phrase is incorrect because of the unclear antecedent of "it."

Another:

In her novels, Nella Larson focused on the problems of young
A
black women which lived in Europe and America during the
B C D
1920's. No error (SAT 57).
E

The student should choose (B) because "which" is incorrect.

Sentence Correction:

Many inferior films earn a great deal of money for their producers, some extremely good ones do not.

- (A) Many inferior films earn
- (B) Many an inferior film earns
- (C) With many inferior films which earn
- (D) However, many inferior films earn
- (E) Although many inferior films earn (SAT 60).

The correct response would be (E) because there is a need for the conditional, "although."

Another:

Being as it was a full moon, the tides were exceptionally high when the storm struck.

- (A) Being as it was a full moon
- (B) With the moon as full
- (C) Due to there being a full moon
- (D) The moon was full
- (E) Because the moon was full (SAT 33).

The correct response would be (E).

In the preceding examples the answers have been fairly evident. What happens when there is some ambiguity in a choice, when several answers might work in the sentence? The example that follows reflects this problem:

A number of parents are concerned about the protection of high school athletes and the many injuries being reported.

- (A) athletes and the many injuries being reported
- (B) athletes and numerous reports of injuries
- (C) athletes because of reports of numerous injuries
- (D) athletes, but many injuries are being reported
- (E) athletes, numerous injuries have been reported (SAT 60).

One might choose (A) or (C) without a problem in sentence form. The student, though, is forced to make a decision, which usually involves trying to second guess the test examiner. The choices made are not of the writing variety but of the "out-thinking" or "second guessing" variety. The problem with multiple choice tests is that the only outlet for choice is in which answer to choose. By their design, the tests do not allow the student to make choices of his own, according to his own writing style. The problem with this is that, as Odell points out, understanding a student's judgment is a part of understanding how to assess his writing abilities.

In addition to being able to recognize error, the students must learn how to interpret and correct errors, which are not skills encouraged by these exercises. It is possible that the application of these same editorial judgments in a student's own piece of work might prove to be more worthwhile in the long run, in predicting the student's writing skills. That these exercises ask the students to become aware of language and sentence structure is good, but the fact is that they also encourage a rigid interpretation of these points of grammar, without considering that there are other options available. Lee Odell believes that "the 'textbook' definition of competence does not account for all of the choices writers make, even when they are on their best linguistic behavior" (102-3).

College Board Achievement Tests

These tests are not as well known as the SAT or the ACT, but they are used by some colleges as placement devices. They are considered in conjunction with a student's high school record, recommendations, and scores on national tests, such as the SAT, for placement purposes. English is among five general subject areas. The English tests are designed for both composition and literature. There are two versions of the composition test, including one which has 90 multiple choice questions and another with the combination of 70 multiple choice questions and a 20-minute essay. The essay version is offered only once a year, in December. The test that is exclusively multiple choice is offered three times a year. It is difficult to understand why these versions are interchangeable.

Like the SAT tests, the Achievement tests are published in a booklet for the public. In the introduction, the College Board says that the English test is designed to measure skills that are acquired throughout a lifetime. The test is designed to gauge skills like **consistency, expressing ideas logically, being clear and precise, and following conventions.** There are three types of multiple choice questions. One type tests students' "ability to detect error in underlined portions of sentences" (College Board 18). These are similar to the usage portion of the SAT.

An example,

If I am reading the editorial correct, the mayor is deliberately
A B
avoiding any discussion of the tax reform bill until after the
C D
November elections. No error (College B 28).
E

The student should choose (B), because "correct" needs an "ly" ending. The student identifies but does not correct the error.

Another type of multiple choice question is one that asks the student to identify errors in sentences and to select a revision. This type parallels the sentence structure questions on the SAT.

Many state universities continued to expand during the 1970s despite clear signs of decline in the number of 18 to 20 year olds.

- (A) despite clear signs of a decline.
- (B) even though clear signs of a decline.
- (C) with clear signs of a decline
- (D) however clearly the declines
- (E) when there was clear signs of decline (College B 31).

The student would choose (A) as the appropriate correction.

The third type of question is not seen on the SAT. The student is given an acceptable sentence and asked to rephrase it in a specific way. The student is also told that if more than one good sentence could be made, he should consider making it retain as much as possible of the original meaning of the sentence; he should choose the most natural phrasing and construction; and he should make sure that it meets the standards of written English. It appears that this exercise wishes to gauge a student's rhetorical choices, given a set of rules to follow.

Owing to her political skill, Ms. French had many supporters.

Begin with: Many people supported.

- (A) so
- (B) while
- (C) although
- (D) because
- (E) and

The student should answer with (D). The sentence then reads,
"Many people supported Ms. French because she was politically skillful."
(College B 34).

Another example:

Coming to the city as a young man, he found a job as a newspaper reporter.

Change: Coming to He came.

- (A) and so he found
- (B) and found
- (C) and there he found
- (D) and then finding
- (E) and had found (College 34).

The answer would include (B) and would read, "He came to the city
as a young man and found a job as a newspaper reporter."

This exercise, in a limited sense, asks students to make similar rhetorical choices to those they might make in their own piece of writing. Yet, as is the case with all of the exercises on national tests, the students are required to make very specific responses to very specific requests.

The essay test, given only one time a year, is a 20 minute piece of writing, requiring students to write about a chosen topic. The topic is flexible enough to allow variety. The essay is scored holistically, by two readers, unless there is some discrepancy, when a third reader is required. The booklet says that the readers look at the tests for spelling, punctuation, organization, choice of words, and "the host of other characteristics of writing that are considered" (21). It is unfortunate that the "host" of other considerations are not mentioned because these are the very criteria which this report is attempting to pin down. The purpose for the essay examination, according to the College Board, is to determine whether the student is writing on the level he will need to in college. There is a dichotomy here, because a 20 minute writing sample may be able to give some indication of a student's ability, but certainly, it cannot represent the type of writing the student will do in college, as will be seen later in this report.

