Evaluating Student-Created Hypertexts: What Do We Do With These Things???

Carl Whithaus  
English Department  
Old Dominion University  
cwhithau@odu.edu

Voices emerge, bend, discover. But how do we judge a voice? How do we judge how writing works? How do we judge the effectiveness of a piece of communication? Does computer-mediated communication provide a space for students’ voices, for a form of writing that not only allows but also encourages student inquiry, risk-taking, linguistic and communicative learning?

Yes and no.  
A home-schooled, twelve-year old in rural Pennsylvania types rapidly in Netscape Composer. He’ll soon open up AOL and send the paper off as a series of attachments to his teacher. The writing is clean, clear, and precisely Christian. He is learning to compose in a hypertext environment. While the content of his essay might be traditional, the medium is not—the form reaches beyond, elsewhere, towards the associative qualities of mind that hypertext theorists have been searching for since Vannevar Bush’s 1945 “As We May Think.” His inquiry turns on a Biblical passage, a moment of revelation, obedient, yet thoughtful in the way that I would like many of my college students to be thoughtful, reflective.

Over and over again the contributors to Cynthia Selfe and Gail E. Hawisher’s *Passions, Pedagogies, and Twenty-first Century Technologies* argue that information technologies alone do not make pedagogy “better.” New information technologies bring change to the everyday dynamics of writing classrooms, but these changes are determined by local conditions. In “Access—The A Word in Technology Studies” in the Selfe and Hawisher book, Charles Moran reminds us that the social class of the students attending an institution has a dramatic effect both on the information technologies used and on the ways in which they are used (206-209). The use of information technologies is determined by teachers’ proficiencies with the new technologies, by students’ access to and proficiency with computer-mediated communication, and by the way these new technologies are used to meet already existing standards for college writing.

As teachers, we must think hard about how we can assess communication and encourage learning in computer-mediated environments. How can we measure students’ abilities in valid and reliable ways without reproducing the engines of assessment that evaluate students’ deficits rather than what they know and can do?

The riches of teaching, of learning, and of being in the classroom are always more than the marks made on a test, the sound and the fury of SAT scores that signify….
In writing courses, students do not struggle toward language but are always already part of complex social and linguistic worlds. Advertisements, sports talk shows, religious organizations, Oprah’s book club, dystopic sitcoms, school. Students know how to listen. They are surrounded by what Don DeLillo describes as the white noise of who we are, where we are, when we are. Now the cul-de-sac of the twentieth century opens into the twenty-first, and within primary language arts classrooms, secondary English courses, and college writing courses language is used over and over and over again—even by the most caring teachers—in what Hugh Mehan long ago identified as the QRE formula. Teacher questions, student responds, teacher evaluates.

In his study of class discussions and teacher preparation, James Marshall identifies two conflicting ideals about interaction with students that many highly successful teachers have. By interviewing secondary school teachers, Marshall found that

on the one hand, teachers felt discussions were an opportunity for “interaction, a chance for students’ ‘self discovery’”… On the other hand, though, teachers also felt that discussion should “go somewhere,” should stay “on track” and away from “irrelevancies.” (41)

While the teachers claim that they want to allow the students to control classroom discussion, Marshall’s observations of their teaching lead him to conclude that these teachers, all with established reputations for excellence, “dominated most of the large-group discussions” (42). That is, the teachers voiced their concerns and maintained classroom order by controlling the ebb and flow of the discourse and by keeping the students focused on the teacher-selected subject matter. By the time students arrive in college composition classes, they have already been socialized into the modes of school discourse. They know the social behaviors that are appropriate. And they know that when an English teacher asks for discussion, it means controlled discussion. In “Getting Together, Getting Along, Getting to the Business of Teaching and Learning,” Margaret Cintorino recalls her own schooling in the “legacy of student silence.” She writes, “We learned to still our young clamorous voices, to be quiet, and to remain quiet for much of the school day. We inherited, from the beginning of our school years, a legacy of student silence” (23).

In this paper I offer thumbnail sketches for four methods of assessing student work in computer-mediated composition courses:

- **distribute** the process of assessment among a variety of readers;
- **situate** assessment within the dynamics and goals of an institution or a program;
- rely on **descriptions** of material in student writing and their own perceptions about how their reading experiences could or should have been different—and better—and;
- **interact** with the “community of practice” created by the students and teachers within a given course by anticipating and questioning the already internalized traditions of assessment students display from prior English classes.

