REFLECTING ON ASSESSMENT: VALIDITY INQUIRY AS ETHICAL INQUIRY

Brian A. Huot & Ellen Schendel

As David Jolliffe points out in "The Moral Subject in Composition," ethics has long been a central part of rhetoric. In recent years, ethics has also become an increasingly popular topic for discussion among composition teachers and writing program administrators, and it has been at the forefront of discussions among those who work in writing assessment. The American Psychological Association, the American Educational Research Association, and the National Council on Measurement in Education regularly publish and update their joint publication, *Standards for Educational and Psychological Testing*, spelling out guidelines on a variety of issues related to educational measurement, including directives about what constitutes a valid measure of student ability, what sorts of credentials professionals in measurement should have, and what passes as ethical behavior in educational and psychological measurement. In recent years, the International Reading Association (IRA), the National Council of Teachers of English (NCTE), and the Conference on College Composition and Communication (CCCC) have published similar documents outlining their positions on appropriate and ethical testing procedures and ethical uses of test results.

While ethics in assessment has been the subject of various codes, guidelines, and positions, recent understandings of what "ethics" means for composition studies have been the subject of ongoing examination. In their introduction to *Foregrounding Ethical Awareness in Composition and English Studies*, Sheryl Fontaine and Susan Hunter point to three trends in the way that educators address ethics: as a "code of conduct," as a site of critical

reflection, and as a "process of inquiry." Published guidelines from professional organizations would certainly fit the first category, and it appears that organizations such as IRA, NCTE, and CCCC have joined the ranks of measurement professionals in composing public documents that outline suitable behaviors for constructing and administering assessments. It is important that organizations representing teachers have begun to articulate what they think should be the roles of assessment in our writing programs and classrooms, and what our ethical obligations are to our students. But while these documents are important public articulations of our responsibilities in constructing assessment procedures, these discussions about ethical assessments are not sufficiently useful to those who make decisions about the design of assessments and the uses of their results. First, these codifications provide a range of principles about what makes an assessment ethical, rather than with a model or method to work through issues in concrete, contextualized assessment situations. And second, presenting ethics as a code of conduct may promote ethics as regulatory discourse imposed upon composition professionals and institutions by outside authorities.

Fontaine and Hunter write that it is possible and productive to understand ethics as a site of critical reflection aimed at "construct[ing] an environment within which students and others become aware of important contradictions and conflicts that will, in turn, contribute to the emergence of an ethical exigency from which to act" (2). This reflection implies a revisiting of various codes, ideas, concerns, people, and situations. Critical reflection asks that we examine individual situations in order to understand how competing ideas can be used to generate best practices as we engage with our culture, our students, and our assessment programs. Donald Schon coined the term "reflection-in-action" to explain the way that professionals work through problems and arrive at solutions as they perform tasks. The role of reflection in the teaching of writing and education has been further articulated (Harkin; Phelps; Freire) to demonstrate that teaching involves not only the dissemination of knowledge, but its production as well.

As critical reflection, then, ethics provides the potential for proactive political and social movement within specific contexts. Ethics as reflection in education in general, and assessment in particular, requires that teachers and administrators examine their own responses to real-life issues and situations in order that they can become agents of change within their educational institutions. Ethical assessments, in this view, are those that allow teachers and administrators to articulate their concerns for literacy education, critically examining the assessment technology, its use, its social consequences, and its effects on all students' ability to learn.

Fontaine and Hunter describe a third way of regaining ethics in composition studies, a "process of inquiry"; in this third view, ethics shifts from a means of instigating ethical behavior to a focus on "adopting an ethical awareness" about teaching, administration, and research activities (4). In this sense, ethics is a way to see assessment as research—a set of procedures for gathering and analyzing information to make important decisions about students and educational practices, with mechanisms built into assessment procedures that enable critical reflection. As inquiry, ethics permeates decisions about who and what will be assessed, how these decisions affect the results of the assessment, and how decisions about the use of results affect stakeholders. Ethics as a "process of inquiry" recognizes that what is ethical may shift from context to context. Rather than establishing a set of standards against which ethical behavior can be measured, ethical assessments foster a means of examining "the process of measuring ethical behavior, the importance of the individual, and the particularity of each situation" (Fontaine and Hunter 4).