The student is told that he should express thoughts clearly and effectively and that, while length is not important, one paragraph will probably not be adequate. The sample topic included is one from December 4, 1982, and it reads, "People seldom stand up for what they truly believe; instead they merely go along with the popular view" (College B 21). The directions that follow the passage are, "Do you agree or disagree with this statement? Write an essay in which you support your opinion with specific examples from history, contemporary affairs, lit., or personal observation" (College B 21). Given the time constraints that are imposed upon a student during testing, Edward White would support the notion that only one general topic should be given. He says, that given the time constraints, the student cannot be expected to have to construct his own writing assignment before he can begin. If the student has to do this, he must face, "considerable guesswork about teacher's implicit demands and concealed criteria" (105). While this essay question, given the constraints, might satisfy some English professionals, others would be asking questions about audience and the "writing situation." This is the crux of the problem. There is yet no widespread agreement upon what criteria must be met in order to conclude whether a test is adequate or not. The following discussion of writing tests, which are used for placement at some colleges, may be more helpful in ascertaining good writing criteria.

Writing Tests (for Placement)

Kathleen Kiefer, who teaches at Colorado State University, says that "if our testing reflects our emphasis on producing focused, fully detailed, and coherent texts, then we must test writing and not recognition of error or accepted usage" (1). She also points out that while many students do well on standardized tests, that does not insure that they can write well. The fact that standardized tests are not always able to predict "actual writing" abilities has made English teachers increasingly more involved in testing student's writing abilities. At the first National Testing Network in Writing conference, held in March 1982, in New York City, Marie Lederman, James Hill, and Edward White, during panel discussion, and later in published conference proceedings, assert that "English faculty must be involved from the beginning in development, maintenance, and control of testing programs" (3). White reinforces this in his recent book on assessment.

Because of what English teachers know about writing, Paul Diederich says that, "as a test of writing ability, no test is as convincing to teachers of English, to teachers in other departments, to prospective employers, and to the public as actual samples of each student's writing, especially if the writing is done under test conditions in which one can be sure that each sample is the students own unaided work" (1). The fact that writing samples are more likely to gain assent with English teachers, might explain why none of the placement tests examined in this report, rely solely on multiple choice testing. Among college writing tests, the most used method is comprised of a written sample in combination with other methods of testing.

Before examining the tests, it is important to be aware of some of the ways that tests are scored because the scoring can provide clues as to what good writing criteria are being sought out. Some methods include **primary trait** scoring of a writing sample, identifying the presence or absence of traits required by the rhetorical situation in the writing assignment; **analytical scoring**, focusing on problems which are not specific to a given writing assignment (mechanics, word choice); **discourse scoring**, using counts of T-units, cohesive ties, and other discourse units; and **holistic scoring**, in which an overall picture is gathered on the piece of writing.

Primary Trait Scoring

Ina Mullis, an associate director of National Assessment at the Educational Testing Service, Princeton, defines primary trait scoring as a method based on the theory that writing is addressed to an audience, and done for a purpose. The degrees of the writer's success are measureable in increments. In her article in Educational Measurement: Issues and Practice, she says there are three modes of discourse which have established the foundation for primary trait scoring. These areas are informational, persuasive, and literary, or imaginative writing where audience needs are different. In addition, the writer's purpose is different. Thus, the prompt used for the test must provide the student with a purpose to address and an audience to recognize. If the students' writing fails to address these aspects appropriately, their writing sample falls short. Mullis gives one example of a prompt that asks the student to write to a landlord about keeping a puppy. The assessment scale includes four levels: level one indicating little evidence of skill; level two indicating marginal evidence of skill; level three indicating competent or solid performance; and level four, indicating a very good performance. Each time the assessment scale is used, the decision has to be made about what a "very good" performance is. This is sometimes done by comparing the essays in one group of papers and setting "range finders" by which to measure all other papers. A college is likely to adapt this type of testing to suit its own needs in terms of student populations and levels of classes. This form of scoring places value on a student's ability to recognize a purpose, and adjust his ideas for a particular audience.

Analytic Scoring

In analytic scoring Ina Mullis mentions that single characteristics of writing are identified and each is rated according to quality. She cites a system of Paul Diederich's where eight items are under consideration, these being "ideas, organization, wording, flavor, usage, punctuation, spelling and handwriting" (3). Each of these is rated on a five point scale and ratings for ideas and organization receive a double score. Because these criteria have been singled out, it isn't difficult to see what good writing involves, in this framework. It is encouraging to see that in Diederich's system, anyway, ideas and organization are emphasized over the mechanical details of writing. The University of Illinois has a system with five areas of assessment which include focus, measuring clarity with which the writer presents ideas; support, measuring the detail and reasoning; organization, measuring logical flow of ideas; mechanics, measuring standard English skills; and an overall "global" judgment of the composition. Each characteristic is graded separately on a six point scale. Once again, this is an adaptable system and varies from college to college.

Mullis emphasizes that results are usually better if the writing samples are directed toward particular purposes and audiences. She admits that often with analytic testing the concern centers on characteristics of grammar and usage, these being the easiest to rate, over such things as focus and organization.

