

Classroom Assessment of Writing: Purpose, Issues, and Strategies*

John Albertini

National Technical Institute for the Deaf
Rochester, New York

Freedom is feeling easy in your harness.

-- Robert Frost

Purpose of Classroom Assessment

As teachers of writing we are expected to assess and evaluate students' writing and to help colleagues in other departments do the same. Whatever our approach to the teaching of writing, we must assign grades and prepare students for programmatic assessment. As Peter Elbow (1993) has noted, "Much of what we do in the classroom is determined by the assessment structures we work under" (p.187). In the field of second language writing assessment, the return to the direct assessment of writing, as seen as an attempt to make the assessment structures we work under more valid. In some school districts and colleges (see for example, Brand, 1992 and Weiser, 1992), portfolio assessment is being used as an alternative to testing. Here, I argue that the best assessment of writing in the classroom is an on-going, descriptive documentation of behavior and attitude, and that conscientious assessment at this level will allow us to move more freely in the harnesses of program-level evaluation.

The assessment strategies considered here are both valid and manageable at the classroom level; that is, they reflect real writing behaviors and strategies that teachers already use for instructional purposes. To assess writing, teacher becomes researcher and observes, records, collects, categorizes and evaluates data. The data includes documentation of changes in attitude and knowledge of oneself as a writer as well as the acquisition of skill. Because we want the student to become involved in the assessment process, the strategies go beyond observation by the teacher and include collaboration with and reflection by the student. Finally, it is important to note that assessment activities may be conducted in more than one language, and the data itself may be recorded in written or videotaped form.

Assessment and Evaluation

When I ask teachers what they remember about evaluation of their own writing, they recall corrections in red ink, marginal comments, and no comments at all. They recall feelings of surprise, when, daring to take a risk, they are penalized for doing so. They also recall critical but encouraging comments from some instructors and not wanting others to lay a finger on their writing. Along with a shift in emphasis from product to process in composition instruction (Hairston, 1982), has come the point of view that not all writing should be evaluated; and, following Weaver (1990, p. 182), it may be useful to

*This presentation was made in June, 1994 in Atlanta, Georgia at *Tools for Language: Deaf Students at the Postsecondary Level*, a PEC-sponsored mini-conference.

distinguish between everyday documentation and periodic judgment. Here, I take assessment to mean an analysis and interpretation of that data for the purposes of grading or placement. Traditionally, we have graded and corrected students' writing. Recent experience and research suggest that we should also consider rating their writing holistically and responding without evaluation.

Grading versus Rating

Grading is the practice of assigning points or letter grades to an essay according to certain criteria. We assign a grade to indicate how much of a certain criterion is present in the writing. Assuming we have the same criterion for all our students, a grade of "B" should mean the same thing from essay to essay. Grading involves a scale composed of intervals; and the intervals or units on this scale may be added and subtracted.

Rating (or "ranking") involves a different kind of scale. Here, judges rate the quality or relative "goodness" of writing samples along an ordinal scale (Hatch & Farhady, 1982). Examples of such scales are holistic rating scales with values ranging from 1 to 6 points, 0- to 100-points, or "poor" to "excellent." Holistic ratings of writing samples are now preferred over indirect, multiple-choice assessments of writing skill for students who are deaf or hard of hearing as well as hearing students of English as a second language (Albertini et al., 1986; Berent et al., 1994; Jacobs et al., 1981).

At NTID, holistic ratings are used to place new students in developmental writing courses (Bochner et al., 1992) and to admit students into degree-prerequisite composition courses at RIT. For the former, students are given thirty minutes to write a short essay about first impressions of NTID. Each essay is rated on a scale of 0-100 points by three experienced raters (English instructors), and the three ratings are averaged to yield a single score. The scoring procedure directs raters' attention to four categories--content, organization, language and vocabulary--and is thus a "modified holistic" rating procedure. Training and the practice of averaging raters' scores increase the reliability of the score. For instructors interested in improving the reliability of classroom ratings, such a procedure is feasible. It only requires collaboration and some consensus among instructors.

When a colleague in civil technology asks for advice about grading deaf students' lab reports and is overwhelmed by the report's grammatical anomalies and departures from the expected format, we suggest that the colleague not assign a single letter grade to the report. Rather, we advise grading content, organization, language, and mechanics separately. At NTID, a concern for the improvement of lab reports has led to cross disciplinary collaboration among English and technical faculty (Shannon, Keifer, & Senior, personal communication).

