

DOCUMENT RESUME

ED 343 156

CS 213 241

AUTHOR Cosgrove, Cornelius
TITLE Text against Text: Counterbalancing the Hegemony of Assessment.
PUB DATE 21 Mar 92
NOTE 13p.; Paper presented at the Annual Meeting of the Conference on College Composition and Communication (43rd, Cincinnati, OH, March 19-21, 1992).
PUB TYPE Speeches/Conference Papers (150) -- Viewpoints (Opinion/Position Papers, Essays, etc.) (120)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Discourse Analysis; Higher Education; Statistical Inference; *Student Evaluation; Testing Problems; *Writing Evaluation
IDENTIFIERS Academic Discourse; Academic Discourse Communities; Composition Theory

ABSTRACT

A study examined whether composition specialists can counterbalance the potential privileging of the assessment perspective, or of self-appointed interpreters of that perspective, through the study of assessment discourse as text. Fourteen assessment texts were examined, most of them journal articles and most of them featuring the common methodology of educational testers--statistical inference. Critical strategies employed were: closely examining the language in which assumptions about writing was couched; considering the scholarly authority used to support an assumption; juxtaposing an assumption made by an assessor with a representative text from the composition discourse community; analyzing the "works cited" list of a discourse; and habitually seeking definitions for those qualities or abilities assessors have set out to measure. Conclusions, admittedly impressionistic, include: (1) it is most fruitful to critically examine the assumptions about writing which inform the tests in question; (2) three of the assessment texts had no citations familiar to someone who has been reading composition literature for almost two decades; and (3) many assessors are less definite and assertive about the process of evaluating writing than composition specialists have feared. Writing teachers could be handicapped by a commitment to individual perception and a reputation for idiosyncratic grading unless composition specialists can help their audiences see the value of those traits, through an informed criticism based on a knowledge of composition theory and a familiarity with assessment texts. (Twenty references are attached.) (RS)

 * Reproductions supplied by EDRS are the best that can be made *
 * from the original document. *

ED343156

TEXT AGAINST TEXT:
COUNTERBALANCING THE HEGEMONY OF ASSESSMENT

Cornelius Cosgrove
Slippery Rock University of Pennsylvania

"PERMISSION TO REPRODUCE THIS
MATERIAL HAS BEEN GRANTED BY

Cornelius
Cosgrove

TO THE EDUCATIONAL RESOURCES
INFORMATION CENTER (ERIC)."

U S DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION
CENTER (ERIC)

This document has been reproduced as
received from the person or organization
originating it

Minor changes have been made to improve
reproduction quality

* Points of view or opinions stated in this docu-
ment do not necessarily represent official
OERI position or policy

A Paper Presented at the
Conference on College
Composition and Communication

March 21, 1992

Cincinnati, Ohio

CS213241

TEXT AGAINST TEXT:
COUNTERBALANCING THE HEGEMONY OF ASSESSMENT

Cornelius Cosgrove
Slippery Rock University

Allow me to begin by explaining how I arrived at my subject for today, as the origin of my interest may have relevance to my subsequent discussion. A few years ago I looked into definitions of "learning disabilities," and at how writing and the teaching of writing were viewed by those involved in that field. As the director of a writing center, I thought it necessary to venture into the literature; I found the way in which writing was often used to evaluate and categorize "language disabled" students particularly fascinating.

Then, less than two years ago, I read an article by Edward White in which he described his unsettling experiences at an event called the Assessment Forum of the American Association for Higher Education. For nearly two decades, White had devoted considerable professional energy to the area of writing assessment, had published frequently, and had become well-known for his efforts among college composition specialists. Yet he found himself, at this conference, among scholars who made "different assumptions about writing than I," made references that were "entirely unfamiliar," and described assessment procedures different from those he performed and prescribed. He concluded that he had stepped into a "wholly different scholarly community of discourse" (White 187).

It struck me that, despite the theoretical difficulties involved

in defining "discourse communities," White was making an important and necessary distinction. While composition specialists identify themselves primarily as teachers and researchers of writing, assessors identify themselves primarily as testers or evaluators who may or may not focus on writing. In other words, assessors concern themselves with methods of evaluation, while composition specialists concern themselves with the thing being evaluated. There are other observable distinctions. Assessors are usually graduates of schools of education, and are practiced in the techniques of psychological measurement. Composition specialists are usually products of programs in English or Rhetoric, and are practiced in the techniques of textual analysis. Assessors publish in journals devoted to educational measurement and "exceptionalities" in learners. Composition specialists publish in journals devoted to the study of written communication and to the teaching of the same. Both concern themselves with writing evaluation, but must bring quite different perspectives to the task.

