

Defining Assessment as Research: Moving from Obligations to Opportunities

Peggy O'Neill, Ellen Schendel, and Brian Huot

Writing assessment is a large and complex part of the administration of a writing program; however, many writing program administrators (WPAs) are not prepared to work with assessment. This lack of preparedness among WPAs is most saliently illustrated by the number of frantic posts sent to the WPA listserv (WPA-L) requesting help in arguing against an assessment already in place or in working through the complexities of setting up an alternative assessment system. The post below typifies this kind of request:

Can you help settle a difference of opinion between essay placement graders and our tests and measurements people? We currently give a 35-minute essay prompt during pre-freshman placement testing during the spring and summer. There are seven different testing dates and we like to supply seven different prompts to be administered. The essays are administered on Saturday and graded holistically on Sunday. There are anywhere from 1200 to 1800 students taking the essay at any given time.

The tests and measurements folks want to reuse a prompt for two sittings in order to determine reliability. Our graders feel that if they have to grade 3,000 essays on the same prompt they will go nuts, glaze over, and, consequently, that this will affect their reliability. It seems to me that they will lose the ability to discriminate sufficiently well by the time the prompt is repeated, but [the] tests and measurements [people] are adamant that only by repeating the prompt can reliability be assured. They are willing to listen to expert opinion, however.¹

In this example, what the WPA identifies as “a difference of opinion” is really a difference in values and theoretical perspective. He seems not to understand the psychometric concept of reliability and the role it plays in the validation of an assessment. He does not seem to understand the concepts and values of the “tests and measurement” specialists, and yet he finds himself responsible for designing and administering the test. This lack of knowledge makes his job as a WPA more difficult and keeps him from communicating with the testing and measurement specialists on his campus effectively and persuasively about the specific needs of writing assessment. In short, while he may have expertise in writing program administration and writing theory—experience that should be of primary importance in the design and implementation of any writing assessment system—he is shut out from being an integral force in shaping the assessment to the particular context of the program, and he lacks the knowledge and discourse necessary to communicate with the other assessment stakeholders on campus.

Here is a second example of a rather typical WPA-L post in which a new WPA is overwhelmed as her department works to design a new placement test:

Here I am—new member and already asking for a favor. Several members of the English department feel we have a window of opportunity (this fall) to institute an essay placement exam. We have been using ASSET. I have relatively little time to put together a proposal for administration and am woefully lacking in assessment expertise. Any ideas would be most gratefully received.

In this second example, the WPA is, in her own words, “woefully lacking in assessment expertise,” and yet responsible for proposing a new placement test. Convincing upper-level administrators and one’s own colleagues that a new assessment system will be better can be difficult. Actually developing a system that is better than the one in place takes time, careful planning, and expertise; convincing others that the system works also takes time, and, as the post makes clear, there is no time to properly assess the assessment situation in order to identify alternatives.

This third example addresses the political aspects of assessment that often influence—or dominate—college writing assessment and underscores the fact that WPAs are expected to participate in the design of large-scale writing assessments whether they have the expertise or not:

My university started a large-scale portfolio assessment project without much of a plan of how they were going to assess the data. Now that the project has been under-

way for four years and the archives are bulging, the assessment committee wants to begin scoring, and they're looking around (with increasing anxiety) for help. They want to examine student outcomes, more specifically, to determine whether or not students become better writers over the course of their undergraduate experience. [. . .] I have no training in large scale assessment, and I'd appreciate any thoughts you have.

In all of these posts, writing assessment is positioned as an administrative activity, not a knowledge-constructing one. In some ways, this positioning is understandable because writing assessment is often seen as an unpleasant task—irrelevant at best and punitive at worst—that is imposed on us and controlled by someone else—a task that WPAs often have to approach without much preparation or expertise. Add to the mix the fact that assessment can cost quite a bit in terms of time, money, and human resources, yet seemingly produces few benefits to individuals within a writing program and can have many detrimental effects (Moss 10-11).

