

Richard L. Larson

SELECTED BIBLIOGRAPHY OF WRITINGS ON THE EVALUATION OF STUDENTS' ACHIEVEMENTS IN COMPOSITION

I present here a selection from the large, and growing, body of literature on responding to and evaluating students' composing skills in courses and tests. The selection, obviously, is my personal choice—the books and essays that seem to me most conspicuously to contribute to our knowledge of how to judge student writing.*

I group my selections according to the purposes that seem to underlie the authors' work. First, and perhaps most helpful to the classroom teacher, come essays designed to advise teachers on how to react to students' writing as part of the instruction in a writing course. Though the teacher can probably never quite escape the role of authority figure and judge in dealing with students' papers, the essays in this group suggest how the teacher can act more as guide or coach than as judge—can suggest to the student what he or she has accomplished, or not accomplished, and what steps he might take to add to the accomplishments, either in revising the paper being looked at or in attempting the next one. The perspectives from which the guidance is given differ (my own piece, for instance, invites the teacher to look at the student's paper, as a whole, as a completed act of using language, and to suggest ways in which that act of using language for a particular purpose might be strengthened); different users of the bibliography will find different approaches to student papers congenial to them, and that varied response will be entirely appropriate, since none of the suggested perspectives or approaches has been shown by research studies to be the best, or a better, way to respond to students' work.

The second group of essays, quite small, deals openly with the task of making judgments about students' writing. Here the judgments discussed

Richard L. Larson, Dean of Education at Lehman College, publishes a selected bibliography of research and writing about the teaching of composition in the May issues of College Composition and Communication.

are essentially administrative acts—decisions made about where to place student papers, and students, on scales that permit assigning the student to a particular class, declaring that he or she has or does not have a particular aptitude for writing or for academic study, determining that he or she is or is not demonstrating required proficiency in written English, or establishing the progress that he or she is making relative to a starting point (particularly if that progress will on some scale or other be deemed adequate, for example, to excuse the student from further work in writing). The emphasis, that is, does not fall in these essays on the teaching that can be done for the student, but only upon giving the student a bit of feedback in the form of a summary score, while enabling school administrators to reach a decision. In this group are two current discussions about testing procedures and procedures for scoring tests. Of these two, the more directly applicable to the work of the classroom teacher is the book by Diederich; the other item discusses and evaluates particular tests that teachers might employ.

The third group of essays listed deals, we might say, with the *processes* of measurement; these pieces investigate the theoretical and practical uses of various means for measuring and various scales or kinds of data that can be derived from testing. This group also contains discussions of national efforts to determine the skills in writing exhibited by students of different ages. The studies in this group do not necessarily advocate, for the classroom teacher or even for the administrator, the general use of the procedures discussed. Instead, as noted, they are contributions to the literature on the theory of testing, or to research on what tests can tell us and on how they can be scored. Much of the information contained here is now used to support, or one day may be used to support, the use of particular techniques to gather information, and thus in the future may affect the lives of teachers who do not now hear these techniques discussed in their schools or their neighborhoods. Knowing about the discussion of these testing procedures, therefore, may be of benefit as teachers talk about the evaluation of students. At the very least, familiarity with these pieces will let teachers see the issues now being debated among those concerned with evaluation, and thus let them sense the state of the art of evaluation.

For the teacher in basic writing, even more than for the teacher in regular writing courses and advanced courses, evaluation is a special concern because of the persistent need to determine when a student is ready to move onward in the sequence of writing courses or has satisfied some minimum standards of accomplishment. The items in this

bibliography will not offer easy answers for teachers who are responsible for these determinations, but they may equip these teachers to make better decisions about how to work with students, and may fortify them against capricious efforts to adopt judgmental techniques that have not themselves been fully investigated and evaluated.

I. SUGGESTIONS FOR CLASSROOM TEACHERS ON RESPONDING TO STUDENTS' WRITING

Bain, Robert, "Reading Student Papers," *College Composition and Communication*, 25 (October, 1974), 307-309.

Offers suggestions about the judging of student papers based on the premise that the student's work should be evaluated mainly on how well he executes his purpose in dealing with his chosen subject.

Beaven, Mary H., "Individualized Goal-Setting, Self-Evaluation, and Peer Evaluation," in Charles Cooper and Lee Odell, eds., *Evaluating Writing* (Urbana, Illinois: National Council of Teachers of English, 1977), pp. 135-156.