In the rest of this article, we examine the roles that assessment[1] plays in changing programs, departments, and culture to show the need for more field-wide discussion about ethics and assessment. We show that this discussion is necessary for three reasons: 1) historical, since assessment has been a powerful force of social change, influencing what gets taught and how; 2) political, since assessments are sites of power and resistance, inscribing and defining individuals, and thereby shaping their

opportunities and regulating their behavior; and 3) disciplinary, for writing assessment has been a legitimate form of knowledge production in English Studies. We link the concept of ethics as a "process of inquiry" to contemporary theories about the validity of assessments, arguing that assessments that do not include explicit reflective discussions among educators about impact of the assessment on the decisions made on the basis of the assessment may suffer a low degree of validity. Current conceptions of validity focus not on the test itself; tests are not either valid or invalid. Rather, validity is a matter of degree, referring to the decisions made on behalf of an assessment (Cronbach; Messick). Further, "examining the intended interpretations and effects (uses), as well as those that are unintended" is central to establishing a test's validity (Moss, "'Testing'" 115). That is, the consequences of an assessment for individuals being tested and for the community in which the procedure takes place are important to consider in determining the validity of any assessment.

Our assumption in this article is that assessment must be discussed in the context of ethics, for the consequences of assessment procedures are closely tied to the political and social contexts in which they take place. Therefore, politics and ethics must be considered if we are to construct socially responsible assessments. We must be aware of the social and political impacts that assessment technologies have on the individuals who are assessed, on the institutions that sponsor such assessments, and on the larger culture in which such assessments take place.

Assessment as Social Action

Assessment is a form of social action, whether that action is to constrain and inscribe individuals or to liberate and empower them. Originally, assessment was designed to be a positive form of social action, disrupting existing social order and class systems. For example, the first formal assessment systems initiated in ancient China (Hanson) and nineteenth-century America (Witte, Trachsel and Walters) were intended to interrupt existing practices of awarding civil service appointments and educational

opportunities based upon social position, family connections, or other priorities unrelated to personal merit, achievement, or ability. But inevitably, these assessments replicated the social injustices of the cultures in which they were used.[2] Michel Foucault's *Discipline and Punish* illustrates how the examination is both hegemonic (a way that individuals are subjected to power) and punitive (a denial of agency and opportunity). A prime illustration of Foucault's theory of examinations is the recent book *The Bell Curve: Intelligence and Class Structure in American Life* (Hernstein and Murray), in which African Americans' lack of access to education and other important cultural institutions is defended based upon their lower test scores. Test scores become simultaneously the means of constructing an identity of deficiency for a group of people and a means of punishing them based on that construction. A second, less recent example comes from the earlier part of this century when Louis Terman, the primary developer of the Stanford-Binet I.Q. tests, renormed the instrument after initial results showed that girls had largely outperformed boys (Darling-Hammond). In this case, when the test ceased to construct males in the way the culture supported, the test was modified to reflect the sexism of the society in which it was administered. Because of testing's problematic history, assessment is regularly viewed by compositionists as a largely hegemonic exercise invested in reinscribing current unequal power relations in American society.

This overall impression of assessment is exacerbated in composition, since one of the driving impulses in the disciplinary formation of the field in the 1970s was to define itself against current-traditional rhetorical practices that emphasized correctness. Writing assessment was particularly implicated because it was the means by which such pedagogies were enforced. Composition's response to the negative effects of assessment has been, for the most part, to avoid it. One way to view this rejection of assessment is as a response to the "code of conduct" form of ethics, since the people who were constructing writing assessments were following a code of conduct in conflict with the

goals of the discipline. The result of these competing values and of composition's disinterest in assessment has been widespread adoption of holistic scoring developed by testing companies with theoretical and epistemological positions irrelevant to composition's values in language acquisition and literacy instruction.