Discourse Scoring

This type of scoring centers on the use of T-units, cohesive ties and other discourse units for assessment. The "T-unit," the most well-known of the discourse scoring tools, is an index of syntactic complexity and was developed by Kellogg Hunt, in 1964. As defined by Hunt, the T-unit is, "a main clause plus all subordinate clauses and nonclausal structures attached or embedded in it." The discourse approach to testing came about as a result of much research in composition studies that was being done in the 60s and 70s. Stephen Witte points out that in the 60s, people such as Francis Christensen, Kenneth Pike and Willis Pitkin were studying levels of discourse, including functional slots, semantic levels of generality, discourse "blocs," and relationships across sentence boundaries. Each "bloc" and each level of generality varies from theory to theory. One can deal with a piece of discourse as small as a t-unit or as large as a paragraph. One advantage to this approach is that no longer does one have to consider writing in terms of sentences only.

Tests that measure t-units, as well as other discourse approaches, are like the multiple choice tests, in that they break writing down into measureable pieces, but unfortunately the pieces don't add up to the whole picture of writing ability or success. It is not difficult to see why holistic grading has become a mainstay in assessment among English professionals because it seeks to measure the whole piece of writing, even if some of the aspects that are being measured remain unnamed.

Holistic Scoring

Edward White says the philosophy behind "holisticism" maintains that "writing must be seen as a whole, and that the evaluating of writing cannot be split into a sequence of objective activities, because holisticism reinforces the vision of reading and writing as intensely human activities involving the full self" (Holisticism 409). While holistic scoring does provide a philosophy that is agreeable to many English professionals, because it does account for the humanistic nature of writing, there are still problems. Davide Charney asserts that "The issue of criteria is central to the problems of holistic ratings as a valid means to evaluate writing ability" (79). Good writing criteria must be available for holistic scoring to be successful. However, Miles Myers explains that even better than having the characteristics of a good piece of writing, is to have a good piece of writing available to compare the students' writing to. He cites Paul Diederich's success with using this approach. While it is undoubtedly easier to use models for grading, without having the criteria for "good" writing available, there is always a question of validity and there is always a question of what is being measured.

Another problem with holistic grading, as Ina Mullis points out, is that while holistic scoring can provide information about the range of overall writing quality exhibited, it cannot provide specific prescriptive or diagnostic information. Edward White would agree that the best test, will engage the student on not only an "individual" but also a "social" level, where both overall writing quality and mechanics are valued and measured. White goes on to say that, "we should attempt to measure anything we call 'writing ability' by more than one writing sample and in more than one writing model" (Assessment 118). Despite the flaws with holistic grading, it is undeniable that the philosophy upon which it is based, will continue to be important in the assessment issue because assessment research continues to lead towards examination of writing as a whole picture approach.

The Tests

These tests have been chosen primarily because the results have been made accessible to the public in articles, and because they represent the variety of tests being done today. The order in which the tests are presented and the choice of tests, is not significant. The purpose of this examination is not to compare the tests to one another, so much as it is to try to locate good writing criteria based in the tests themselves. Because of the nature of assessment, the fact that it requires ongoing studies into the best methods available, many of these tests are still in the development stages.

Colorado State University's Placement Test

This test is based solidly on holistics. Not only is the writing graded in this way but also the student himself is evaluated holistically. The program that Kathleen Kiefer describes includes a pre- test and post-test, each written during a one hour session. The students are given a rhetorical situation, such as writing about a teacher that they admire or dislike. The pre- and post- tests are graded at the same time, with no distinctions made, and the teachers use a holistic scale, considering six features of the writing. Unfortunately, these six items are not made available in the article. Even without access to the scale though, some aspects of the test situation are clear. For example, the test is considered along with all of the work the student has produced during the semester, as well as his attitude and perceived achievement, so that the criteria that are established for good writing not only encourage viewing the writing as a whole but also the individual as a whole. After all of the results are evaluated, the students are placed, during the second semester of school, into various classes, according to their needs. One must bear in mind, that understanding students as individuals is a luxury that national standardized tests do not have but that teachers, because they're in contact with each student, can take advantage of in assessing them.

The Ontario Test of English Achievement

Unlike the Colorado State test, the Ontario Test of English Achievement was designed for all of the universities and colleges in Ontario, with each university individualizing its use. In addition, this test has several components, including a language and skills section and a writing sample. With the wide variety of students entering the universities, it was decided that skills had to be assessed before classes could be formed. Since the committee perceived that there were more aspects of English than could be tested, they chose four components they thought to be critical to success in college and these are, reading ability, writing ability, ability to plan and develop a theme, and competence in using mechanics of language. The latter three sections of the test will be the focus here. In the thesis development section, there is a series of multiple-choice questions, with a four options response. The questions ask students to identify the essay outlines, supporting arguments, theses, and topic sentences. The committee explains that because some people feel that "writing flows naturally and should not suffer obstruction as a result of rules of format, structure, etc.," (Evans 7) this test will not seem appropriate to some. After heated debate on whether or not to include this section on the test, it was decided that, "... it is not that all good writing follows a set of implicit or explicit rules; rather it is that the majority of university students benefit from knowing how to structure a formal essay" (Evans 7). In the text of the report, the authors do not allude to what makes up a "formal essay." One might call to mind a piece of writing with an introduction, paragraphs with thesis statements and support, and a conclusion that restates the thesis, but one cannot be sure without seeing the prompts.

The language section of the test consists of 40 multiple-choice items administered during 25 minutes. The four sub-sections of the test include spelling, grammar, diction/usage and sentence structure. The writing section provides a topic that is expository in nature, allowing for persuasion or argumentation, that is written in a 50- minute period. After some experimentation with different rhetorical situations, the committee chose three different prompts for students to write in response to. These prompts are not included in their report. The essays are scored holistically on a scale of one to ten by two scorers. The writing prompts and the criteria involved in this scale are not mentioned in the report.