Michel Foucault has remarked that disciplines assume, as they develop, "the privileged or exclusive right to speak of a particular subject" (216). It's a short step from that assumption to an exclusionary tendency Foucault characterizes as "the will to truth" (219). Another tendency, in the modern academy, is to concede that "knowledge" is almost always discipline-specific. However, such a concession can be dangerous when two distinct disciplines concern themselves with the same area of study. The perspectives of one discipline may end up suppressing those of the other. One way to combat such suppression is for both disciplines to read, to analyze, and to criticize the texts of the other. As practiced students of

texts, composition specialists could be ideally suited for such textual conflict. Could we counterbalance the potential privileging of the assessment perspective, or of self-appointed interpreters of that perspective, through the study of assessment discourse as text?

To find out, I have been examining a sample of 14 assessment texts. Most of the sample consists of journal articles, with the predominance being experimental papers featuring the common methodology of educational testers--namely, statistical inference. The articles are published in journals primarily devoted to "measurement" and/or "exceptionalities." As a textual critic, the conclusions I have so far drawn from this reading are admittedly impressionistic, and occasionally intuitive.

My first is that valid criticism will not emerge from a close reading of that which is the assessors' strength. Their experiments are often elegantly constructed, and their statistical models frequently struck me as either ingenious or impenetrable. Even though I once took a year-long graduate level course in statistical inference, I doubt my ability to comprehend, much less evaluate, the following sentences from a 1988 article in Educational and Psychological Measurement:

The parallel analysis was performed by generating a 100 X 10 data matrix from a multivariate normal distribution of uncorrelated variables, and then factoring the correlation matrix for the random data set. The eigenvalues from the random matrix were then compared to those of the real data correlation matrix (Perkins et al. 1117).

It is far more fruitful to critically examine the assumptions about writing which inform the tests in question, and the experiments intent on measuring the validity and/or reliability of those tests. Assessors function within a necessarily closed universe. It may be

their job to see, for instance, whether a test measures those "writing abilities" it says it measures. At the same time, it may be our job to ask whether the abilities in question are relevant or of great significance in the execution of a given writing task, and whether there might be other elements to the writing which are of greater relevance or significance.

Let us examine a few instances in which assumptions about writing that govern study designs appear to have escaped critical attention. Here is a sentence taken from the introduction to a study of "data based procedures" meant to identify "written expression disabilities at the university level:"

The study analyzed twelve language components of university compositions, covering the three major aspects [Underlining is my own.] of writing cited previously: syntactic maturity, fluency, and vocabulary (Gajar and Harriman 253).

I suspect the editors of the journal in which this article appeared may be less sensitive to the exclusionary nature of "the" in this sentence than textual critics like ourselves. Nor are they as likely to consider the possibility that syntactic maturity, fluency, and vocabulary have been designated "the ...major aspects of writing" because they are aspects for which psychometricians have devised quantifiable measures.

An article published in Educational and Psychological Measurement argues for use of something called the "Standardized Test of Essential Writing Skills" or STEWS. Part of the argument involves "a series of studies concerning the reliability and validity of the STEWS (DeShields et al. 101)." But another, equally crucial, element is the assertion that brief paragraphs used as writing samples within the test are as valid as "full-length essays" in assessing student writing

(104). This assertion is supported by an appeal to the authority of such late 19th-century rhetoricians as Bain and Genung and to a statement made in a 1909 textbook that a paragraph "may be regarded by itself as a separate and complete composition in miniature (qtd. in DeShields 104)." No authoritative source cited in this section of the article is less than 63 years old; Bain's book was published in 1870. Nevertheless, the authors contend that these works form "the basic theoretical foundation upon which many of our current ideas about composition are based...(104)."

