SPECIFYING THE WRITING DOMAIN FOR ASSESSMENT: RECOMMENDATIONS TO THE PRACTITIONER

Linda Polin

CSE Report No. 135

December 1980

Paper presented at the Annual Meeting of the National Council on Measurement in Education, Boston, 1989.

Center for the Study of Evaluation
UCLA Graduate School of Education
Los Angeles, California 90024

X3CSE/A

The preparation of this paper was supported by funds from the National Institute of Education (NIE), Department of Health, Education and Welfare. Opinions expressed do not necessarily reflect official NIE position or policy.

PART I: ISSUES IN WRITING ASSESSMENT

Overview

Although writing is cited as one of the three basic skills, it has received much less research attention and much less instructional and assessment emphasis than reading and mathematics, the other two basic subject areas. Our understanding of student writing skills is inadequate and the research and theorizing in the field is somewhat disordered.

In contrast to the reading and mathematics domains, there is little consensus among professionals about what constitutes "good" or even "adequate" writing performance, or even about what knowledge and sub-skills it subsumes. Accordingly, instructional methods and content, and the assessment of writing achievement have been arbitrarily determined or referenced loosely to any one of a variety of competing definitions of writing.

A review of research and theory in writing reveals three predominant conceptual categories: writing structure, writing function, and writing process. Studies in writing structure assume a product orientation. Writing ability is evidenced in the features of the written product and inquiry is directed primarily to linguistic features of syntax. Studies in writing function assume a vehicular orientation in which writing is viewed as a tool adaptable to writing purpose. Competence is evidenced by the writer's ability to adapt to a variety of audiences and rhetorical purposes. Writing process studies exemplify the cognitive psychology orientation in which writing is viewed as the interaction of task requirements and writer strategies. Inquiry is primarily into hypothesized cognitive component processes as they are revealed through

behavioral observations and introspection. Writing competence is viewed as the ability of the writer to cope with the task demands.

Despite the diversity of these perspectives on writing, they are not adversarial or antithetical. In fact, it appears that they might each present a partial explanation of the complex domain of writing. Taken together, their theories and research findings often complement each other. For example, many of the features taken as indicating a good text are explainable outcomes of component processes characterizing good writing behaviors.

Unfortunately, the urgent need among practitioners for a reliable system for assessing student writing ability has led to a narrow view of the product features of writing skills. The emphases in test practice have been upon rating scales and the tasks and procedures to be followed by raters. One of the issues this narrow perspective raises is the instructional (and perhaps construct) validity of procedures. For example, prewriting activities and revision, which are presumed to affect text quality, are heavily endorsed components of writing instruction. Yet current predominant assessment strategies do not directly assess the writer's ability to revise nor the quality of text revision efforts. Nevertheless, much research indicates that students' competence in and understanding of the notion of revision is a salient distinguishing characteristic between skilled and unskilled writers. For instance, English teachers report difficulty in getting unskilled or basic student writers to conceive of revision as more than cosmetic editing of punctuation and word level errors.

This paper aims, then, to encourage a broader, integrative view of the writing skills domain. This view will serve as reference for

developing a sound assessment program which (1) recognizes a variety of teachable component skills and (2) describes a student's writing ability as a profile of his/her competence or achievement in each of those skills.

Issues in Writing Assessment

The UCLA Center for the Study of Evaluation has described several qualities that should be considered in the development and installation of any assessment program (Baker, 1977). Briefly described, these include:

publicness: The public, including students, should have access to and an understanding of the assessment domain.

economy: The test program should function to minimize the time, money, and opportunity costs to students and staff.

instructional sensitivity: The test program should support the instructional program by testing content amenable to and included in instruction.

meaningfulness: The testing experience should be seen as important by the student test takers.

equity: The testing program should provide equal opportunity to succeed or fail to all students, based upon their learning.

technical adequacy: The testing program should provide reliable, valid estimates of student performance.

Applied to the assessment of writing skills, these criteria raise the following concerns:

publicness: What is being tested as "writing" and how is it being tested? What do the assessment scores mean?

economy: What is the most cost-effective way of assessing writing: multiple-choice or writing sample? In collecting writing samples, how are they most efficiently gathered and scored? How many should be gathered? How many rated?

instructional sensitivity: What (diagnostic or prescriptive) instructional information do the writing tests yield? What recommendations for remediation do they provide? How do they tap the success of classroom instruction or district curricula in writing? What nonwriting or prerequisite skills are assessed?

meaningfulness: What is the significance of the testing experience for the student test takers? What does the test convey to them about writing as a skill? What information is fed back to the student?

equity: Does the relationship between the obtained score and the true score in writing ability hold for all test takers? Is the test task the same for all student test takers? What are the major equity issues in writing assessment (e.g., nonstandard English)?

technical adequacy: Does the test task provide a reliable indicator of students' level of achievement in writing? Does the test really measure writing skills? Is the definition of writing complete and appropriate to the interpretation of the test scores? How confidently and to what extent may the outcomes be generalized to writing in other areas or for other functions and topics? Are different writing topics equivalent? Are different error types of equal importance?