As illustrated by these WPA-L posts and reinforced by our experiences at conferences, discussions about assessment tend to emphasize practice and pragmatics over theory and assumptions, which some assessment scholars lament. It is not uncommon for a panicked message like this one to appear on WPA-L:

I'd like information ASAP (we're meeting next Monday on this!) on the following: 1) If you have prompts that you've used successfully for timed essays that would work for a large scale portfolio assessment, I'd love to see them; 2) If your department has developed useful criteria for such prompts, I'd love to see them; 3) If you know of a good article that addresses this issue, send me the citation!

Other subscribers to WPA-L usually rush to the rescue, citing relevant articles, offering to fax documents related to the assessment procedures at their own universities, referring readers to an appropriate website, directing the person to the WPA-L archives, or offering this-is-what-we-do-here suggestions to help the colleague who is in "trouble." The implication of this sort of response is that WPAs who now find themselves in charge of writing assessment need only to follow the models of assessments in other contexts and read an article or two to know enough about assessment to establish an effective system for the institution.

Focusing solely on writing assessment's technical aspects (e.g., designing a prompt) and the logistics of processing data (e.g., organizing a scoring session) often reinforces a simplistic approach to writing

assessment. While the composition community's generosity in helping each other is laudable, what is often lacking in these exchanges is an acknowledgment that there is more to assessment than knowing how to gather and score student writing. And even when that is acknowledged, the immediate demands and short time frame usually require that the WPA in need of help jump into the assessment process without time for careful planning and research. Although we may help each other satisfy our immediate needs in responding to calls for help, we are also promoting an uncomplicated, practical approach to the assessment of writing that cannot only belie the complexity of assessment but also make ourselves, our programs, and our field vulnerable to the whims of administrators and politicians because issues of power, values, and knowledge-making converge on assessment sites, with very real consequences to all stakeholders.

While the work of WPAs certainly encompasses more than writing assessment tasks, assessment needs to be considered a central concern of WPA work because the effects of writing assessments on students, teachers, and curriculum are far-reaching: assessments define good and bad writing; they promote certain pedagogies and discourage others; they have real-world consequences for students and teachers; they function to endorse certain positions or sites with authority; and they define our values—whether accurately or not—to others. Writing assessment does have many administrative aspects to it, but there are other aspects as well. It has the potential to contribute even more toward establishing the discipline of composition and changing the realities of teachers and students. Writing specialists need to be knowledgeable about assessment theory and practice whether or not they are involved in administration because assessment has material effects for us as well as our students; however, WPAs must have this knowledge to participate in assessment's intellectual work and to ensure the validity of writing assessments.

Recasting the role assessment plays in writing programs and in composition studies requires that WPAs revise their own approaches to assessment theory and practice. Richard Haswell, Michael Williamson, and William L. Smith are among those who have done scholarly work in assessment that acknowledges its complexity and potential, but these scholars are not a representative sample of those who are responsible for creating, implementing, or overseeing most college-level writing assessments (Murphy et al; Huot "Survey"); the dominant view of writing assessment among WPAs seems to be that it is only an administrative obligation.

We argue here that by conceptualizing writing assessment as research, WPAs will be able to engage in intellectual work that makes knowledge about writing, writers, programs, and pedagogy. Further-

more, by acknowledging that writing assessment is a form of social action because it results in real-world consequences for people and programs, we can promote our values and theories while limiting the costs of assessment to ourselves, our students, and our programs.

ASSESSMENT AS RESEARCH: ASKING QUESTIONS, FINDING ANSWERS

In writing assessment—as with other areas of assessment—the emphasis has been on the technical aspects, such as calculating interrater reliability, designing prompts, organizing norming sessions, and running scoring sessions. This focus on the technologies of writing assessment has affected the way it has been used and studied. For example, until quite recently the overwhelming objective in writing assessment was to produce scores that approximated a certain level of interrater reliability—that is, consistency of scores on the same essays from independent raters. Interrater reliability is a problem that has plagued—and preoccupied—writing assessment since the 1910s (Starch and Elliott; Cherry and Meyer). Over several decades, procedures such as the construction of a scoring guide or rubric and the training of raters on that rubric became common practice, and it is through these procedures that we have been able to produce consistent scores. These procedures—what Brian Huot calls the technology of writing assessment (“Toward”) and what Edward M. White calls “the machinery of holistic scoring” (96)—have been the main concern in writing assessment, and their predominance leads to the fallacious but widespread conclusion: If we follow these procedures—or recipes—and produce enough consistency in scoring, then we have done a good job and our assessment is valid. Validity, however, is a much more complicated and contentious concept that is not the property of a particular test. Arguing for validity may include addressing reliability issues, but reliability is not the only issue involved. In fact, Pamela Moss argues that in writing assessment, we can have validity without reliability. Many WPAs, however, don’t demonstrate a nuanced understanding of validity and its central role in assessment and therefore overemphasize or misconstrue issues of reliability.