Correcting versus Responding

The efficacy of correcting students' writing continues to be debated. One report from the University of Minnesota (Semke, 1984) indicates that accuracy, fluency, and general language proficiency

in the writing of students of German was enhanced by practice, not error correction. In a study of home and school influences on low-income children's literacy, researchers report that

instructional techniques that rely heavily on teacher corrections, that stress producing mechanically perfect texts, and that fail to provide an appreciative audience for even the poorest writers' efforts, may be especially frustrating to children who confront the writing task with little confidence in their ability to say something of interest to others (Snow et al., 1991).

Anecdotal reports from adult deaf writers indicate that the sight of school papers "bleeding with red pencil" adversely affected their motivation to write (Gustason, 1992, p.64). On the other hand, teachers of writing to hearing second language learners (for example, Reid, 1994) argue that, as editors, mentors, and surrogate audiences for academic writing, they cannot abrogate the responsibility to correct unsubstantiated conclusions or departures from standard form and acceptable usage.

With adult students, negotiating an appropriate time and context for correction is one solution. Also, selective correction of only those errors related to the main objective of an assignment will reduce student frustration and increase learning. In the civil technology lab report, the steps of a particular test must be reported accurately. Clear description of grammatical errors related to clarity and intelligibility was suggested several years ago by Burt and Kiparsky (1974). They categorized sentence level grammatical errors as either "global" or "local" mistakes; and claimed that missing or inappropriate clausal connectors and tense inflections were "global" in that they affected overall intelligibility more than missing noun inflections and articles ("local" mistakes). *Figure 1* includes a list of the major error types according to their definitions.

Figure 1

- I. Global Mistakes - Those that confuse the relationship among clauses, such as:
 - A. Use of connectors
correction: change conjunctions, relative pronouns
 - B. Distinction between coordinate and relative clause constructions, or the order of constituents
correction: put relative clause immediately after its antecedent head noun
 - C. Parallel structure in reduced co-ordinate
correction: add missing subject
 - D. Tense continuity across clauses
correction: change endings
- II. Local Mistakes
 - A. Articles
 - B. Inflections
 - C. Auxiliaries
 - D. Prepositions
 - E. Vocabulary

Burt & Kiparsky, 1974

What to correct is one question; another is, when. For several years now, advocates of the writing-as-process approach have suggested that we reserve grammatical correction to final, "pre-publication" stages of writing. A new writer, they argue, should be allowed to focus first on content and arrangement, then on style and mechanics. An instructor and other readers can promote continued writing and revision

by reflecting what is seen, heard or felt in the piece. Such feedback may be simply the reiteration of striking works, phrases or ideas and is decidedly non-evaluative. Comments such as "I like ...," "I don't like...," and "You should...," are withheld. Responding without correction in teacher-student conferences and in writer groups may be particularly effective with writers whose concern for correctness interferes with concept formation and fluency. Peter Elbow (1993) reports that the creation of "evaluation-free zones" at the beginning of each semester improves both students' writing and his own attitude towards it. He reports "liking" students' writing better and, as a consequence, being more able to criticize it constructively.

Non-evaluative response establishes a connection between writer and reader. A student writes to satisfy a requirement; a writer writes to connect with a reader. Responding to a student's text with experiences of our own shows a personal connection to that text. Such connections should motivate a student to continue writing. I once asked a colleague to comment on a personal piece of writing concerning a student's violence at home. My colleague's response began, "I remember when I was so angry that I ..."

Strategies

Classroom assessment of writing begins with the instructor but involves the student as soon as possible. The instructor may use logs, checklists or grids. Logs and checklists document work completed but also milestones and problems along the way. A grid, which is a list of criteria plus a simple rating, is a useful way of summarizing an assessment of one piece or a collection of pieces in student-teacher conferences. Peter Elbow (1993) uses an analytic grid (shown in *Figure 2*) to comment on student papers and to provide evaluation.