What is especially interesting about the report on this test is that the committee thought it important not only to show what criteria are being measured, in the various sub-tests, but also to show that a body of criteria had to be left out in the interest of time and resources. Clearly, the committee is not unaware of the criteria they felt it most important to include. The first sub-test, that focusses on thesis development, emphasizes the importance of having something to say, knowing what it is and being able to recognize whether the idea is being fully supported or not. This might be captured under the single word "focus." The second sub-test measures the mechanics of writing, and is not unlike the SAT, or any other standardized test, but is at least not used exclusively. The third section, with a writing sample, is difficult to judge because the scale for holistic scoring is not made public. What is clear, though, is that the test designers see value in the students having a rhetorical situation to which the purpose and the audience are outlined and the appropriateness of the student's written response to these is measured. The committee gives students a prescribed topic to follow rather than turning them free to design their own writing situation.

New Jersey Basic Skills Placement Test

All freshman entering New Jersey Public colleges are required to take the New Jersey Basic Skills Placement Test, which was first introduced in 1978. The test, primarily multiple choice, also includes a 20-minute essay question. Many private colleges have elected to use the test as well. Like the Ontario test, this is not an admissions but rather a placement test. It consists of five multiple choice sections with a 20-minute free-response essay question. The responses to the essay question are scored by two readers on a four point scale. The composition score is composed of scores from the sections on sentence structure, logical relationships, and the essay. An example of an essay response question is, "The communities and neighborhoods we live in have various concerns, such as high taxes, repeated power failures, water pollution, the clearing of roads in winter, new highway construction, new parks and playgrounds, better street lighting, or increased police and fire protection.

"In a well-organized essay, write about a concern in your community or neighborhood, preferably one that has affected or may affect you, etc" (Hecht 2). What is meant exactly by "well-organized" is not made clear in the report.

One can see that the test allows the students to choose something that is particularly important to them. The goal here is, as Frank D'Angelo says, to let the student find his own pressing matters to write about. Indirectly, the student is provided with an audience to write to, because it will be those who also live in her community, that will share in the concern being shared by the student.

The logical relationships section of the exam asks students to categorize ideas, use appropriate connectives, make analogies and recognize principles of organization through multiple choice testing. These are primarily criteria of coherence and organization, which have been seen before.

College's Placement Test

Joseph Harris, the spokesman for Parkland community college, explains that the testing occurs during a two-week class, English 100, and after two weeks, the administrators total the results and place students in the appropriate classes. During the class students are subjected to an essay test, skills tests and a cloze test. Students are exempt from the class who have successfully completed English 099. During the first day of classes, students write what is called a "practice paragraph," which is handed in and returned before the student writes the diagnostic test paragraph, on the fourth day of the course. An example of a topic is the "advantage of apartment living," or the "disadvantage of apartment living." The students are asked to underline the sentence which states the main idea of the paragraph, and to list only one single supporting idea for the paragraph. Finally, a suitable conclusion is to be reached. For the diagnostic paragraph written in class, students are given a topic, are asked to develop their ideas, and before turning in the paper, are to number each sentence in the essay. There are three major areas that are examined in the paragraphs: content and structure, grammar and mechanics, and style. After the papers are examined, they are graded as either unacceptable, acceptable, or superior.

Besides writing a practice and final paragraph, students take what is known as a "cloze test," which asks them to fill in the blanks in a passage from which every fifth word has been removed. It is thought that the "cloze" test will provide students with practice in making rhetorical choices. They do this by deciding which words will complete the sentences already provided. Following this, they are asked to complete an exercise where they look at thesis statements and paragraphs that are "successful" and compare them to paragraphs that are not successful and ascertain the reason. A paragraph is successful, according to the directions, if "the topic statement states the main idea and is general; the body of the paragraph develops the topic sentence by giving specific information and examples; and the concluding statement brings the development of the main idea to a satisfactory close" (J. Harris 16). The students write out answers to such questions as "Why did this topic sentence fail to capture the rest of the ideas in the paragraph?"

These concepts about successful paragraphing, presented in this two- week course at Parkland Community College, are present also in the paragraph that Alexander Bain described over a hundred years ago. There is nothing wrong with keeping a system if it works, no matter how old it is, and one can see uses for the Bain paragraph. The problem develops when students believe that this type of paragraph, the "theme in miniature," is the only acceptable means of putting ideas on paper. If one picks up any number of books, articles, or essays, he can see that most paragraphs are exceptions to Bain's rule. There are paragraphs that have no "topic sentence" in the pure sense and there are paragraphs that are one sentence long. What results from a narrow concept of paragraphing is that even if a student is given a "free" topic to write on, he will be unable to let his ideas dictate style. In this sense, telling someone how to paragraph limits his rhetorical choices.

The last test that students take is an English Usage test, very much like that in many of the standardized tests, where they're asked to identify proper English usage.

Placement Test

Motivated by the reported success of using the cloze test to measure non-native speaker's English abilities, the authors of the report on Texas A. & M.'s placement test, Helmut Esau and Carlson Yost, are in the stages of designing a cloze test for native speakers of English. Their version of the cloze test removes twenty-five words from a prose passage. The items deleted are chosen randomly. This test was first administered to one hundred fifty first and second semester freshman students during 1977 and 1978. In addition to taking this test, students wrote an essay on a topic related to the cloze passage. Some of the essays were class assignments in English that ended up being used for placement as well. The essays were graded holistically, by two readers, with the scores of high pass, pass, marginal pass, and fail. These categories would be determined somewhat, by the papers that are turned in to the readers. They compared the results of the cloze tests to the results with the essays and discovered that there was no correlation in scores. Their conclusion was that, with the cloze test, they were not measuring "students' overall writing ability, but rather some quantity like verbal ability, which is necessary but not sufficient to write a passing essay" (Esau 5). In conclusion, the skills measured by the cloze test, identifying patterns and proper usage, are only a small part of the overall skills needed for a good piece of writing. These other skills are what we continue to look for.