By viewing writing assessment as research, as a way to ask and answer questions about our students, their writing, our teaching, our curricula, and the other factors that constitute effective writing instruction, we can move beyond reliability and toward constructing a validity argument. This different orientation also provides a way to reimagine writing assessment. Instead of it being something imposed upon us, something we have to do or have done to us, assessment becomes a way we can research answers to legitimate questions about how instructors, students, administrators, and programs are doing. In this approach, emphasis on

technical concerns such as reliability is reduced, although that does not mean that we should ignore the importance of scoring consistently. In fact, we might start to rethink what is meant by consistency, as William L. Smith did (“Assessing” 187-97). From over a decade of research on placement at the University of Pittsburgh, Smith discovered that raters who teach similar classes and have similar experiences often agree at a greater rate than those only trained on a holistic scoring rubric (“Assessing” 189-92; “Importance”). Based on his findings, Smith abandoned holistic scoring and developed alternative assessment methods that relied on the raters’ teaching expertise. His research points to an understanding of consistency that is based in a teaching community’s consensus of what counts as successful writing rather than individual readers’ ability to come to agreement via a rubric. While Smith’s work gave rise to new assessment methods, the real significance of his research is that it highlights the need to revise our approach to assessment. Instead of focusing on how to maintain the kind of standardization necessary for training raters to agree (a focus on the technology of writing assessment), we should ask other questions about reliability (questions focused on assessment as inquiry and ethical practice): How can we be fair to students? How can we produce consistent decisions that are fair? What kind of reading environment helps to produce consistently fair scores?

Smith’s research grew out of his understanding of assessment practices and technologies as well as his commitment to posing questions and looking for answers (“Assessing”; “Importance” 314-16). By approaching assessment as research—as an opportunity to learn more about programs, pedagogy, students, and writing by asking questions—we can make valuable contributions to the field and contribute to the formation of composition as an academic discipline. The assessment research that Smith conducted involved articulating the purposes and content of each of the different courses the program offered; designing and testing prompts to determine if they appropriately sorted students for the courses; examining how variables such as location, proctoring, and time affected student performance; examining methods of reading and scoring of essays; determining factors that influenced readers such as the most recent course taught; and tracking students through various courses to determine the effectiveness of the courses (“Assessing”; “Importance”). As mentioned above, Smith’s research depended on his nuanced understanding of concepts such as reliability and validity as well as theories about writing, but he also relied on a variety of research methods, including surveys, interviewing, think-aloud protocols, and participant observation. He conducted double-blind studies so that student writing was read and sorted various ways to test how reliable the sorting mechanisms were (“Assessing”). As with other types of research,

much of what Smith learned along the way created knowledge useful not only to him but potentially to the field. He discovered that his readers actually placed students, instead of scoring essays; that is, instead of focusing on the text as an artifact to be assigned a number, the readers used their knowledge of the course and the students in the program to determine which class would best serve the students' needs. This process is very different than merely assigning a number from a rubric to an essay. Smith also concluded that adequacy of placement could only be determined through multiple data sources, such as the number of students moved, student grades, teacher impressions of students' placement, exit exams or post tests (when used in the program), and student satisfaction surveys gathered after the end of semester. All of this information is useful to WPAs not only in planning a placement test but also for program and faculty development.