Focuses on formative evaluation of writing, and after outlining six major assumptions underlying the author's approach (including her attention to the links between development of a student as a writer and his or her overall development as a person), describes procedures for teachers to use in making comments on students' writing: to work with students on setting (and achieving) their own goals for improvement; to have students evaluate their own work; and to have students' work evaluated by peers working in groups. Offers a rationale for each procedure, and suggests questions that a teacher can help students to learn in using it. Draws extensively on published research, particularly from educational psychology and psychology of personal development and group processes.

Cooper, Charles R., "Responding to Student Writing," in Walter Petty and Patrick Finn, ed., *The Writing Processes of Students* (Report of the Annual Conference on Language Arts, State University of New York at Buffalo, 1975), pp. 31-39.

Urges an emphasis in the composition curriculum on regular and frequent writing—in which students go through the complete process of composing—and recommends supportive, helpful responses to students' work (including responses by other students). Particularly advises discussion of the student's rhetorical emphasis

(including audience, voice, and design) in the paper, the “intellectual strategies” employed, and characteristics of the syntax.

Larson, Richard L., “The Whole is More Than the Sum of Its Parts: Notes on Responding to Students’ Papers,” *Arizona English Bulletin*, 16 (February 1974), 175-181.

Suggests that in evaluating students’ writing, teachers look at the success of the piece as a complete work, offers some questions to aid in making this assessment, and shows how the questions apply to some student writing.

Lamberg, Walter J., “Feedback on Writing: Much More Than Teacher Corrections,” *Statement: The Journal of the Colorado Language Arts Society*, 12 (May, 1977), 33-38.

Defining “feedback” as “information about performance,” reviews various theories about responding to student writing, cites the results of some research on the subject, and asks whether feedback should be different for different purposes and in different conditions.

Odell, Lee, “Responding to Student Writing,” *College Composition and Communication*, 24 (December, 1973), 394-400.

Drawing on analytical techniques developed by Kenneth Pike and others, identifies some habits of thinking and responding to experiences that appear in students’ writing, and suggests how teachers can help students develop new strategies and procedures for thinking about their subjects.

Kantor, Ken, “Evaluating Creative Writing: A Different Ball Game,” *English Journal*, 64 (April, 1975), 72-74.

Proposes six criteria, drawn from psychologists’ discussions of creativity, for the evaluation of students’ creative writing, and applies the criteria to a brief story written by a student.

Shaughnessy, Mina P., *Errors and Expectations* (New York: Oxford University Press, 1977).

Working from an analysis of papers by thousands of students, classifies and interprets the “errors” found in these students’ work, suggesting the sources of these errors—the reasons (or reasoning) that lead students to make them—and proposing ways of responding to them. Not primarily a book about testing or evaluation, but nonetheless a book that can help make the evaluation of papers containing errors become wiser and more humane.

II. SUGGESTIONS ABOUT TESTS AND MEASURES OF ABILITY IN COMPOSITION AND GROWTH IN COMPOSITION SKILLS

Braddock, Richard, "Evaluation of Writing Tests," in Alfred H. Grommon, ed., *Reviews of Selected Published Tests in English* (Urbana, Illinois: National Council of Teachers of English, 1976), pp. 118-126.

Cites the shortcomings of standardized tests of writing, taken as a group, and then comments on four specific tests, at least two of which are useful for their stated purposes.

Diederich, Paul B., *Measuring Growth in English* (Urbana, Illinois: National Council of Teachers of English, 1974).

Proposes procedures for the reliable evaluation of students' writing, and offers advice on how to assure the reliability of evaluation through the use of appropriate statistical procedures. Includes sample examination papers and a statement of criteria by which students' writing can be judged. Important book for those who are involved in testing and want to learn about interpretation of the statistics often given in manuals that explain the scoring of tests.

III. RESEARCH STUDIES, ESSAYS ON THE THEORY OF TESTING, AND DISCUSSIONS OF ISSUES IN TESTING AND MEASUREMENT IN COMPOSITION

Breland, Hunter M., "Multiple-Choice Test Assesses Writing Ability," *Findings (ETS)*, V, 1 (1977), 1-4.