Some of the major problems in writing assessment have arisen from our failure to define writing, from our lack of understanding about its development, and from the indicators we use to judge competence. Unresolved questions, therefore, permeate all features of a test of writing ability.

1. task or item: Should the student task be to select a multiple-choice response or to produce a writing sample? What should the students write on? What mode of discourse or rhetorical function, genre and audience?

2. directions: What should the task explicitly ask for and what should it require students to make decisions on?

3. setting: How much time should be provided and under what context should writing be assessed? Should students have the opportunity to revise?

4. evaluative criteria: What features of writing should be evaluated? Should the evaluation process be analytic or impressionistic? Should criteria be differentially weighted?

5. evaluative process: How many writing samples should be corrected per student: How many raters should read each paper? How should raters be checked for stability or consistency?

6. interpretation of scores: What determines competence? How generalizable are scores? Should different subscales be differentially weighted? What are the instructional implications of, e.g., a "2" in "organization"?

Most notable among these issues and the questions they raise is the need for a sound decision-making basis, for a rationale and referent. Domain-referenced testing is a useful vehicle for addressing the matter of assessment criteria. The specification, to which test items are referenced, describes what is meant by writing, the range and depth of the knowledge required of the student, the conditions under which such knowledge is to be demonstrated, and the criteria by which quality of performance is judged. The domain specification, then, publicly communicates the content and nature of the test. By describing the testable material within the subject area, the domain specification is also a useful device for communicating to teachers the instructional content they are being held accountable for teaching. In many ways domain-referenced testing also encourages and reinforces the link between evaluation and instruction. Further benefits of domain-referenced testing stem from the existence of rules governing the generation and scoring of test items. This allows the test developed to control or

exclude factors known to or suspected of affecting test difficulty in ways unrelated to instructional content, i.e., biasing effects. Domain specifications also enchance the reliability of the testing program; the clarity of definition and rules for item generation should allow for results to be replicated using different test items or at different times of administration. Test validity may be judged in comparisons of items to domain, and domain to instructional or curricular contents or objectives.

While it is easy to argue for the domain-referenced testing approach, the actual creation and use of domain-referenced tests of writing present a major challenge. As mentioned briefly above, there is little consensus on the boundaries and contents of the domain of writing skills. However, research and theory from the various fields of inquiry provide some useful information and clues for specifying the domain for assessing writing skills.

PART II: RESEARCH AND THEORY IN WRITING

Building a Writing Domain

A distinction in research perspectives arises immediately in the definition of writing. Some researchers and theorists view writing as a noun, that is, as "text," a written product. Others consider writing a verb, a set of behaviors or overt and covert processes of communication. Obviously, this distinction affects the specification of the writing skills domain.

When writing is seen as a product, test tasks and their instructional prerequisites attend to text features. This focus upon the text sends

the test developer off in search of rating scales for evaluating text and rules for generating rhetorical tasks and appropriate topics. Instruction in this perspective leans toward concern with text features such as organization, syntactic fluency, style, mechanics, and usage (among others). Often this falls out in practice as student exercises in outlining and writing topic sentences, or sentence-combining exercises designed to move students toward greater syntactic sophistication.

When writing is viewed as a subset of a range of communication activities, the text is viewed as an outcome of the interaction between the writer and reader, and the focus is upon performance of necessary writing behaviors presumed to affect communication rather than on the text itself. Instruction emphasizes planning and revising behaviors related to audience, purpose, and content of the writing.

The first perspective, and its implications, describe the majority of current instructional and assessment practices, and their rationales. The second represents the direction taken by research theory efforts in psycholinguistics and social and cognitive psychology. In most instances, work in these areas has been less concerned with evaluation and instruction than with close scrutiny of the writer-in-process to obtain accurate descriptions of behaviors distinguishing skilled from unskilled writers.