Because Smith did not just administer the placement test but made the evaluation of it a central and ongoing research concern, he was able to refine the system over time, making valuable contributions to the field's understanding of placement and to the knowledge about assessment and research in general. Smith's work exemplifies the potential in assessment as research. Other composition professionals, such as Richard Haswell, have also approached assessment this way, but many WPAs and writing programs have yet to see the potential in assessment research.

Viewing assessment as inquiry as Smith and Haswell have done also changes the options available in constructing assessments to suit our particular contexts. For example, if we see assessment as research, then it is no longer appropriate to determine the adequacy or validity of a placement program merely by looking at whether or not we have the necessary number of students in developmental courses versus first-year composition or whether the raters have agreed with each other enough. Rather, we need to explore what actual benefits students receive from their placement: Do students who are placed into basic or remedial non-credit-bearing writing courses benefit from this instruction? Do they remain in school? What kinds of grades do they receive when they enter intermediate and advanced courses? These sorts of questions not only refigure assessment as research, they also contribute to the validation of an assessment's results. In this paradigm, validation inquiry is about building an argument using multiple sources of evidence to demonstrate that the results of the assessment improve teaching and learning. Furthermore, data gathering and evaluation are integral parts of the entire writing program, not something calculated at the end of the scoring sessions to justify the placements.

Reconfiguring assessment as research also has the potential to change the power dynamics inherent in assessment. Often, assessment reinforces a hierarchical structure where assessment is done *to us* (teachers, students or programs) by *them* (experts, upper administration or governing boards) contributing to an overall sense of surveillance or dominance. By focusing on assessment as inquiry, these relationships of power can be substantially changed, in several ways:

More people are qualified to do assessment: While most writing programs have no one who is an assessment expert, most do have people with experience and expertise in empirical research.

Multiple methods of assessment are available: No longer does the holistically scored timed essay, a portfolio, or another student-generated product that gets scored become the primary data gathering tool. Multiple methods of data gathering are available depending on the questions asked.

Standardized testing becomes limited: Outside assessments purchased from testing companies have limited value because programs must document how well they can answer questions they wrote that are specific to their program.

Everyone participates in the assessment's design and implementation: Assessment cannot be left in the hands of a few people in a department; it takes an entire community to know what needs to be asked and to provide answers to those questions.

The technological aspects of writing assessment are de-emphasized: Posing questions is the primary task because the questions determine what methods are needed. Technical aspects of an assessment become secondary to the asking and answering of questions a community deems important. Technical issues such as reliability are still included but are not the driving force of the assessment.

Issues of power and authority, while present in any program, are especially significant in writing programs where most courses are taught by contingent labor (i.e., adjuncts, lectures, graduate students). While it is impossible to alleviate these power differences, assessment can help subvert or undermine them. For example, at the University of Louisville, the writing program conducted a self-assessment as part of an accreditation review that included all instructors from tenure-track faculty to adjuncts and graduate teaching assistants.² All instructors created a course portfolio for every class they taught. The portfolios included the syllabus and all course handouts, samples of student work, and a short

introductory reflection about the class. All materials were identified only by the course and section; instructor names were not on any materials. These portfolios were read by a small representative group of instructors who taught the course, whether they were tenure-track faculty, adjuncts, or graduate students. The groups also determined their own process for reading the portfolios within very general guidelines established by the department's composition committee. Each group documented its reading in a report that included a general description of the course, its goals, and the student work. The reading reports identified strengths and weaknesses from the portfolios as a group and also described the method the readers used to accomplish the task. From the individual course reports, the WPA created the overall program assessment document identifying the overall strengths and weaknesses of the program. In this system, all instructors participated by collecting materials, and their voices were included through the course documents and self-reflections; many others were also included in the reading groups (which consisted of three or four instructors, depending on the number of sections of the course offered), although the WPA compiled the reports and synthesized the data. The process extended over almost a full academic year and was messy at times; however, the results not only satisfied the accrediting agency but also provided information useful for the program that led to changes, such as criteria for those teaching business writing.