Ferris State College's Placement Exam

During orientation period in the summers, incoming freshmen at Ferris State College write an essay that is used for placement. This test is then graded holistically. John Alexander and Fred Swartz, the test designers, hope that having incoming freshmen write, before they are even officially students, will give the public the message that writing rather than merely correct usage should be the judge of English skills. Alexander, who is the head of the Department of Languages and Literature, also says that this evaluation process is a way of getting English teachers involved in assessment, on a larger scale. One of the conclusions drawn, when establishing this assessment program, was that a writing sample alone was insufficient to judge abilities. Some of the problems encountered with the holistic grading of writing samples, were that average samples received higher ratings when preceded by a block of poor essays, and handwriting actually became influential in determining writing ability. The biggest problem was agreement on "good writing" criteria.

After these initial difficulties the group decided to attempt computer analysis of writing samples, when such discrepancies were found in the holistic evaluation. T-units were introduced into the assessment procedure because of the conversion that could be made to computer analysis. ACT scores, holistically graded writing samples, and evaluation by computer, were all combined to try to ascertain students' abilities, but nothing was resolved except that, as the group decided, "In the minds of most members of the English profession, placement decisions cannot rest solely on scores gathered from national testing practices of the SAT and ACT. . . on the other hand, for those specialists in educational administration, writing samples will not gain easy acceptance since for all the effort and time expended no studies have as yet appeared revealing any significant increase in placement accuracy" (Alexander 12). Certainly, without knowing what it is we're looking for, any form of assessment will fall short. The one clear conclusion established by this report, is that a standardized test does not reflect a student's writing potential.

Conclusion

The majority of these tests provide a for a writing sample with some other form of testing. This says two things: first, that the English professionals that design placement tests, want to test writing ability with a writing sample; secondly, that because there is not yet consensus on what should appear in that writing sample, the testers back these up with other sorts of tests. The evaluation of these tests together, does provide an outline, a rough sketch, though, of a few good writing criteria. Among these criteria are a need for focus, organization, audience and clarity of purpose. Indeed these criteria should be held in equal if not superior importance to mechanical detail. A minority, unfortunately, still place a large emphasis on mechanics and correctness. Most English professionals would agree, that mechanics and correctness are only a small part of the overall picture of writing success. Making sure that writing skills are tested by writing samples is an important step in the right direction.

Composition Textbooks

Edward White says that "... writing research, like research in any emerging field, is extremely important ... The portion of it that turns out to be significant, and no one can quite be sure which research that will be, becomes the new knowledge of the discipline, determining its textbooks as well as its directions" (Assessment 171). Each textbook includes many criteria for "good writing." The predominant criteria tend to be duplicated across the texts. With the many threads running through composition research, two are at once apparent. One of these is a renewed interest in the principles of ancient rhetoric, and with that an interest particularly in audience and in the way ideas are generated. This interest has led to a focus on invention techniques, or heuristics, by which students can generate ideas for writing. Another major thread involves the philosophy that writing is a "process," and as such, a writer goes through a prewriting, writing and revising stage sometimes simultaneously, sometimes not. The "process" approach is connected with the rhetoric to the extent that invention is important in both cases.

Both of these threads are seemingly at odds with assessment, as we now know it, because each calls for the opportunity to plan, think, write and rewrite, actions which are not provided for on limited, timed exams. With this theory in mind, some English teachers, such as Russell Meyer, have tried to come up with tests that allow for a writing environment that includes time for invention and revision. Meyer, who teaches at the University of Missouri at Columbia, has been involved in a study in the last few years, that allows placement exams to be written at home, under the same conditions that a normal composition assignment would be written. Of course, the obvious problem with this approach, is that you run the risk of a student using someone else's work. If this problem can be solved, it might provide an interesting alternative.

The interest in these two threads of thought have given a noticeable uniformity to the purposes with which many authors have approached writing composition textbooks in recent years. Some of the predominant criteria will be categorized and divided, with authors appearing in the various appropriate categories.

The Writing Situation

The "writing situation" simply means that a writer must have a purpose for writing. This sounds simplistic but, in fact, it is an idea that is ignored in many writing assignments and situations. Because one must have the means of finding ideas and situations, invention theory became important.

Process and Thought in Composition, by Frank D'Angelo, tells the student that "a writing situation must take into account a need (an urgency), the occasion, the place, the writer, the writer's purpose, etc . . . " (4). Janet Emig, Janice Lauer, Gene Montague and Andrea Lunsford would agree with D'Angelo, and their own advice to students is that "primary to meaningful writing is the discovery and communication of a significant focus to which aims and modes are subordinate . . . that meaningful writing grows from a writer's own pressing questions" (xvi). Because it is the nature of tests to be timed and restricted, they rarely afford students the opportunity for free choice and consequently for "immediacy." One can understand a testers' motivation for dictating a uniform assignment because of some students' tendency to sidestep challenging themselves. Even though avoiding a challenging topic, the unambitious student may be as equally rewarded as the ambitious student. This translates into a testers' nightmare, because of the obvious discrepancies in the level of difficulty between two essays. If a student cannot respond well to the given writing situation, there is a need for him to have the opportunity to write several essays, addressed to various writing situations and topics. Edward White, among others, supports this view.