Using data drawn from studies at Educational Testing Service, argues that the Test of Standard Written English is as useful as a written essay, and much easier to administer and score, for purposes of placing students in writing courses and thus for the planning of instructional activities. Suggests also that many students completing composition courses are not, in the judgment of scorers who worked on the study, producing satisfactory essays.

Britton, James, *et al.*, *The Development of Writing Abilities (11-18)* (London: Macmillan Education, 1975).

Reports on an extended research project conducted under the sponsorship of the Schools Council, concerning the kinds of writing done in school (in all subjects) by students aged 11-18. Constructs a fresh procedure for classifying writing according to the relationship

of writer to reader(s) and according to the function served by the writing. An important book for those interested in writing in the schools and in procedures for doing research on writing.

Cohen, Arthur M., "Assessing College Students' Ability to Write Compositions," *Research in the Teaching of English*, 7 (Winter, 1973), 356-371.

Reports on how a group of two-year college teachers participated in an experiment to determine whether their students' ability to write well was improving in response to instruction, and on the instrument devised by the group to measure that improvement. Useful discussion of procedures for engaging classroom teachers in the conduct of research.

Cooper, Charles R., "Holistic Evaluation of Writing," in Charles Cooper and Lee Odell, eds., *Evaluating Writing* (Urbana, Illinois: National Council of Teachers of English, 1977), pp. 3-31.

Argues for the superiority of using an essay test rather than a multiple-choice test in the evaluation of writing, and then reviews a number of holistic procedures (to be differentiated from the making of counts of particular elements of syntax, diction, mechanics, and so on), showing how each procedure should be managed. Discusses the use of scales made up of graded complete essays, feature analysis (judging one feature of the writing only), primary trait scoring, general impression marking, and "center of gravity response" (the term is from Peter Elbow's *Writing Without Teachers*, London: Oxford University Press, 1973), before discussing the development and use of analytic scales (where the desired qualities of a piece of writing are enumerated, and the characteristics of high, average, and low papers are specified). Explains the procedures essential to effective use of analytic scales, and includes illustrations of such scales.

Fagan, William T., Charles R. Cooper, and Julie Jensen, "Measures: Writing," in *Measures for Research and Evaluation in the English Language Arts* (Urbana, Illinois: National Council of Teachers of English, 1975), pp. 185-206.

Lists, describes, and gives data on validity and reliability about fourteen measures useful in research on writing, many of them not previously published. Includes tests, evaluation scales, indices of such features as syntactic maturity, analytical tools, and a corpus of

American expository essays. Likely to be of value to persons engaged in serious research about writing and the teaching of composition.

Godshalk, Fred I., Frances Swineford and William E. Coffman, *The Measurement of Writing Ability* (New York: College Entrance Examination Board, 1966).

Reports on an extended study to determine which of three kinds of tests—a series of objective tests, a series of objective tests together with an interlinear exercise (in which the student is given a passage of prose with errors, and is asked to indicate corrections between the lines of the passage), and a series of objective tests together with a short essay—gives the most reliable predictor of students' writing ability. Concludes that the objective tests with essay are the most valid predictors of writing ability, but supports the assertion that scores on objective tests alone, if the tests are well chosen, are themselves valid predictors. Discusses the establishment of a criterion against which to judge the validity of the various kinds of tests. A complex study, buttressed with extensive statistical data and tables.

Hunt, Kellogg W., "Early Blooming and Late Blooming Syntactic Structures," in Charles Cooper and Lee Odell, eds., *Evaluating Writing* (Urbana, Illinois: National Council of Teachers of English, 1977), pp. 91-104.

Reports on some recent research which confirms that older children and adults in general have more words per T-unit and more embedded S-constituents (analogous to what were formerly called kernel sentences) per sentence, and enumerates kinds of syntactic structures that appear more frequently in older writers than in younger writers. Comments on the implications of these findings for research and for the making of curricula.

Illo, John, "From Senior to Freshman: A Study of Performance in English Composition in High School and College," *Research in the Teaching of English*, 10 (Fall, 1976), pp. 137-146.

Reports on a study of the correlations between various factors and indices (including the results of a questionnaire concerning home activities, high-school instruction, and family) concerning freshmen's ability at an experience with writing, and measures of performance in writing in college courses. Suggests that the emotion, purpose, and will of the student is more directly responsi-

ble for the success of the student in college courses than other factors, such as intelligence and quality of preparation in writing. Explores the implications of this argument.