In considering these examples, we have employed two common critical strategies. One is to closely examine the language in which an assumption about writing is couched. Another is to consider the scholarly authority used to support an assumption. A third could be to juxtapose an assumption made by an assessor with a representative text from our own discourse community which may be saying something quite different. A recent comparison of the Woodcock-Johnson Writing Tests to other measures describes the writing sample section as one-sentence responses to 15 prompts over the course of 15 minutes. The author then asserts that "Writing Samples measures higher-order writing skills...by generating meaningful sentences at increasingly more complex levels of vocabulary and abstraction (Mather 36)." Equating higher-order "skills" with the generation of single-sentence ideas may seem reasonable unless the casual reader can compare that definition to one that is richer and more commonsensical. Like this one from a book on training peer tutors in writing: "The four priority concerns we have dubbed HOCs [Higher Order Concerns] are thesis or focus, appropriate voice or tone, organization, and development. Weakness in these areas can devastate a paper..."

(Reigstad and McAndrew 11)." The Woodcock-Johnson's Writing Samples section may assess a student's ability to rapidly generate a sentence from a prompt. but it clearly does not assess the ability to focus a discourse, to develop and organize ideas, or to adopt a consistent and appropriate linguistic voice. Juxtaposing the two texts makes this conclusion obvious.

A fourth critical strategy, analyzing the "works cited" list of a discourse, allows us to judge whether an assessor has remained within or gone beyond the boundaries of her/his own scholarly community. I examined the list of my sample of 14 for sources from the composition literature which would be familiar in terms of the scholar, the journal, the publisher, and/or the subject matter. I found three assessment texts which had no citations familiar to someone who has been reading the composition literature for almost two decades, including a 21-page bibliography from a book entitled Academic Skills Problems: Direct Assessment and Intervention. Of course, this means that more than two-thirds of my sample did contain familiar citations. Three scholars--Charles Cooper, Paul B. Diederich, and Richard Lloyd-Jones--were cited at least three times. Articles from Research in the Teaching of English were cited seven times and books published by NCTE were cited 15 times. While many of the citations from the composition field had been published much before the discourse in which they were cited, one must conclude that a tradition for consulting scholars of the act of writing, as opposed to the act of testing, has been established among assessors. It is also no coincidence, I believe, that assessment discourse which cites composition sources is more likely to acknowledge the problems with construct and content validity which plague attempts to measure

specific components of writing ability.

A final critical approach to evaluating the assessment literature, and one that may serve as a guiding principle for the entire enterprise, is to habitually seek definitions for those qualities or abilities assessors have set out to measure. Unfortunately, assessors have a history of rushing in where angels would fear to tread. Tests are proposed to measure critical thinking when there is no clear definition of critical thinking (Smith 237-8). School districts employ standardized tests to determine if a child is "gifted," even though specialists in "giftedness" have taken pains to distinguish between "proficiency in lesson learning and test taking on the one hand and innovative behavior and creative productive accomplishments on the other (qtd in Engel 9)." In the case of learning disabilities, the failure to match tests with an original definition appears to have led to the creation of definitions which can accommodate the tests psychometricians have been able to develop.

Nowhere is the impulse to define quality or level of accomplishment more obviously destined to failure than in the evaluation of writing. The opinions of readers concerning what makes writing "good" are notoriously disparate, and judgements readers pass about pieces of individual discourse may be even more varied. Undaunted, assessors create instruments that measure "success," "competence," "fluency," and "thematic maturity," while assuming that readers' evaluative difficulties can be overcome through measurement of specific and quantifiable "components" of writing. A critical reading of assessment literature insists on precise definitions of that which is being tested, and clear warrants linking definitions to reported outcomes.

Which is not to say that there aren't plenty of examples of precision and clarity and circumspection in the assessment literature. Moreover, journal articles often conclude with something very distant from a ringing endorsement of the tests under inspection. What follows is a sampling of final sentence(s) from the abstracts of four experimental papers:

The results indicated that the holistic and objective assessments were relatively independent measures of writing ability, indicating support for a two-factor theory of ability assessment. The holistic assessment was found to be more reliable than the objective assessment. (Perkins et al. 1111)

It is suggested that practitioners interested in reliably measuring all aspects of the proposed writing process continuum, as characterized by this cognitive model, use both indirect and direct methods. (Ackerman & Smith 117)

While both types of the tests had advantages, the results of this study underscore some of the advantages of the Diederich method of direct assessment using a writing sample test and support the view that direct assessment should not be replaced by indirect tests for evaluating writing skills... (Sabban & Kay 61)

A recommendation is made that the multiple samples writing approach to the direct measurement of writing ability be subjected to further experimental study and be considered as an alternative method to writing assessment (Mather 84).