These methods are implied and used to various degrees in many contemporary electronic portfolio projects; however, more often than not, they are discrete, separate methods. Nowhere have I seen them applied across an institutional assessment system. Still it is
classroom teachers who can begin to think about these methods, use them, develop them in relationship to classrooms and learners.

In the last five years, discussions about electronic portfolios have begun to connect assessment and computer-mediated communication in K-12 and higher education circles. Within critical pedagogy, assessment and evaluation are often seen as secondary issues to discussions about promoting social change through educational practices. However, the questions of literacy practices and social values that writers such as Ira Shor, Patricia Bizzell, and Barbara Herrnstein Smith explore bear directly on questions of evaluation and assessment. For instance, we see the convergence of these questions in the work of Brian Huot on context-based assessment. By drawing on the work of measurement experts such as Pamela A. Moss, Samuel Messick, and Lee Cronbach, Huot shows how validity has begun to move away from an empiricist notion of objectively determined accuracy toward acknowledging “the social construction of knowledge” (“Towards” 550). Huot quotes Cronbach’s claim that validity “must link concepts, evidence, social and personal consequences and values’ (4)” to demonstrate the redefinition of validity within the assessment community(550). This work within the assessment community suggests links with literacy scholars’ research and the localized writing testing programs used at the University of Pittsburgh and Washington State University (553-554). For Huot, assessment theory, literacy and reading theory and some composition testing programs suggest “emergent methods” of assessing writing that stretch older, accepted definitions of valid, reliable, and authentic assessment.

These emergent methods can be view under a new theoretical umbrella, one supported by evolving conceptions of validity that include the consequence of the tests and a linking of instruction and practical purposes with the concept of measuring students’ ability to engage in a specific literacy event or events…. These methods are sensitive to the importance of interpretation inherent in reader response and psycholinguistic theories of reading. (561)

Despite the five years that have passed since Huot articulated the basis for a “new theory of writing assessment,” the uses of validity and inter-rater reliability in the literature on writing assessment have remained bound by their traditional connotations. Validity tends to mean that the assessment tool matches the content to be covered; and, reliability tends to mean that different readers will produce the same score for the same essay or type of essay across contexts.

The discontent among composition specialists with current-traditional methods of writing assessment, and a more general discontent with standardized educational testing, has begun to be felt. Research within educational assessment and measurement by Eva Baker, Pamela A. Moss, Samuel Messick, and Lee Cronbach as well as the work with writing portfolios by Kathleen Blake Yancey, Pat Belanoff, Liz Hamp-Lyon, and Bill Condon has strengthened the theoretical groundwork that Huot pointed to in 1996. Taken in combination, this work suggests the need to continue to push on the conventional definitions of validity and reliability. The importance of redefining these terms for the assessment of student compositions in computer-mediated writing classrooms is, perhaps,
An interest in context-based assessment speaks directly to concerns within K-12 assessment and K-12 teaching communities. The increasing numbers of standards-based assessments imposed on students and teachers during the last five years are shaping our students’ experiences of education. These experiences socialize students into passive roles and lower their expectations for the relevance of literacy education to their daily lives. A move toward communication-based assessments in K-12 assessment could enable students to see themselves as active—rather than passive—agents in the construction of literacy, learning practices and social values.

In her discussion of K-12 testing, Eva Baker notes that current systems and theories of assessment do “not meet our expectations for guiding practice and improving learning” (15). As an expert on assessment, she makes it clear that her complaints about testing do not emerge from the usual ideological positions which resist “the capital letters TESTING INDUSTRY” and valorize “the wisdom and accuracy of classroom teachers’ judgments” (15). Even from her position within the testing community, Baker urges that current methods of assessment are failing to meet the needs of students, teachers, communities, and policymakers because they center on measurements of discrete skills. System validity and multipurpose testing are “heretical” ideas for the assessment community (16). Research in computer-mediated composition and electronic portfolios suggests that students draw on a variety of communication skills and interact with other writers and audience members in multifaceted ways. For the evaluations—either classroom-based or large-scale—of computer-mediated compositions to be valid and authentic, then, our assessment tools must take effective communication as their benchmark instead of discrete skill-based standards. In turn, this shift in assessment methods and criteria returns us to the questions of literacy activities and social values that underlie the work of Shor, Bizzell, and Herrnstein Smith. Teachers who use electronic portfolios incorporate principles of interaction and description into their evaluations;