This product-process distinction is useful primarily for describing the domain of writing and suggesting assessment possibilities. If we can begin to distinguish the processes and skills that affect product features, and if we can begin to understand their interrelationships and the variables (such as developmental stages) affecting them, then perhaps we can begin to design assessment tasks that will help us describe student competencies in a more meaningful way. Perhaps also the reliability

and validity of our judgments about student achievement in writing will improve. Obviously, such behaviorally based (rather than text-tied) assessments would provide a greater wealth of instructional information for diagnosis and remediation of student writing problems. While descriptions of student competence that rely on an essay's product features tell us what gross errors a student is making, these descriptions do not tell us why the student succumbs to problems nor how to assist him/her in overcoming them. As Bereiter, Scandamalia, and Bracewell point out:

> It is impossible, simply by analyzing the composition that someone has produced in response to an assignment, to infer both (a) what the actual task was, as construed by the writer, and (b) what the writer's competence was in respect to the requirements of that task. If you assume one, you can infer the other, but you cannot infer both at once. Developmental research in writing has generally proceeded by assuming that the task was construed the same by all subjects and therefrom inferring differences in competence. But this is quite an inadequate way of going after any deep understanding of cognitive development in writing. (1979, p. 5)

This paper recommends that we view the writing domain in a manner that is more in accord with our current views of reading. In reading, we now have the research and theory-based sophistication to describe a student's reading ability as a profile which describes his/her competence in oral decoding, literal and inferential comprehension, word attack, vocabulary, and critical or evaluative reading skills. We can construct test tasks we believe measure these skills and provide instructionally relevant definitions of competence and error types. In writing, too, we need to consider the value of adopting a process orientation that recognizes the developmental quality of component skills and subprocesses, the interrelationships among those skills, and the behaviors or outcomes that demonstrate competence.

What is Writing?

Despite the variety of disciplines represented in the literature, most theory and research on the processes in writing appear to distinguish two large process categories. The first category includes those behaviors and thoughts occurring before the drafting of the essay. The second category of activities includes the drafting of the text and any reworkings. These divisions of labor have been described by Stallard (1972, 1976) as "composing" and "transcribing." Composition refers to the invention of the message content; to activities occurring before writing. Transcription refers to the encoding of the message; to the actual production and refinement of the message. These two process categories subsume many tasks and subskills.

During composing, the writer plans. Analyses of writers-in-process and theory-based speculations have been employed to decompose the prewriting experience of composing into related subtasks. However, because these subtask behaviors are so much more covert than writing and revising activities that may be traced in the text, there is less consensus about what does and should constitute the subskills in prewriting. Generally, the differences of opinion related students' detailing skills and knowledge in two levels of prewriting activities. The first is the recall and use of appropriate metaplans; that is, plans for plans, heuristics to guide the task-specific decisions and behaviors. In writing, such plans might begin with the determination of the rhetorical nature of the task. Given that determination, the next step in the plan might be to recall and adapt the appropriate schemata (Bereiter et al., 1979). These schemata, for example, might describe elements and relationships among elements in a business letter of inquiry, an expository essay, a

persuasive editorial letter, or a factual narrative. A next step in prewriting might be to select the schemata describing appropriate audience-relevant writing elements. A final activity in prewriting might be to incorporate or integrate information and decisions into a set of intentions that will guide the further development and transcription of the essay response (Flower & Hayes, 1979; Nold, 1979). Such a plan may affect organization, tone, syntax, etc. in the final text. The second category of prewriting activities, "planning to say" (Flower, n.d.), is concerned with decisions and determinations about the task-specific requirements. This second planning type might be evidenced in jotted notes or outlines indicating student awareness of content plans, of audience, and of purpose. For good writers, at least, these content decisions seem to include decisions about audience information needs and the purpose of the endeavor.

During transcription, the writer drafts a written text. This drafting process has also been described as comprising several interrelated subtasks. Much of the research in writing behaviors has concentrated upon this part of the writing domain. Descriptions of error categories, structural levels of revisions, and repertoires of revision skills are often used to distinguish effective from ineffective writers. Observational and verbal protocol analyses have also been used to uncover the transcription processes that result in effective versus ineffective texts (Hayes & Flower, 1978; Matsuhashi & Cooper, 1978). These processes include recursive planning and revising of the text during writing. Skilled writers have been found to pause during writing for rereading the text they have produced "so far." This rereading seems to serve two purposes. First, writers can monitor and maintain their planned course

of action in the essay. In monitoring their text, writers are hypothesized to compare their intentions and understandings of audience and task function (developed during prewriting) with what they've written. Where dissonance arises between these two representations of meaning (intended and textual), skilled writers are cued to perform revisions (Sommers, 1979). The second function of pausing to reread appears to be to plan the "next step" in a successively refined message. The prewriting plan, then, may be a rather vague guiding intention that is elaborated and specific during the actual writing process.