Writing to Audience

Perhaps because of the new - found interest in the ancient tradition of rhetoric, audience is a major focus in many recent textbooks. The audience is to be moved, changed and ultimately persuaded to adopt the writer's perspective. In fact, it might seem that when anyone bothers to communicate an idea, the hope is that someone will be changed. One can reach an audience in a variety of ways, such as with intellect, with personality, or with language. In the second edition of The Contemporary Writer, W. Ross Winterowd tells students that "The writing instruction in the book is based on the simple but crucial notion that a writer conveys meanings and intentions through language structures to an audience of readers" (vi). Jeff Rackham simply says that "writing is an act of communication with a reader." Linda Flower, in Problem Solving Strategies for Writers, suggests that if writers are offered writing options for addressing readers in the real world, the result is that "such writing . . . goes beyond traditional persuasions and means helping readers to comprehend and remember fully what you have to say" (vi). Flower also emphasizes that good writing is functional and goes beyond "mere" correctness to meet the needs of the reader. When readers are impatient, language difficulty may distract them from content. Writers should make their prose "easy to read, their points easy to see, and their arguments logically presented" (2). The difficulty for Flower's readers is that she writes about a "logically presented argument," but does not provide the criteria necessary for understanding what she means.

Audience is so important that, some teachers would say that the success of a piece of writing is based entirely upon audience response. Barbara Walvoord mentions that "writing is effective when it accomplishes its purpose for its audience" (2). And some say, that if the language is not precise, the effect will be lost for the audience. Leggett, Mead and Charvat, in the eighth edition of the Prentice-Hall Handbook For Writers, 1982, push students to consider language especially: "Written English communicates through the precision of its diction, the orderliness of its sentence and paragraph structure and the relative fullness of detail. If it is to be taken seriously by a general audience, it must observe the conventions of spelling, punctuation and grammar. In short, writers do conform to certain standards to get their meaning across" (9). Because audience is important, both in the classroom and in testing the students must be made aware of an audience they are writing to and their response to this audience should be noted.

Writing as Discovery

The value that writing has for the individual is a difficult thing to weigh indeed, especially because the growth goes on subconsciously most of the time. Yet, it is a major trend in textbooks to direct the writer towards writing for learning, and for self evaluation, as well as for communication. The "process" approach to writing, which encourages a writer's self growth, derives from the work of D. Gordon Rohman at Michigan State. Rohman's work is described by Richard Young in "Invention: A Topographical Survey." Rohman's defines the concept of pre-writing, which like classical invention, views the composing process as "linear and mentalistic." This concept presumes that, "... Different things happen at different stages in the process of putting things into words and words onto paper. . . . We divided the process at the point where the 'writing idea' is ready for the words and the page: everything before that we called 'Pre-Writing,' everything after 'Writing,' and 'Re-Writing. . . ." (Young, 17). In the "process" philosophy, the writer's exploration of his own feelings and attitudes, in the "prewriting stage," is an important part of the communication. In accordance with Rohman's approach, Janet Emig and others say that "meaningful writing grows from a writer's own pressing questions" (xvi). Frank D'Angelo is emphatic about self-growth saying that even if writing were not an important social or public activity, it would still be of value for the individual because it "facilitates thought." He says that writing can "enable you to make connections, see relationships, deepen perception, solve problems and give order to your experience" (3). Anne Ruggles Gere explains that writing can help students see things in themselves that they didn't know they knew. Linda Flower says that because writers build their own representations of meaning, "writing is a form of thinking." In testing, it is difficult to measure this aspect of writing -- in fact, I suspect that this is the area of "good writing" that eludes us most because measuring growth in someone's inner self is something none of us is able to do.

Revision in Writing

This is another area difficult to measure in testing because, up to now, very few tests have allowed students the luxury of revising. White points out that any effort to create a test that allows revision has failed because many students did not know what to do. They set about recopying word for word the text in front of them. This becomes a vicious circle, because even though revision has worked its way into the philosophy of composition, and it is very definitely encompassed in the "process" approach, it has not been around long enough for most students to know what to do. Nevertheless, many of the textbooks place great emphasis on revision. One of the ways that the revision process takes place has been described, and Linda Flower uses this in her textbook, as the transformation of "writer-based prose" to "reader-based prose." That is, sometime during the process of writing, after the writer has established with himself what it is he wants to say, he transforms the piece of writing into something an audience can read and understand. With most writers, this happens through constant revision. Charles Cooper and Rise Axelrod mention that "Experienced writers know that strong writing does not always emerge in a first draft. Writers are revisers" (5). Gere suggests that if one accepts the connection between writing and learning, "it makes it impossible to learn to write all at one time" (26). One can apply this to the revision process, when each time the writer learns to write a little better. Finally, Lenora Woodman and Thomas Adler suggest that because writing is "a dynamic process that weaves back and forth between thought and words . . . revising occurs during the actual composing process, as well as after the first draft" (7).

Most teachers realize, that even with the emphasis placed on revision, many students sit down the night before a paper is due and write the final draft. It will not be until revision is firmly established in composition, and students see the benefit in revision, that it will be transferred successfully to testing.

Conclusion

The recent interest in competency testing has revived some old questions about testing, questions that surfaced as long as a hundred years ago at Harvard. While it is sometimes tempting to throw up our hands and give up the issue of testing writing completely, as Edward White explains that isn't possible because assessment is expected and desired by students and teachers alike. Likewise, White asserts that a teacher who concerns himself with assessment, will not only learn a lot about testing but also about what he's asking students to do with their writing. The point is, if we can take what we already use in the classroom, and apply it to new testing situations as they develop, we will test what we want to teach, instead of teaching what we want to test.

In addition to this, everytime the assessment crisis is renewed, we must fight the urge to go "back to basics," if what that means is testing correctness only. The better approach is to make assessment a part of ongoing research so that when crisis comes, there is more to work with than multiple choice tests. As has been pointed out in this report, multiple choice tests are insufficient measures of English skills when used alone. Instead, tests should include a writing sample, and in addition to this, tests should include more than one type of measurement. After all, writing is a complex process, with many components that contribute to its "success." Still, the problem of how to judge the writing sample persists. We have discovered, that while the holistic approach has its flaws, it is a step in the right direction because it doesn't seek to reduce the piece of writing to a series of "sub skills." In addition to grading holistically, experts have also found that the next step must be to look at the person holistically, and that means evaluating several pieces of writing. This is valuable from the standpoint that one can measure a student's ability to recognize different audiences, different purposes and different modes of writing.