However, compositionists have recognized that leaving assessment in the hands of non-composition professionals is problematic. Kurt Spellmeyer argues against suspect claims that education is in "crisis" or is in "decline," since much of the support for those arguments centers on test scores that do not test what we value in reading and writing.[3] Arguments about a literacy crisis often rely on assumptions of "teachers as lazy and incompetent and of students as lazy and ignorant" (179). Spellmeyer points to how these conservative attacks on new forms of teaching and new topics of study are ways that a powerful elite can control who has access to education. Spellmeyer recognizes the inherent nature of social action in assessment, since debates about assessment theory and practice are often conflicts over who will establish and maintain what counts as knowledge in our classrooms. Understanding writing assessment as a site of high-stakes political struggle is especially important because "unless we [as professionals] can play an active role in the fashioning of policy, there is little chance that we will retain control over what we do as teachers or as scholars" (179). Such involvement in assessment procedures means that professionals within the field need to become educated in the political and legal issues surrounding local, state, and nationwide assessments. Spellmeyer's call to action is part of a larger, common call in composition studies for writing teachers and administrators to become more active in writing assessment. Many recent articles about validity inquiry, portfolios, and program assessment end with the argument that more teachers need to actively change assessment practices.[4]

However, recognizing assessment as a political arena in which teachers need to become active constructors of knowledge is not enough. Assessment is the site at which decisions are made on

policy and the distribution of things of value, such as decent wages for teachers and material resources for schools and students. Recognizing assessment as a site of power and acknowledging the material consequences of assessments make clear how important it is for teachers of writing to engage in producing this sort of political discourse.

Focusing on assessment as a means by which teachers, along with students and parents, can effect progressive political action from within institutions offers a way to configure assessment in new and empowering ways. Viewing assessment proactively is a departure from educators' responses to assessment that are largely reactive—viewing assessment as something teachers have to do to satisfy university administrators, professional organizations, and governmental organizations. If teachers learn to be proactive rather than reactive, constructing their own assessment agendas and practices with their theories about literacy in mind, then they can control the politics of what will be valued within programs and institutions. Assessment can become a means for programmatic, institutional, and cultural change that values and invites teachers and students to study and critique discourses of assessment and act as rhetoricians of social change—goals that fit teaching writing as social change.

Making Assessments Proactive: Assessment as Reflective Practice

The goal of writing instruction in post-process classrooms is to help students to become critical thinkers about the discourses that make up the world around them (Berlin; Bizzell; Cooper). When students think critically in the classroom about the discourses that shape their world, they can become agents of social change within their own institutions and in the culture at large: "The study of rhetoric. . . may be said to contribute to social change. If people learn how better to control persuasive discourse, they can use it better to make the changes in the world that they desire" (Bizzell). Historical overviews of assessment (Hanson; Williamson; Yancey) make clear that it is one of these "persuasive

discourses" that should be examined, for it is often the means by which our work within universities and our roles as teachers and students are constructed. Patricia Bizzell contends that "[V]alues and cultural content must be addressed in the classroom—they are neither out of bounds nor merely incidental (e.g., to be used as 'prompts'), but very much part of the central pedagogical business." It is important, therefore, to address the discourses of assessment—perhaps the practice that most defines the school context—with our students and others. As Pamela Moss argues, "It is important to study the actual discourse that occurs around the products and practices of testing—to see how those whose lives a testing program impacts are using the representations (interpretations) it produces" ("Testing" 119). Studying the discourse and assumptions that make up our assessment practices raises a number of questions: What are the implications of our assessment techniques? How does the discourse surrounding a student being "placed" into a writing course or a program assessing itself implicate the work that happens within a program? What do different forms of assessment mean to students and teachers? What do students and teachers view as the implications of the assessment for literacy? Education? Our culture?