This brief description of the writing domain presents writing as a complex activity, involving many subskills and processes which draw upon an individual writer's limited resources of attention and effort, and capacities of long- and short-term memory, in addition to his/her knowledge about topic and audience implications. The effect of these task demands for competent writing performance has been termed "writer overload" (Nold, 1979). However, although writer resources and capacities are limited, competent writers appear to be able to cope with the load. Cognitive theory suggests two possibilities. First, competent writers may become more adept at some of the subtasks and thus are able to "pay less attention" to them. This concurs with the descriptions of skilled writers (Hayes & Flower, 1978; Matsuhashi & Cooper, 1978; Stallard, 1972). Or, second, the competent writer may adopt a metaplan or employ strategies for efficient deployment of resources across tasks (Odell, 1978; Rose, n.d.; Young, Becker, & Pike, 1975). Thus, our standard assessment practice of evaluating student writing competence on the basis of the judged quality of essay features is perhaps an assessment of the ultimate criterion, that of competence at putting all the subskills

together for any given writing task. If, on the other hand, assessment proceded from a writing domain that moved forward toward the complex, integrative demonstration of competence from demonstrations of competence in the en-route or prerequisite subskills, we might have an assessment program that could satisfy most of the issues raised in Part I of this paper; that is, an assessment program that provides information of immediate and prescriptive use to district and classroom personnel.

While we must wait for further research and refinement of theory to help us decide upon defensible, valid task analyses and competence markers, some encouraging work in this direction has been done. For example, in revision, we know that there are specific skills that some students either do not know they should do, do not know when or how to do, or are incompetent (inadequate or incomplete) at doing (Bridwell, 1979; Flower, n.d.; Perl, 1979; Sommers, 1979; Stallard, 1972). This suggests a familiar cognitive developmental explanation of the difference between being able to recognize and being able to recognize and correctly respond to a situation. We also know that some subprocesses are bound to others; for example, objective rereading for revision, considering audience when planning content (Beach, 1976; Flower, n.d.). We have evidence (Kroll, 1978) suggesting that at lower grade levels, audience awareness is closely tied to cognitive maturity (ability to take on the role of others, recentering). We also know that different writing topics and purposes require different kinds of thinking, e.g., ability to generalize, to use levels of organization, to use sequences or contrasts (Davis & Nold, 1980; King, 1978). For example, there are developmental differences in the ease of responding to descriptive, narrative, expository, or argumentative writing (Perl, 1977; Perron, 1978).

In addition to research on processes and variables affecting processes, we have some theory-based notions about the kinds of competence markers we might employ for various subskills. Bruce, Collins, Rubin, and Gentner (1978) devote several pages of a technical report to descriptions of "intermediate tasks" that deal with the variety of subskills including discovering ideas, manipulating ideas, producing texts, editing texts, and self-editing texts. Odell (1978) describes linguistic cues in writing that might measure cognitive processes in writing and also suggests a developmental view of these cues. Bereiter (1979) has experimented with young children (grades 4 and 6) in teaching them heuristics for particular rhetorical purposes and has found them able to use those devices in later discourse of their own.

Summary Recommendations

Attention to writing assessment is new. Research and theory are still widely scattered among fields of inquiry and differing perspectives within those fields. As with most new rushes to hot issues, much current and recent work is split in one of two directions: basic research or practical stopgap measures. Applied research will require some integration of available knowledge into a viable framework which prescribes for current practice and recommends future research.

This paper has adapted a domain-referenced view of instruction and assessment. This view stresses definition of a subject, delimitation of instructional goals and efforts, and sampling of criterion behaviors to describe competence vis a vis those instructional goals. The writing skills domain may provide the appropriate framework for integrating basic research and theory on writing with implication and recommendations for instruction and assessment of writing achievement. The writing

domain, however, is not a term I have found anywhere in reviewing the research and theory on writing. In fact, if we choose to describe the writing domain according to current practice, we would have to draw a rather static picture of writing skills in terms of desired features of the product (e.g., mechanics, cohesiveness) and describe criterion performance as the appearance of those qualities in texts addressing different tasks (e.g., exposition, narration). Recent research and theory, however, in contrast to the practitioner experience, have begun to explore the writing domain with an active or rather an interactive model of composing and transcribing skills and subskills. Descriptions of behavioral differences in the performance of these subskills are now used to distinguish competence.