1 Ira Shor has noted that
language arts are constant requirements for students from elementary grades through college, making language instruction the biggest and most closely-watched enterprise in mass education. From childhood through early adulthood, official language arts help to socially construct how students see the world and act in it (Pattison; Rouse). This socialization through curriculum (what Paulo Freire called “the banking models” of pedagogy) uses assessment and instruction as vast “sorting machines,” to borrow Joel Spring’s metaphor…. The tradition of complaint in first-year college writing is a product of the contention faced by the status quo in reproducing itself in each new generation. (“Illegal Literacy,” 105)

2 Baker argues that her thesis, that there is something wrong with our system of K-12 testing, does not flow from the same impulse as many such analyses. It is not developed as a critique of the factory model of education, the one that sees children as outputs and that is a vestige of the industrial age. It does not attack tests and their results as reductionist oddities. It does not compete with the findings of tests developed by the capital letters TESTING INDUSTRY against a sometimes more romantic view of the wisdom and accuracy of classroom teachers’ judgments of their students’ performance. Last, it will not deny that policymakers have the right and responsibility to demand testing programs that shed light on school progress and real policy options, and that such programs be developed on a schedule shorter than the Pleistocene era. (15)
however, the social contexts surrounding their evaluations and the students’ perspectives on each other’s work have not yet altered the processes of assessment. Assessment remains an activity based on standards established before the students begin to communicate and establish their own interpretative community. In a full-realized model of communication-based assessment, the four characteristics or attributes need to co-exist, and interrelate.

In “Toward a New Theory of Writing Assessment,” Brian Huot argues that it is through the growing theoretical concern for the context of writing in composition studies that many university writing programs are beginning to redefine valid assessment. For Huot, we are beginning to see the emergence of methods of writing assessment that eschew inter-rater reliability and instead base claims for accurate assessment upon contextually rich assessment schemes. Huot’s article as well as the claims he and Kathleen Blake Yancey make in “Construction, Deconstruction, and (Over)Determination: A Foucaultian Analysis of Grades” suggest that evaluation must be driven by the situation within which students work rather than outside standards. Proponents of portfolio assessments have long argued that portfolios are highly valid tools for evaluation because they can provide not only a context-rich but also a student-centered means of measuring learning (e.g., Roberta Camp “The Place of Portfolios in our Changing Views of Writing Assessment”). The argument that context-rich and student-centered assessments lead to better classroom practices is a central idea in much of the recent work on portfolio methods of assessment beyond college writing programs. For instance, Giselle Martin-Kniep worked with a variety of teachers on the Hudson Valley Portfolio Assessment Project, and observed that portfolios provided both teachers and students with increased control over the products that were assessed. That is, students were no longer limited by the format of a timed assessment in which they presented merely one performance to the evaluators. They were now able select from a variety of different performances developed over time. Students’ selection of appropriate works for evaluation helped link assessment with student choice. Portfolio assessment—particularly the grading sessions where evaluators can ask questions of the classroom teachers—presents one means of increasing the consideration of contexts in assessment processes as both Hamp-Lyons and Syverson have pointed out. Syverson cites a case in which elementary students had been working on research projects as part of their learning-record portfolios. An evaluator was particularly concerned about the paucity of sources used in a number of essays. When the outside reader asked the classroom teacher about the lack of research citations, the teacher was able to explain that for part of the semester the library had been undergoing renovations and students were unable to check out books. While the opportunity for classroom teachers to provide contextual information is not inherent in all portfolio systems, most portfolio assessments include a cover letter from the student to the reader. These cover letters attempt to provide contextual evidence for the reader-evaluators.
The developing body of work on electronic writing assessment suggests that there is the possibility—perhaps even the necessity—of creating methods of distributive, interactive, descriptive and situated evaluation that reflect the risks and complexities of student-generated, computer-mediated compositions. Each of these methods of assessment reflects a growing understanding of evaluation as a process of communication. Distribution acknowledges that different readers will respond differently to texts and does not smooth out these differences but instead incorporates multiple and distinct responses. Interaction argues that assessment should include feedback and negotiation. Description allows a reader to respond from her experience as a reader instead of through a filter of criterion- or standards-based assessment. And situated readings actively work to include—rather than exclude—social contexts. None of these methods outline discrete skills or areas for measuring writing ability; rather as an amalgam of approaches they sketch the groundwork for treating evaluation as communication.

Works Cited