To highlight how such questions can guide our assessment practices, we discuss the implications of viewing ethics as a "process of inquiry," linking ethics to validity theory and to Donald Schon's notion of reflection-in-action, which we described in the introduction. We propose a methodology for assessment validation that is ethical, context-bound, and rooted in practice.

The guiding principles behind sound measurement practices have revolved around establishing their validity, which until the 1950s was usually represented by simple statistical relationships. Traditionally, validity had been the degree to which a test measured what it purported to measure. Since the 1950s, however, theories of validity have undergone important changes —changes that are compatible with evolving notions of literacy and language in a postmodern world. Unfortunately, however, as Pamela Moss ("Testing") points out, compositionists seem to be

isolated from the larger educational assessment community. Although theories of validity have changed significantly over the past fifty years, composition is to either rely upon old theories of validity or ignore measurement theory altogether. But measurement theory offers a basis for constructing sound assessment practices that allow for ethics as a "process of inquiry."

Lee Cronbach and Samuel Messick have been separately revising their notions of validity since the 1950s. For Cronbach, validity "must link concepts, evidence, social and personal consequences and values" (4). Messick's version is similar: "Validity is an integrated evaluative judgment of the degree to which empirical evidence and theoretical rationales support the *adequacy* and *appropriateness* of *inferences* and *actions* based on test scores or other modes of assessment" (13). These two definitions provide a much broader basis for thinking about the validity of decisions based upon an assessment measure than did earlier definitions that focused only on whether a test reliably (consistently) tested what it purported to test. Importantly, measurement theorists now talk not of validating tests, but rather of validating inferences and actions (note Messick's use of italics in the above quotation to emphasize this). In this newer conception of validity, it is a process by which researchers and teachers can reflect on the decisions they make based on an assessment's results. That is, this understanding of validity links the evidence gathered from an assessment with how stakeholders are affected by the way the assessment's results are used. For instance, while an older way of establishing the validity of a placement assessment might focus only on how accurate the placements were compared to previous methods of placement, a newer way of establishing validity might also focus on the experiences of students and teachers in the program after the placement occurs, as well as on how that means of placement results in progressive social action within the writing program.

This sort of validity inquiry relies upon reflective practice. Theorists like Cronbach and Messick refer to this reflective activity as "rival hypothesis testing," in which alternative inferences

are offered as explanations of the evidence available from the assessment. Moss extends this concept of rival hypothesis testing (which she calls "epistemic reflexivity") as something that happens "by subjecting the beliefs and practices of researchers to the same critical analysis of those of the researched" (121). Understanding ethics as reflective practice, "a process of inquiry," means that as we validate the decisions we make about our students' performance as writers, we must also look to our actions as teachers, administrators, and test makers.

The results of any given assessment are not straightforward but complicated by contextual and social issues surrounding the test, and recent discussions of validity require a consideration of these environmental and contextual factors. In fact, Messick stipulates that validation inquiry should include an examination of the "relevant aspects of the examinee's environmental and experiential background" (14). For English studies, contextual factors would include some commonly held principles about reading and writing texts. Theories about how language and learning are acquired differently in various contexts become important considerations for those looking to validate decisions made on behalf of assessments that purport to measure validly the reading and writing skills of students in our schools and colleges.