Figure 2

Analytic Grid

Strong	OK	Weak	
			Content, Insights, Thinking, Grappling with Topic
			Genuine Revision, Substantive Changes, Not Just Editing
			Organization, Structure, Guiding the Reader
			Language: Syntax, Sentences, Wording, Voice
			Mechanics: Spelling, Grammar, Punctuation, Proofreading
			Overall (Note: this is not a sum of the other scores)

Elbow, 1993

Students are drawn into the assessment of their own writing through conferences and interviews. Donald Murray (1985) suggests that instructors begin individual conferences by asking students to write about the pieces under consideration. The student should describe what was attempted, what was achieved, and what the next step will be. A dialogue journal is an interactive context where writing may be discussed in writing. Three ground rules for dialogue journal writing are that 1) teacher and student are partners, 2) that the content of the journal is negotiated, and 3) that the writing is never corrected. In this context, the teacher responds on a personal level to what the student has written. In writing classes a dialogue journal may become a writer's notebook where experiences are traded, past writing experiences are recalled, and

view of writing are discussed. This interactive writing may be used to seed other, more formal pieces of writing outside of the journal. For some students, the journal context helps trigger recall of experience and reflection on the writing process. If students are willing to comment regularly on their strengths and weaknesses as writers, a longitudinal self-assessment is compiled by the end of the course.

Self-assessment and reflection are our ultimate goals. Questionnaires, student logs and journals will prompt students to consider their strengths and weaknesses as writers. The writing portfolio is another context where reflection is appropriate. According to Yancey (1992), the inclusion of written reflection is what distinguishes writing portfolios from art or investment portfolios. Like these others, writing portfolios are longitudinal in nature, diverse in content, and collaborative in ownership and composition (assuming that the student has received feedback on various drafts from instructor and classmates). Unlike artists or investors, however, the writer is asked to reflect on content and process and to provide some sort of introduction to the pieces. Such commentary may take the form of a "letter to the reader" or short process descriptions preceding each piece. Inviting students to narrate the contents of their portfolios should elicit evaluative comments like, "One of my strengths in writing is ..." and "My writing style has changed so much within the last year or so!"

In practice, writing portfolios combine several forms of documentation and evaluation. A writing portfolio can include product, process and reflection. Inclusion of product and process allows others to evaluate the acquisition of skill or strategy; reflection reveals the writer's attitude and point of view. As an assessment tool, the writing portfolio compares favorably with standardized indirect measures of writing with regard to validity. Given the complexity and variety of real writing tasks, a collection of final drafts written on different topics at different times is more valid than a single sample of writing as well. To the extent that the creation of a portfolio mirrors the writing process followed in other college courses, it is a valid and relevant assessment of a student's academic writing ability (Hamp-Lyons, 1991, p. 263).

On the other hand, the individuality and variety inherent in this method make it difficult to estimate the reliability of portfolio assessment. Comparability and replicability of the evaluations of the portfolios is the issue. Ratings across portfolios become more stable to the extent that we can elicit the ratings of colleagues who read the final drafts with the same criteria as we do. One additional rating greatly improves the reliability of a final evaluation; and an additional set of comments provides the student with objective feedback in the sense it is a response from someone who has not been involved in the process of writing the pieces.

Teachers who use portfolios generally have a preference for process writing approaches and their use is more common in social science and humanities courses than in natural science courses (Johns, 1991). Instructors who use portfolios as an assessment tool report that they have had a salutary effect on their teaching (Yancey, 1992). In writing courses, they proved both a focus of study and a record of growth. For students (and teachers) unfamiliar with the method, it is advisable to conduct periodic portfolio checks and even to assign preliminary grades based on quantity and quality. A working portfolio implies but does not ensure genuine revision. Portfolios may include a variety of languages and media. Graphics may add

to appearance and interest, and depending on the readership, bilingual pieces and reflection on the writing process may add depth. If we ask colleagues to rate and comment on our students' portfolios, we need to provide explicit rating guidelines and a reasonable number of portfolios to read. On a rating scale of 1 to 5, for example, what does a "5" mean? Suitable for publication in a campus literary magazine?

Conclusion

The techniques described here are simply ways of recording observations and gathering samples. Their use serves a dual purpose: to document learning and to foster writer maturity. Use of these techniques will make students ware of the process and problems, solutions and changes in their own writing. It will also help us loosen the reins without losing sight of the goals. We can encourage risk-taking and also teach editing. Most importantly, when the time comes for evaluation, we can provide students and supervisors with multiple assessments of performance and documentation of change.

Descriptive methods may be used by the instructor, by instructor and student together, and by the student alone. In using descriptive modes of assessment versus standardized tests, we trade uniformity, balance and sometimes breadth for variety, individuality, and depth. Descriptive assessments can augment the reliable but shallow information we get from standardized test scores. Used alone, standardized tests become blinders, fixing our gaze on a narrow path, a limited characterization of a student's abilities. Some have suggested that the use of portfolios may provide the desired link between classroom assessment and large-scale testing (Freedman, 1991) or that placement essays be used in conjunction with portfolio assessment (Brand, 1992). Thus, if we use longitudinal assessment of student writing we may be able to work more easily within institutional structures or ultimately, we may be able to change them.