Beyond this, understanding the criteria for good writing adds all the more credibility to holistic scoring. However, even though these change and our opinions about writing change, it is still possible to administer writing tests. As long as students and teachers are aware of what tests are measuring, how the results will be used, and what the limitations are, testing can still be used effectively. In order to improve assessment, however, collections of writing samples should be added to the tests, and, more importantly for the long term, assessment research must continue.

All of this is not to downplay the importance of usage and grammar. In fact, there is no doubt that grammar and style are crucial for they serve as the package for our ideas; however, if that package is empty, it doesn't matter how pleasingly it is wrapped. If a student has something to say, and he is aware of the audience he wants to reach, we must count the message even more crucial than his delivery. This is a convincing enough reason to test writing with writing, and to continue to better understand what makes writing successful.

Works Cited

- Alexander, John, and Fred Swartz. "The Dynamics of Computer-Assisted Writing Sample Measurement at Ferris State College." Paper presented at the Annual Meeting of the Michigan Council of Teachers of English. East Lansing, Michigan. 29-31 October, 1982.
- Axelrod, Rise B., and Charles Cooper. The St. Martin's Guide to Writing. New York: St. Martin's Press, 1985.
- Baker, Sheridan. The Complete Stylist, 2nd ed. New York: Thomas Crowell, 1972.
- Berke, Jacqueline. Twenty Questions For The Writer, 4th ed. New York: Harcourt Brace Jovanovich, 1985.
- Brooks, Cleanth, and Robert Penn Warren. Modern Rhetoric. New York: Harcourt Brace Jovanovich, 1972.
- Brossell, Gordon. "Rhetorical Specification in Essay Examination Topics." College English 45 (1983): 165-73.
- Brossell, Gordon, and Barbara Hoetker Ash. "An Experiment with the Wording of Essay Topics." College Composition and Communication 35 (1984): 423-25.
- Charney, Davida. "The Validity of Using Holistic Scoring to Evaluate Writing: A Critical Overview." Research in the Teaching of English 18 (1984): 65-81.
- Cohen, Arthur. "Assessing College Students' Ability to Write Compositions." Research in the Teaching of English 7 (1971): 356-71.
- Cooper, Charles, ed. The Nature and Measurement of Competency in English. Illinois: N. C. T. E., 1981.
- D'Angelo, Frank. Process and Thought in Composition, 3rd ed. Boston:

- Little, Brown, 1985.
- David, Carol, and Thomas Bubolz. "Evaluating Students' Achievement in a Writing Center." Paper presented at the Annual Meeting of the Writing Centers Association. West Lafayette, Indiana, 5-6 May, 1983.
- Diederich, Paul B. Measuring Growth in English. Illinois: N. C. T. E., 1974.
- Emig, Janet, Janice M. Lauer, Gene Montague, and Andrea Lunsford. Four Worlds of Writing, 2nd ed. New York: Harper & Row, 1985.
- Esau, Helmut, and Carlson Yost. "Measuring Writing Ability With the Cloze Test is not Closed." Paper presented at the meeting of the Linguistic Society of America. Urbana, Illinois, summer, 1978.
- Evans, P., and others. "The Ontario Test of English Achievement. The Final Report on its Development." Toronto: Ministry of Colleges, 1983.
- Flower, Linda. Problem Solving Strategies For Writing, 2nd ed. New York: Harcourt Brace Jovanovich, 1985.
- Gere, Anne Ruggles. Writing and Learning. New York: Macmillan, 1985.
- Gere, Anne Ruggles. "Written Composition: Toward a Theory of Evaluation." College English 42 (1980): 44-58.
- Gorrell, Donna K. "Defining the Basic Writing Student by Count." Paper presented at the Annual Meeting of the Conference on College Composition and Communication. Dallas, Texas, 26-28 March, 1981.
- Gould, Stephen Jay. The Mismeasure of Man. New York: W. W. Norton, 1981.
- Harris, Joseph B. "English 100 Composition Workshop." Illinois: Parkland College, 1977.
- Harris, Maverick. "Texas English Teachers' Perception of Entering Fresh-

- mens' Writing Skills." Paper presented at the Annual Meeting of the Texas Joint Council of Teachers of English. Fort Worth, Texas. 4-6 February, 1982.
- Hecht, Lawrence W. "Validation of the New Jersey College Basic Skills Placement Test." Princeton: Educational Testing Service, 1980.
- Hirsch, E. D. Jr. The Philosophy of Composition. Chicago: University of Chicago Press, 1977.
- Homburg, Taco Justus. "Holistic Evaluation of ESL Compositions: Can It Be Validated Objectively?" TESOL Quarterly 18 (1984): 87-105.
- Kiefer, Kathleen E. "Testing Basic Writers' Proficiency: An Effective Model." Paper presented at the Annual Meeting of the Wyoming Conference on Freshman and Sophomore English. Laramie, Wyoming. 27 June- 3 July, 1983.
- Kitzhaber, Albert R. "Rhetoric in American Colleges: 1850-1900." Diss. U of Washington, 1953.
- Leggett, Glenn, David Mead, and William Charvat. Prentice Hall Handbook for Writers, 8th ed. New Jersey: Prentice Hall, 1982.
- Meredith, Vana Hutto, and Paul L. Williams. "Issues in Direct Writing Assessment: Problem Identification and Control." Educational Measurement: Issues and Practices 3 (1984): 11-15.
- Meredith, Vana Hutto. "Writing Assessment in South Carolina: Past and Present." Educational Measurement: Issues and Practices 3 (1984): 19-21.
- Meyer, Russell. "'Take - Home' Placement Tests: A Preliminary Report." College English 44 (1982): 506-18.
- Modu, Christopher, and Eric Wimmers. "The Validity of the Advanced Placement English Language and Composition Examination." College