In her review of the literature on validity, Lorrie Shepard argues that investigations of validity should begin with the values and theories that inform a given method of assessment and then move directly into a discussion of what those values mean for the consequences of that assessment: "[T]o the extent that contending constituencies make competing claims about what a test measures, about the nature of its relations to subsequent performance in school or on the job, or about the effects of testing, these value-laden questions are integral to a validity evaluation" (428). Like Shepard, Moss is ultimately concerned with the ways in which evaluations impact stakeholders and the community at large:

> We need to expand our conception of validity to include questions about why particular methods of inquiry are

privileged and what the effects of that privileging are on the community. The choice of one or more methods of inquiry must not be taken for granted. It reflects, as Cherryholmes (1988) has noted, effects and exercises of power. ("Shifting" 254)

These measurement theorists recognize that assessment is a political act, for as Messick observes, "validation occurs more and more frequently in a political context" (14). Ethically speaking, then, we are encouraged by current standards in educational measurement and composition studies that argue that investigations into the validity of decisions made on behalf of a test should also become investigations into the theories and practices surrounding the test. Examining the validity of an assessment means looking at the assessment as only one part of a political and social act within a system of education that has consequences beyond those immediately visible to the people involved in the testing situation.

Richard Haswell articulates this process of validating assessments as "probing" into assessment. He argues for a four-part methodology to study the validity of program assessments. His methodology includes a focus on "use" or "social utility"; obtaining several different "perspectives" that allow for "the complexity of the [assessment] situation"; "cross-checking" in order to triangulate data; and "probing" or "bringing back partial returns from partial samples that need to be synthesized with samples from parallel but not exactly replicating probes" (93-94). Validation such as Haswell's system of multiple inquiry generates lots of data and allows conflicting viewpoints and contradictory results to come to light. Working out the messiness of multi-vocal, multi-faceted assessment allows for the generation of many competing hypotheses and theories even as it works toward validating a particular method of assessment.

In her response to Haswell's system for validation, Moss ("Testing") points out that missing from Haswell's methodology is a consideration of "any evidence of what the test scores actually

mean in terms of students' writing capabilities and whether this meaning is coherent with other aspects of the program" (116). Moss goes on to explain that without understanding the context of Haswell's particular situation—what writing instructors value in their students' writing, and what the differences and similarities are among teachers in determining these kinds of values—and without understanding the effects Haswell's form of placement have on students (the value of the education they receive in whatever course they are placed in, how these placements help or hinder students' educations), it is difficult to assess the validity of the assessment. The elements Moss adds to Haswell's method of inquiry are important, for they force the issue of ethics to the surface, allowing for discussions about what each instructor values in student writing, the assessment method at hand, and the role the instructor plays in that assessment and its effects. Concerns about what individual instructors value in literacy education go straight to the heart of linking individual pedagogies to one another, to the writing program and its goals, and to future permutations of the assessment methodology. Together, Haswell's and Moss's views of validity make this "process of inquiry" a means of facilitating program-wide reflective practice in assessment. Assessment becomes a way that teachers can reflect upon their teaching and on their students' accomplishments. If teachers ask themselves and one another what they value about reading, writing, pedagogy, and theory, and if teachers reflect upon their classroom experience and their individual theories about teaching and learning in order to discuss what works and doesn't work about a particular curriculum and what sort of role assessment plays in all of these considerations, then assessment becomes a means way of fostering reflective teaching—better teaching and better learning.

Dialogue and debate are perhaps the most important elements of validity inquiry, for it is through such interactions that individuals reflect on the motivations, values, and beliefs that influence the way they construct testing procedures. Discussions that center on making changes to assessments, curricula, and

pedagogies are, in essence, discussions that make assessment an ethical "process of inquiry." Professionals, Schon argues, function by "knowing-in-action," through reflection and intuitiveness that comes through experience (49). Hillocks, in *Teaching Writing as Reflective Practice*, stresses the ways in which writing teachers can use their intuitive and experiential knowledge to formulate a theoretical base to inform their pedagogical repertoire. Hillocks describes the dialectical nature of reflective practice:

> Any moment in teaching may be thought of as a trial with similar ramifications. Every question has multiple meanings and implications. Every planning and every teaching decision allows some things and denies others. Each deserves subjection to comparable analysis of relationships and benefits and, whenever possible, the careful evaluation that reflective planning and teaching allow. (190)