References

- Albertini, J., Bochner, J., Cuneo, C., Hunt, L., Nielsen, R., Seago, L., & Shannon, N. (1986). Development of writing tests for deaf college students. Teaching English to Deaf and Second-Language Students, 4(2), 5-11.
- Berent, G., Samar, V., Kelly, R., Berent, R., Bochner, J., Albertini, J., & Sacken, J. (1994, March). The valid assessment of deaf college students' writing competency. Paper presented at the 28th Annual Convention, Teachers of English to Speakers of other Languages, Baltimore, MD.
- Bochner, J., Albertini, J., Samar, V., & Metz, D. (1992). External and diagnostic validity of the NTID Writing Test: An investigation using direct magnitude estimation and principal components analysis. Research in the Teaching of English, 26, 299-314.
- Brand, A. (1992). Portfolio and Test essay: The Best of Both Writing Assessment Worlds at SUNY Brockport. (Report No. EDO-CS-92-10). Bloomington, IN: Indiana University, ERIC Clearinghouse on Reading and Communication Skills.

Burt, M. K. & Kiparsky, C. (1974). Global and local mistakes. In Shumann, J. H. & Stenson, N. (Eds.), New Frontiers in Second Language Learning. Rowley, MA: Newbury House. 71-80.

Calfee, R. & Perfumo, P. (1993). Student Portfolios and Teacher Logs: Blueprint for a Revolution in Assessment. (Technical Report No. 65). Berkeley, CA: University of California, National Center for the Study of Writing.

Elbow, P. (1993). Ranking, evaluating, and liking: Sorting out three forms of judgment. College English, 55 (2), 187-205.

Freedman, S. (1991). Evaluating Writing: Linking Large-scale Testing and Classroom Assessment. (Technical Report No. 27). Berkeley, CA: University of California, National Center for the Study of Writing.

Goodman, K., Goodman, Y., & Hood, W. (Eds.). (1989). The Whole Language Evaluation Book. Portsmouth, NH: Heinemann.

Gustason, G. (1992). Wall , after wall, after wall. In Garretson, M. (Ed.), Viewpoints on Deafness: A Deaf American Monograph, 42, 63-66.

Hamp-Lyons, L. (1990). Second language writing: Assessment issues. In Kroll, B. (Ed.), Second Language Writing: Research Insights for the Classroom. New York: Cambridge University Press.

Hamp-Lyons, L. (Ed.). (1991). Assessing Second Language Writing in Academic Contexts. Norwood, NJ: Ablex.

Hairston, M. (1982). The winds of change: Thomas Kuhn and the revolution in the teaching of writing. College Composition and Communication, 33, 76-88.

Hatch, E. & Farhady, H. (1982). Research Design and Statistics for Applied Linguistics. Rowley, MA: Newbury House.

Jacobs, H., Zinkgraf, S., Wormuth, D., Hartfiel, V., & Hughey, J. (1981). Testing ESL Composition: A Practical Approach. Rowley, MA: Newbury House.

Johns, A. (1991). Faculty assessment of ESL student literacy skills: Implications for writing assessment. In Hamp-Lyons, L. (Ed.), Assessing Second Language Writing in Academic Contexts. Norwood, NJ: Ablex.

Murray, D. (1985). A Writer Teaches Writing, Second Edition. Boston: Houghton Mifflin.

Reid, J. (1994). Responding to ESL students' texts: The myth of appropriation. TESOL Quarterly, 28, 273-292.

Semke, H. D. (1984). Effects of the red pen. Foreign Language Annuals, 17(3), 195-202.

Snow, C., Barnes, W., Chandler, J., Goodman, I., & Hemphill, L. (1991). Unfulfilled Expectations: Home and School Influences on Literacy. Cambridge, MA: Harvard University Press.

Weaver, C. (1990). Understanding Whole Language from Principle to Practice. Portsmouth, NH: Heinemann.

Weiser, I. (1992). Portfolio practice and assessment for collegiate basic writers. In Yancey, K. B. (Ed.), Portfolios in the Writing Classroom. Urbana, IL: National Council of Teachers of English.

Yancey, K. B. (Ed.). (1992). Portfolios in the Writing Classroom. Urbana, IL: National Council of Teachers of English.