This direction holds much more promise for the practitioner, although she/he will require patience for this greater payoff. If we want to teach students a skill that transcends particular classroom instances of instruction, or task characteristics, it might behoove us to consider the teaching and assessment of composing and transcribing behaviors and their prerequisites. Perhaps after mastery of these skills, students will be prepared to refine the skills or advance beyond these minimal competencies to deal with sophisticated distinctions arising from differences in writing purpose, setting, and topic. If we can begin to distinguish the processes and skills that affect product features and if we can begin to understand their interrelationships and the variables affecting them, perhaps we can begin to ferret out assessment tasks that will help us describe student writing competencies in a more meaningful way. It appears that these "ifs" are possible; these recommendations may then be the long-term research payoff for the practitioner. Meanwhile

because she/he is on the line now, current work on practical assessment problems should continue to be addressed.

References

Baker, E. L. Long Range Plan for the UCLA Center for the Study of Evaluation, Submitted to the National Institute of Education, September 30, 1977.

Barritt, L., & Kroll, B. Some implications of cognitive developmental psychology for research on composing. Chapter 4 in C. Cooper & L. Odell (Eds.), Research on composing: Points of departure. Urbana, IL: National Council of Teachers of English, 1978.

Beach, R. Self evaluation strategies of extensive revisers and non-revisers. College Composition and Communication, 1976, 27, 160-164.

Bereiter, C., Scardamalia, M., & Bracewell, R. An applied cognitive-developmental approach to writing research. Paper presented at the annual meeting of the American Educational Research Association, San Francisco, CA, 1979.

Bruce, B., Collins, A., Rubin, A., & Gentner, D. A cognitive science approach to writing (Tech. Rep. No. 89). Cambridge, MA: Bolt, Beranek and Newman, Inc., 1978.

Davis, B., & Nold, E. The discourse matrix. College composition and communication, in press 1980.

Dilworth, C., Reising, R., & Wolfe, D. Language structure and thought in writing composition: Certain relationships. Research in the Teaching of English, 1978, 12, 97-106.

Flower, L. Good writing: Evaluating the writer's process. Pittsburgh, PA: Carnegie-Mellon University, n.d.

Hayes, J., & Flower, L. Writing as problem solving. Paper presented at the annual meeting of the American Educational Research Association, San Francisco, CA, 1979.

Hayes, J., & Flower, L. Protocol analysis of writing processes. Paper presented at the annual meeting of the American Educational Research Association, Toronto, Canada, 1978.

King, M. Research in composition: A need for theory. Research in the Teaching of English, 1978, 12, 193-202.

Kroll, B. Cognitive egocentrism and the problem of audience awareness in written discourse. Research in the Teaching of English, 1978, 12, 269-281.

Loban, W. Language development: Kindergarten through grade twelve (Res. Rep. No. 18). Urbana, IL: National Council of Teachers of English, 1976.

Matsuhashi, A., & Cooper, C. A video time-monitored observational study: The transcribing behavior and composing processes of a competent high school writer. Unpublished paper, State University of New York, Buffalo, 1978.

Murray, D. Internal revision: A process of discovery. Chapter 7 in C. Cooper & L. Odell (Eds.), Research on composing: Points of departure. Urbana, IL: National Council of Teachers of English, 1978.

Nold, E. The writing process. Unpublished manuscript, Stanford University, 1979.

Odell, L. Measuring the effect of instruction in prewriting. Research in the Teaching of English, 1978, 12, 228-240.

Perl, S. The composing processes of unskilled college writers. Research in the Teaching of English, 1979, 13, 317-336.

Perron, J. Written syntactic complexity and the modes of discourse. Paper presented at the annual meeting of the American Educational Research Association, New York, 1977.

Rose, M. Strategies, audience, exposition and freshmen's progress--A cognitive/contextual theory of instruction for college composition. Los Angeles, CA: University of California, Los Angeles, Department of English, the Writing Project, n.d.

Shaughnessy, M. Errors and expectations: A guide for the teacher of basic writing. New York: Oxford University Press, 1977.

Sommers, N. The need for theory in composition research. College Composition and Communication, 1979, 30, 46-49. (a)

Sommers, N. Revision strategies of student writers and experienced writers. Paper presented at the annual meeting of the National Council of Teachers of English, San Francisco, CA, 1979. (b)

Stallard, C. An analysis of the writing behavior of good student writers. Unpublished doctoral dissertation, University of Virginia, 1972.

Stallard, C. Composing: A cognitive process theory. College Composition and Communication, 1976, 27, 181-184.

Stallard, C. An analysis of the writing behavior of good student writers. Research in the Teaching of English, 1978, 12, 206-218.

Young, R., Becker, A., & Pike, K. Rhetoric: Discovery and change. New York: Harcourt, Brace and World, 1970.