Hillocks's description of reflective practice sounds strikingly like Moss's and Shepard's descriptions of construct validity, where the effects of a particular technique are analyzed, the results interpreted and reflected upon, and the technique modified in order to more accurately reflect the needs of the stakeholders. Just as Moss ("Shifting"; "Testing") and Shepard view an assessment procedure as but one consideration of its validity, Hillocks contends that a particular teaching method is only one element of an entire situation that should be reflected upon. However, Hillocks also points out that "most of the reflection that occurs in teaching is not available to anyone other than the teacher" (202). Framing test validation as an ethically driven "process of inquiry" makes reflection a public act required of an entire community involved in the construction and administration of an assessment and the use of its results in educational decision-making.

In validity inquiry, then, reflection becomes a collaborative act, a means by which members of a community participate in the social construction of knowledge, power, and values. As such, reflection can become a means of social change. As a group of

people come to an understanding about an assessment and its effects within their community, the expectations and standards about what is important to them as individual teachers and administrators change. This potential of collaborative reflection is reminiscent of Anson's assertion that teachers writing reflections for inclusion in their teaching portfolios benefit when those reflections are shared in groups. In the discussing and revising of teaching portfolio documents, Anson sees the possibility for "changed attitudes and improved teaching strategies" (189). Likewise, as it encourages reflective practice, validity inquiry allows for the assessment to shift as the values and theories of those instituting/constructing the assessment shift. Assessment as reflective practice is fluid rather than immutable or resistant to change; instead, assessment becomes a site where reflective teachers can shape future assessments as they reflect upon those in the past.

Inquiry into the validity of a test yields much more than an answer about whether that test is an adequate assessment tool; it also brings to light the values of the particular testing situation, assessment in general, and the field of composition and English instruction as a whole. Investigating validity is inquiry into both the theories that shaped the test and the effects of its results on stakeholders and the community at large. Therefore, validity inquiry can be both research into what has already happened as a result of the assessment and an exploration of what might need to change in the future. In this sense, validity inquiry is reflective, informative, and fostering a spirit of ethics as a "process of inquiry." It is the means by which assessors can examine their own values and ethics in relation to those of the assessed. It is a way to open up dialogue between those who administer an assessment and those who are affected by it—teachers and students. And, inquiry into the validity of an assessment becomes the means by which a community—assessors, an English department, researchers—can socially construct knowledge about a procedure through dialogue and debate, even *revising* the discourse of assessment for that community. This constant reflection and reevaluation of

assessment and its discourse construct assessment as a "process of inquiry" that is inherently collaborative, public, responsible and ethical.

Conclusion

Clearly ethics should be an important consideration for all educational practices and especially within the practice of assessment. As we have tried to demonstrate in this essay, conversations about ethics and assessment need to go beyond the mere construction of codes of conduct. Recognizing assessment as a means to foster critical inquiry into literacy is an important step in using assessment to promote equity of opportunity for all students. It is also an important step in helping to reduce the hegemonic potential for educational assessment. But assessment is productive social action only if we recognize it as a "process of inquiry"—a process dependent on reflective practitioners mindful of the power of discourses and practices to influence the material realities of their jobs and their classrooms. This negotiable, reflective conception of assessment is an important counterpart to the monolithic, mandated nature of many assessment programs.

We hope that our reconception of assessment serves as an invitation for teachers and administrators who usually shy away from being involved in assessment to construct localized, reflective practices for their own institutions. Assessment can be a way to open up dialogue among professionals who want to become agents of change in their programs and institutions. Through validity inquiry—by looking backwards on what has happened and forward towards what is possible within specific contexts—stakeholders can come to a community understanding of what that particular assessment means for their particular context and what the effects of that assessment can and should be on their teaching and their students.