Ultimately, reimagining assessment as Smith, Haswell and others have done changes the focus of assessment from the *products* of scores and consistency of scoring to the *process* of assessing writing and writing programs. In this approach, stakeholders are forced to confront questions about what needs to be addressed in a program, what information is crucial to improving teaching and learning, and why assessment data matters by using appropriate research methods to respond to the questions and document findings as in any type of research project. Assessment, then, is no longer a machine through which to run our students and their writing, but a fundamental disciplinary activity; it becomes a process for inquiry and reflection, a way of knowing who we are, how we're doing, and where we want to go.

CONSEQUENCES OF ASSESSMENT: PERIL OR PROMISE

Most teachers and WPAs have stories that attest to the ways that assessment drives what we do in our classrooms and how it complicates our purposes and goals for the teaching of writing. Being told by administrators or accrediting agencies that a program must assess writing through psychometric, nonsensitive instruments or scoring guides can put WPAs in an ethical and professional dilemma. Such standard forms of assess-

ment will not measure the success of the program—because they will not treat students as writers—and therefore the results, which may be widely shared with administrators, colleagues in other departments, and even the public, could implicate students and teachers in unfair, even dangerous ways. However, it could be politically risky for the WPA to refuse such an assessment—such a refusal wouldn't even be an option in many circumstances.

But while WPAs may not be able to refuse mandated assessments, they do have an obligation to respond to them appropriately. The problem is that in order for WPAs to establish credibility for their programs, their assessment techniques must be such that they produce convincing, useful results while also treating students, teachers, and the practice of writing with sensitivity and respect. This problem becomes more complex when an administrator or outside agency is demanding accountability of the writing program or data to prove that quality teaching and learning are happening within a program.

Understanding writing assessment as a means of research or inquiry is an effective way to reimagine writing assessment because it puts assessment to work for us, rather than us to work for the sake of assessment. It also forces us to ask different kinds of questions in creating assessments, such as these: What work is the assessment doing to improve teaching and learning? What are the consequences of the assessment for individuals (teachers, students, the WPA) and the program at large? And, importantly, what does the way we assess our programs reveal about what we value about literacy and education?

After all, writing assessment, which is undoubtedly political—ideologically informed and often tied to public policy—is a means of contributing to university-wide or public discourse about what is happening with our composition programs. As such, it is a form of social action—whether or not we acknowledge it as such—because its consequences have far-reaching effects on teachers, students, classrooms, and writing programs. For example, the data generated during a program assessment can have positive or negative consequences for real people, individually and collectively. If the methods of assessing the program are invasive and threatening to students or teachers (for example, a student's graduation status could be affected by his or her participation in the assessment; or an instructor's job security may hang in the balance based on the results of the assessment), it may not support good teaching and learning (Hillocks; Madaus 96-97). Moreover, if the assessment data are gathered in a way that does not accurately reflect what happens in the program (pre and post-testing, for example, in a program in which many instructors use a process approach to writing and portfolio grading), the results could do more than disrupt the teaching and learning

going on in the program; the assessment might also actually give results about student writers and teachers that are inaccurate and damaging, especially if those results become the impetus for sweeping changes. Assessment results could even lead a WPA to fear for his or her job or tenure status.

And the damage to a program doesn't end with what happens to students, teachers, and the WPA(s): the results of an assessment are often discussed across campus, by colleagues in other disciplines, by administrators, and often with accrediting agencies. The failure of a WPA to control the design and direction of an assessment—whether it be a program evaluation, a placement exam, or an exit assessment—could damage his or her credibility and that of the program.

By seeing assessment as an opportunity to research a program, however, much of the fear of assessment and the problematic consequences can be eliminated. In the example cited above from the University of Louisville, the course portfolio assessment design reflected the values of the program: that all teachers—not just the WPA and one or two others—have important contributions to make to the program; that the successes of a course can best be documented by diverse artifacts from the course; that what is actually happening in classrooms can best demonstrate what a program is doing well and point out what might need improvement; and that the conversation about what is particularly successful or problematic about a course must happen not between two or three people, but across a program or department.