- English 43 (1981): 609-20.
- Mullis, Ina. "Scoring Direct Writing Assessments: What Are The Alternatives?" Educational Measurement: Issues and Practice 3 (1984): 16-18.
- Myers, Miles. A Procedure For Writing Assessment and Holistic Scoring. Illinois: NCTE, 1980.
- National Testing Network in Writing. Proc. of a Conference on Writing Assessment. March 1982. New York: City University of New York, 1983.
- Odell, Lee. "Defining and Assessing Competence in Writing," The Nature and Measurement of Competency in English, ed. Charles Cooper. Illinois: NCTE, 1981, 95-138.
- Prater, Doris, and William Padia. "Developing Parallel Holistic And Analytic Scoring Guides For Assessing Elementary Writing Samples." Journal of Research and Development in Education 17 (1983): 20-24.
- Purves, Alan, Beryl Baily, Robert A. Bennett, and others. Common Sense and Testing in English. Illinois: NCTE, 1975.
- Purves, Alan. "In Search of an Internationally Valid Scheme for Scoring Composition." College Composition and Communication 35 (1984): 426-38.
- Quellmalz, Edys S. "Toward Successful Large - Scale Writing Assessment: Where are we now? Where do we go from here?" Educational Measurement: Issues and Practices 3 (1984): 29-32.
- Rackham, Jeff. From Sight to Insight, 2nd ed. New York: Holt, Rinehart and Winston, 1984.
- Ruth, Leo, and Sandra Murphy. "Designing Topics For Writing Assessment: Problems of Meaning." College Composition and Communication 35

- (1984): 410-22.
- Shearer, Ned A. "Alexander Bain And The Genesis of Paragraph Theory." Quarterly Journal of Speech 58 (1972): 408-17.
- Sternglass, Marilyn. "Assessing Reading, Writing and Reasoning." College English 43 (1981): 269-75.
- Strunk, William, and E. B. White. The Elements of Style, 3rd ed. New York: Macmillan, 1979.
- Ten S. A. T. S. New York: The College Board, 1983.
- The College Board Achievement Tests. New York: The College Board, 1983.
- Yopat, James. Comment and Response. "Do I Dare Disturb the ETS?" College English 44 (1982): 532-39.
- Yopat, James. "Going APE: Reading the Advanced Placement Examination in English Composition and Literature." College English 43 (1981): 284-92.
- Walvoord, Barbara. Writing Strategies For All Disciplines. New Jersey: Englewood Cliffs, 1985.
- White, Edward. "Holisticism." College Composition and Communication 35 (1984): 400-409.
- White, Edward. Teaching And Assessing Writing. San Francisco: Jossey-Bass, 1985.
- Winterowd, W. Ross. The Contemporary Writer, 2nd ed. New York: Harcourt Brace Jovanovich, 1981.
- Witte, Stephen P. "Topical Structure and Writing Quality: A Study of the Argumentative Texts of College Writers." Paper presented at the Annual Meeting of the American Educational Research Association. Montreal, Canada. 11-15 April 1983.
- Woodman, Lenora, and Thomas P. Adler. The Writer's Choices. Illinois:

Scott, Foresman, 1985.

Young, Richard E. "Invention: A Topographical Survey." Teaching Composition: 10

Bibliographical Essays. ed. Gary Tate (Texas: Texas Christian U. Press, 1976) 1-43.

THE NECESSITY OF GOOD WRITING CRITERIA TO THE TESTING OF WRITING

by

DEBBIE DIETRICH LEASURE

B. A., Kansas State University, 1982

AN ABSTRACT OF A MASTER'S REPORT

submitted in partial fulfillment of the

requirements for the degree

MASTER OF ARTS

Department of English

KANSAS STATE UNIVERSITY
Manhattan, Kansas

1986

ABSTRACT

Testing has become an inevitable part of the educational system because of our need to rank and verify students' abilities. With the importance placed upon tests, students and teachers must know the criteria being tested. Good writing criteria are not yet established, which makes testing difficult, unless one chooses to rely only upon mechanical matters. Mechanics are easily evaluated on multiple choice tests. However, a rising number of English professionals now designing and implementing placement tests argue that a test measuring writing skills must include writing samples. The present situation reveals a dichotomy between criteria stressed in classroom writing and criteria stressed in testing situations. This is a frustrating, perhaps even a defeating situation for students trying to learn to write.

Two national tests, the SAT and the College Board Achievement Test, which are primarily multiple choice, are reviewed here. They contrast significantly with placement tests, designed and implemented at various universities by English faculty. These tests are composed of a combination of measurement devices, namely writing samples and multiple choice questions. The designers of these tests, while not agreeing on the criteria of good writing, have established that writing criteria cannot be measured by either multiple choice or written test alone. The last section of the report examines the criteria for good writing presented in composition textbooks and identifies those areas where a consensus exists.

This study yields the following conclusions:

- 1) It must be understood that without some kind of consensus about good writing criteria, there will be no test of writing that is valid. Professionals should work together to establish a common set of criteria.
- 2) Teachers of English must extend their involvement in assessment to insure that writing tests adequately test the criteria of good writing.
- 3) Because writing is such a rich and complex task, testing methods by themselves will always be unable to completely determine a writer's ability.