What the above suggests is that many assessors are less definite and assertive about the process of evaluating writing than composition specialists like ourselves may have feared. Our true nemeses may be those administrators and practitioners in the many areas of educational psychology who apply the assessors' findings with an unwarranted amount of certainty and resolution. It is with them that we most frequently clash when issues concerning writing evaluation and its uses arise. One elementary language arts teacher has observed that "when it comes down to meetings about students...it is the test

scores and the recommendations of 'support service' personnel which are generally given far greater weight than teacher observation" (Martin 490). Edward White remembers "a group of admission deans and specialists in educational measurement" attributing a significant discrepancy between the essay and multiple choice scores of minority students to the probable laxity of the composition faculty who had rated the essays (188).

The secondary audiences for discourse on writing evaluation--students, parents, and legislators, for example--may be tempted to suppress conflicting perspectives on the issue. Tentative and shifting knowledge has not the appeal of clear, specific numerical results produced by those who have been "trained to test." Writing teachers could be handicapped by a commitment to individual perception and a reputation for idiosyncratic grading. Unless composition specialists can help their audiences see the value of those traits, through an informed criticism based on a knowledge of composition theory and a familiarity with assessment texts.

WORKS CITED

- Ackerman, Terry A., and Philip L. Smith. "A Comparison of the Information Provided by Essay, Multiple-Choice, and Free-Response Writing Tests." Applied Psychological Measurement 12 (1988): 117-128.
- Benton, Stephen L., and Kenneth A. Kiewra. "Measuring the Organizational Aspects of Writing Ability." Journal of Educational Measurement 23 (1986): 377-386.
- Boodoo, Gwyneth M., and Patricia Garlinghouse. "Use of the Essay Examination to Investigate the Writing Skills of Undergraduate Education Majors." Educational and Psychological Measurement 43 (1983): 1005-1014.
- DeShields, Shirley M., Hui K. Hsieh, and Douglas Frost. "The Measurement of Writing Skills: Some Problems and a Probable Solution." Educational and Psychological Measurement 44 (1984): 101-111.
- Engel, Brenda S. "Education: A Universe of Discourse." Teaching and Learning 1 (1986): 4-12.
- Foucault, Michel. The Archaeology of Knowledge and the Discourse on Language. Trans. A.M. Sheridan Smith. New York: Pantheon, 1972.
- Gajar, Anna H., and Nancy Harriman. "Identifying Data Based Procedures for Written Expression Disabilities at the University Level." Education 107 (1987): 252-258.
- Graham, Steve. "Is New Necessarily Better? A Review of the TOWL-2." Learning Disabilities Focus 5 (1989): 47-49.
- Isaacson, Stephen. "Assessing the Writing Product: Qualitative and Quantitative Measures." Exceptional Children 54 (1988): 528-534.
- Martin, Anne. "Screening, Early Intervention, and Remediation: Obscuring Children's Potential." Harvard Educational Review 58 (1988): 488-501.
- Mather, Nancy. "Comparison of the New and Existing Woodcock-Johnson Writing Tests to Other Writing Measures." Learning Disabilities Focus 4 (1989): 84-95.
- Perkins, Kyle, John T. Pohlmann, and Sneila R. Brutton. "A Factor Analysis of Direct and Indirect Measures of English as a Second Language Writing Ability." Educational and Psychological Measurement 48 (1988): 1111-1121.
- Poteet, James A. "Informal Assessment of Written Expression." Learning Disability Quarterly 3 (1980): 88-98.
- Reigstad, Thomas, and Donald McAndrew. Training Tutors for Writing Conferences. Urbana, IL:ERIC/NCTE, 1984.

Sabban, Yitzchak, and Patricia M. Kay. "Distinction between Essay and Objective Tests in Assessing Writing Skills of Underprepared College Students." Journal of Research and Development in Education 21 (1987): 61-68.

Shapiro, Edward S. Academic Skills Problems: Direct Assessment and Intervention. New York: Guilford, 1989.

Smith, Frank. Insult to Intelligence: The Bureaucratic Invasion of Our Classrooms. New York: Arbor House, 1986.

Teale, William H. "Developmentally Appropriate Assessment of Reading and Writing in the Early Childhood Classroom." The Elementary School Journal 89 (1988): 173-183.

Weiner, Eva S. "Diagnostic Evaluation of Writing Skills." Journal of Learning Disabilities 13 (1980): 48-53.

White, Edward M. "Language and Reality in Writing Assessment." College Composition and Communication 41 (1990): 187-200.