On a local level, writing assessment can have a positive impact on teaching and learning if we view it as an opportunity to inquire into our programs and then proactively respond to what we discover. At the school or program level, we can view assessment as an occasion for teachers to discover and share with each other what they value in writing instruction, giving both teachers and curriculum an opportunity to change. We not only should construct writing assessments by talking as a program about what we want to know, but also by talking about how the design of the writing assessment will support or positively challenge the writing curriculum and the pedagogies of the teachers in the program. If we reconceive of writing assessment as an occasion for conversation, we can view assessment as an opportunity to know more about what the various stakeholders—students, teachers, administrators—think about writing instruction and assessment, enlarging the dialogue and expanding our understanding of how writing assessments impact the people who teach and learn in our programs. For example, by analyzing how writing is assessed, we can view assessment as a means of ensuring that our programs and schools respect diversity (Mountford) and encourage the concept of multiple literacies by asking questions

such as these: How do the designs of our assessment systems value or not value diverse literacies? How can we as a writing program ensure that a broad range of literacy practices is respected and incorporated into the assessment system?

But these benefits to programs and universities extend beyond academia. Once we view writing assessment both as research and as an occasion to be proactive within our institutions and programs, it becomes an opportunity for change on our own terms rather than a means for control from outsiders. However, making writing assessment a priority issue in our work as administrators isn't enough: we must effectively communicate our findings to all stakeholders, and even use the data we generate to enter into more public discussions about literacy education. Peter Mortensen's "Going Public" argues that believing "our work—our teaching, researching, theorizing—can clarify and even improve the prospectus of literacy in democratic culture" means that "we must then acknowledge our obligation to air that work in the most expansive, inclusive forums possible" (182). That is, if WPAs and teachers of writing want our work to connect to the culture outside of academia, then we must engage publicly the conversations about literacy and education that circulate through the media, local school board meetings, civic events, and casual discussion. Information we gather about our programs, teachers, and students—about the rich work that goes on in our classrooms—can show others how what we do is relevant and important. And it can, potentially, help to change the tenor of the conversations already going on about the work we do in teaching college-level writing. As Mortensen argues,

Journalists, essayists, polemicists, policy analysts, and others are writing about the same literacies we study. Indeed, sometimes they critique our teaching and research practices as part of their efforts to locate their observations and arguments in culturally and historically familiar territory. [. . .] [F]or in failing to [learn from these public discussions on literacy], we consign ourselves to mere spectatorship in national, regional—and most importantly—local struggles over what counts as literacy and who should have opportunities to attain it. (182-83)

While "Going Public" focuses mainly on literacy research, what Mortensen writes applies also to the data we generate through writing assessment and the decision-making we go through in practicing writing assessment ethically. As a method of inquiry, assessment can give us important knowledge about literacies as they are practiced within our institutions. As research reports, the results of our assessments

are presented to administrators, faculty from other disciplines, and the public at large as statements about what we value and why. And as public documents, they have the potential to begin important, localized discussions in communities about the roles of literacy in culture and in higher education. But re-envisioning writing assessment as an opportunity for public discussion means thinking about writing assessment locally, contextually, and as a public articulation of values.

By taking control of writing assessment in their own programs and institutions, WPAs can meet demands for assessment head-on, with data that is useful to compositionists and to those who demand assessments, and with knowledge that allows them to talk publicly about the practices and theories valued by the composition community. By staying informed of public policy initiatives and public opinion on writing assessment, and by joining organizations such as the NCTE and CCCC committees on assessment, WPAs can keep abreast of new assessment innovations (theories as well as practices) and how they fit with or undercut compositionists' goals for writing instruction.

WRITING ASSESSMENT AND THE DISCIPLINE OF COMPOSITION

The connection between the assessment of writing and the field of composition studies has been demonstrated by scholars such as Robert Connors, James Berlin, John Brereton, and others. Historically, assessment has had a long-lasting effect on writing, pedagogy, and the development of composition studies by identifying writing (especially first-year composition) as a gatekeeping activity, by positioning college composition as a remedial course, by situating assessment as an administrative apparatus, by establishing writing as a decontextualized activity instead of a communicative act, by promoting writing as a private exchange instead of a public activity, by defining "good" writing as that which is devoid of mechanical and grammatical errors, and by cultivating the notion that just about anyone can teach and evaluate writing. Furthermore, writing assessment has shaped the academy since it has been used to deny access to particular groups of people, such as minorities, women, and working class students (Agnew and McLaughlin; Moss; Sternglass).