END NOTES

[1] Although we recognize and understand the importance of classroom-based assessment and the potential applications of our discussion for the classroom, in this article we limit our focus to those assessments such as placement, exit, and proficiency that typically occur outside a classroom context.

[2] In defining assessment as a technology, George Madaus writes that "technology is not by nature socially unjust. It is, however, inextricably intertwined with the distribution of wealth, race, and gender relations (Sclove, in press). Since technologies are a product of the existing structure of opportunities and constraints in society, they are likely to extend, shape, rework, or reproduce this structure" (79).

[3] Grant Wiggins notes that perceptions of lowered standards are often conflated with the move to standardize curricula, instruction, and assessment.

[4] See the final paragraph of the following articles for a representative sample: Selfe (1997), Huot (1996), Durst, Roemer, and Schultz (1994), and Lowe and Huot (1997).

[5] Durst, Roemer, and Schultz's article "Portfolio Negotiations: Acts in Speech" highlights how assessment can be reflective practice. In the portfolio text assesement they describe (in place at the University of Cincinnati), trios of teachers read and discuss portfolios to determine whether students in the first-year writing curriculum should successfully pass out of the program. But Durst, Roemer, and Schultz also found the talk generated by the assessment— wherein teachers voiced concern with the assessment system, reflected on their own teaching, and examined their own student's writing in light of portfolios generated by students across the program—has been a useful addition to their ongoing teacher education program.

WORKS CITED

American Psychological Association, American Educational Research Association, 1985. *Standards for Educational And Psychological Testing.* Washington, DC: American Psychological Association, 1977.

Anson, Chris. "Portfolios for Teachers: Writing Our Way to Reflective Practice." *New Directions in Portfolio Assessment: Reflective Practice, Critical Theory, and Large-scale Scoring.* Ed. Laurel Black, Donald Daiker, Jeffrey Sommers, and Gail Stygall. Portsmouth, NH: Boynton/Cook. 1994. 185-200.

Berlin, James. "Rhetoric and Ideology in the Writing Class." *College English* 50.5 (1988): 477-95.

Bizzell, Patricia. "Rhetoric and Social Change." *Conference on College Composition and Communication*. Phoenix. http://www.hu.mtu.edu/cccc/97/bizzell.html. 1997.

Conference on College Composition and Communication. "Writing Assessment: A Position Statement." *College Composition and Communication* 46 (1995): 430-37.

Cooper, Marilyn. "Postmodern Ethics in the Writing Classroom." *Conference on College Composition and Communication*. Phoenix. http://www.hu.mtu.edu/cccc/97/cooper.html. 1997.

Cronbach, Lee J. "Five Perspectives on Validity Argument." *Test Validity*. Ed. Harold Wainer. Hillside, NJ: Lawrence Erlbaum, 1989.

Darling-Hammond, Linda. "Performance-based Assessment and Educational Equity." *Harvard Educational Review* 64 (1994): 5-30.

Durst, Russel K., Marjorie Roemer, and Lucille Schultz. "Portfolio Negotiations: Acts in Speech." *New Directions in Portfolio Assessment: Reflective Practice, Critical Theory, and Large-scale Scoring*. Ed. Laurel Black et al. Portsmouth, NH: Boynton/Cook, 1994. 286-300.

Fontaine, Sheryl, and Susan Hunter. "Ethical Awareness: A Process of Inquiry." In *Foregrounding Ethical Awareness in Composition and English Studies*. Ed. Sheryl Fontaine and Susan Hunter. Portsmouth, NH: Boynton/Cook, 1998.

Freire, Paulo. *Education for a Critical Consciousness*. Trans. Myra Bergman Ramos. New York: Continuum, 1973

-----. *Pedagogy of the Oppressed*. Trans. Myra Bergman Ramos. New York: Continuum, 1990.

Foucault, Michel. *Discipline and Punish: The Birth of the Prison*. New York: Pantheon, 1977.