Lloyd-Jones, Richard, "Primary Trait Scoring," in Charles Cooper and Lee Odell, eds., *Evaluating Writing* (Urbana, Illinois: National Council of Teachers of English, 1977), pp. 33-66.

Differentiates "holistic scoring" from "atomistic scoring," and then, citing the theories of discourse underlying the procedure, defines "primary trait scoring" as, essentially, the judging of how well the writer responded to the audience, purpose, and occasion stipulated in doing a writing exercise or test. Describes procedures for developing exercises with which primary trait scoring can be used, and describes the preparation of scoring guides for these exercises. Includes illustrative exercises and the (occasionally complex) scoring guides that accompany them. Suggests applications for primary trait scoring in research and teaching.

Mellon, John, "The Writing Assessment," in *National Assessment and the Teaching of English* (Urbana, Illinois: National Council of Teachers of English, 1975), pp. 14-38.

Describes the first round of tests of writing conducted by the National Assessment, reports some of the findings, and evaluates both the testing procedures and the findings. Notes particularly the suggestions in the assessment that students in high school may not be improving greatly their ability at composing. Notes, too, the imprecision of phrasing in the writing assignments used in the first round and offers suggestions about the effective focusing of assignments. Finally, notes the absence of attention, in the first round of testing, to expressive writing and writing that deals with feelings. Useful introduction to the accomplishments of the National Assessment in writing and to its problems.

Mellon, John C., "Round Two of The National Writing Assessment—Interpreting the Apparent Decline in Writing Ability: A Review," *Research in the Teaching of English*, 10 (Spring, 1976), 66-74.

Examines a report of the National Assessment on "Writing Mechanics: 1969-1974." Notes the deficiencies in the procedures used by the National Assessment, and urges caution in interpreting the data presented, but accepts the general conclusion that there has been a decline in the writing ability of students (in the age groups

considered) in the five-year period. Notes that some of the decline is perhaps due to lessening of attention to writing in the teaching of English in schools. Poses policy questions regarding the role that writing should play in the curriculum.

Moslemi, Marlene H., "The Grading of Creative Writing Essays," *Research in the Teaching of English*, 9 (Fall, 1975), 154-161.

Although reporting an experiment with ninth-graders in which pre- and post-tests surrounding an intensive unit (using audio-visual materials) on creative writing showed no significant improvement in students' ability at creative writing, the article does describe procedures for selecting and training judges of creative writing that resulted in high reliability of scores among the judges.

Odell, Lee, "Measuring Changes in Intellectual Processes as One Dimension of Growth in Writing," in Charles Cooper and Lee Odell, eds., *Evaluating Writing* (Urbana, Illinois: National Council of Teachers of English: 1977), pp. 107-132.

Recapitulates the analysis by Pike and others of the intellectual processes in which people engage, citing the acts of focusing, contrasting, classifying, noting change, relating change, relating events to physical context. Then suggests how students' writing can be analyzed to discover the kinds of processes at work in each piece and the frequency of their occurrence, suggesting that such analysis helps the diagnosis of students' writing problems. Suggests that measures of change in the use of these processes may be important to a comprehensive evaluation of growth in writing.

Purves, Alan, and the Task Force on Measurement and Evaluation in the Study of English. *Common Sense and Testing English* (Urbana, Illinois: National Council of Teachers of English, 1975).

For persons not much acquainted with testing practices and procedures, describes main concepts in testing, kinds of tests, kinds of scores, and uses to which test results can be put, along with ways of defining the uses and limitations of different tests so that persons interested in employing the tests can understand what they are up to, and proceed more wisely than they might otherwise do.

Sanders, Sara E. and John H. Littlefield, "Perhaps Test Essays Can Reflect Significant Improvement in Freshman Composition," *Research in the Teaching of English*, 9 (Fall, 1975), 145-153.

Points out weaknesses in the procedures typically used to test improvements in composition—through impromptu pre- and post-test themes—and reports a study in which students *did* demonstrate improvement when allowed, on both pre-and post-test, to engage in research, undertake pre-writing activities, and revise the first drafts of their essays. Differentiates between the standard “expository” approach to the teaching of writing and the “aims” approach (following Kinneavy’s *A Theory of Discourse*), but in this experiment reports significant improvement between pre- and post-test for students working with both approaches.