Hanson, F.Allan. *Testing, Testing: The Social Consequences of the Examined Life*. Berkeley: University of California Press, 1993.

Harkin, Patricia. "The Postdisciplinary Politics of Lore." *Contending with Words: Composition and Rhetoric in a Postmodern Age*. Ed. Patricia Harkin and John Schilb. New York: Modern Language Association, 1991, 124-138.

Haswell, Richard. "Multiple Inquiry in the Validation of Writing Tests." *Assessing Writing* 5 (1998): 89-109.

Hernstein, Richard, and Charles Murray. *The Bell Curve: Intelligence and Class Structure in American Life*. New York: Free Press, 1995.

Hillocks, George. *Teaching Writing as Reflective Practice*. New York: Teachers College Press, 1995.

Huot, Brian. "Toward a New Theory for Writing Assessment." *College Composition and Communication* 47 (1996): 549-566.

International Reading Association and National Council of Teachers of
English. *Standards for the Assessment of Reading and Writing.* Urbana: National
Council of Teachers of English, 1994.

Jolliffe, David. "The Moral Subject in College Composition: A Conceptual
Framework and the Case of Harvard, 1865-1900." *College English* 51
(1989): 163-173.

Lowe, Teresa, and Brian Huot. "Using KIRIS Writing Portfolios to Place
Students in First-Year Composition at the University of Louisville."
Kentucky English Bulletin 46 (1997): 46-64.

Madaus, George. "A Technological and Historical Consideration of Equity
Issues Associated with Proposals to Change the Nation's Testing Policy."
Harvard Educational Review 64.1 (1994): 76-95.

Messick, Samuel. "Validity." *Educational Measurement, third edition.* Ed. Robert
Linn. New York: American Council on Education and Macmillan, 1989.
13-103.

Moss, Pamela. "Shifting Conceptions of Validity in Educational Measurement:
Implications for Performative Assessment." *Review of Educational Research*
62 (1992): 229-58.

-----. "'Testing the Test of a Test': A Response to the Multiple Inquiry in the
Validation of Writing Tests." *Assessing Writing* 5 (1998): 111-22.

Phelps, Louise Wetherbee. *Composition as a Human Science: Contributions to the
Self-Understanding of a Discipline.* New York: Oxford, 1988.

Schon, Donald. *The Reflective Practitioner: How Professionals Think in Action.* New
York: Basic Books, 1983.

Selfe, Cynthia. "Contextual Evaluation in WAC Programs." *Assessing Writing
Across the Curriculum: Diverse Approaches and Practices.* Ed. Kathleen Blake
Yancey and Brian Huot. Greenwich, CT: Ablex, 1997, 51-68.

Shepard, Lorrie. "Evaluating Test Validity." *Review of Research in Education* 19
(1993): 405-450.

Spellmeyer, Kurt. "Response: Assessment as Surveillance." *Assessment of
Writing.* Ed. Edward M. White, William Lutz, and Susan Kamusitri. New
York: Modern Language Association, 1996, 174-181.

Williamson, Michael M. "An Introduction to Holistic Scoring: The Social,
Historical, and Theoretical Context for Writing Assessment." *Validating
Holistic Scoring for Writing Assessment: Theoretical and Empirical Foundations.*
Ed. Michael M. Williamson and Brian Huot. Cresskill, NJ: Hampton,
1993.

Wiggins, Grant. *Assessing Student Performance.* San Francisco: Jossey Bass, 1993.

Witte, Stephen P., Mary Trachsel, and Keith Walters. "Literacy and the
Direct Assessment of Writing: A Diachronic Perspective." *Writing
Assessment: Issues and Strategies.* Ed. Karen Greenberg, Harvey Wiener and
Richard Donovan. New York: Longman, 1986, 13-34.

Yancey, Kathleen Blake. "Looking Back as We Look Forward: Historicizing
    Writing Assessment." *College Composition and Communication* 50 (1999):
    483-503.