Despite writing assessment's problematic place in the history of composition and the academy, it offers us opportunities, opportunities to make knowledge about writing, writers, and the teaching of writing. Changing our perceptions about and uses of assessment requires that we resist the idea that assessment is something done to us *by them* or done by us *for them*—and *them* could be accrediting associations, state legislatures, or institutional administrators. Assessment is about asking questions and looking for answers. It is research and inquiry. It demands

theoretical and practical knowledge. It requires the systematic collection of data. It needs rigorous analyses through a variety of lenses, from various perspectives. And it needs to be grounded in an informed notion of validity, not a narrow interpretation of reliability. To realize this new definition of writing assessment as action research that is central to the administration of a writing program, compositionists need to

- read and review assessment literature critically, including work outside of composition studies, so that we understand the complexity of concepts such as reliability and validity;
- understand qualitative and quantitative aspects of research;
- demand evidence for claims of reliability, validity, and “success” instead of relying on self-reports or our own “intuitive” sense;
- include formal instruction in assessment as part of graduate work;
- provide opportunities for WPAs and others who may not have had previous preparation in assessment to participate in theoretically informed professional development; and
- encourage and support ongoing research into our assessments by those not directly invested in the assessment or its results.

While some of these suggestions are beginning to happen—the Council of Writing Program Administrators attached a one-day assessment institute to the 2002 annual conference—there still needs to be a sustained effort by practicing WPAs and their composition colleagues to develop theoretical as well as practical expertise in writing assessment. Conceptualizing assessment as research rather than an administrative, bureaucratic activity is one way to encourage WPAs to engage seriously in writing assessment as a disciplinary activity. After all, research is not only what defines a discipline but also what is often most privileged in academic institutions.

Redefining assessment as research and re-envisioning the role that assessment plays in composition studies can also address some of the problems with the field that Richard Haswell mentioned on the WPA-L listserv. Haswell explained that “a large part of the comp/rhet world in the last twenty years has systematically resisted the development of a research agenda and body of research that the rest of the world can recognize as befitting a serious disciplinary study.” Haswell continued by way of explanation: “[. . .] imagine the body of professional medical journals consisting of 95% editorial opinion and 5% research studies—and contrafactually imagine the body of rhet/comp journals consisting of

5% opinion and 95% research studies.” Haswell wasn’t referring specifically to writing assessment, but his comments seem particularly relevant to writing assessment, where we often depend more on opinions than empirical evidence in making decisions or drawing conclusions.

Haswell went on, referring to Robert Scholes’s point that “many college and university teachers of language and literature would be embarrassed if they had to explain to their legislature or trustees exactly what methods [. . .] they are advocating in class.” Again, Haswell isn’t referring specifically to assessment, but there is a connection because writing assessment offers a way to investigate and explain what we do in our classrooms. Writing assessment is a means of building our theories and pedagogies around research-based work. However, for it do this, we must imagine writing assessment as systematic and rigorous inquiry into what we do—inquiry that has real-world consequences for stakeholders both within and beyond the academy— instead of just another bureaucratic nuisance that gets in the way of our “real” work. Assessment is our real work.

Once enough compositionists begin to reimagine the possibilities for assessment, perhaps frantic posts like the ones which open this article will disappear and be replaced with a disciplinary conversation about writing assessment theory and practice, one focused on posing and answering questions about writers, writing, and pedagogy. In fact, there may be some movement in this very direction with a recent WPA-L discussion on a portfolio assessment moving beyond responding to someone’s request for practical assessment materials to more substantial issues about the theories informing the suggested practices. For this type of change to happen on a larger scale, we need to make a sustained effort—as individuals and as a field—to not only value assessment research but to reposition assessment as research.

NOTES

¹ This and other posts are from the WPA-L archives available at <<http://lists.asu.edu/archives/wpa-l.html>>. We purposely are not identifying them by author, date, or subject heading because we are not trying to criticize individuals but rather illustrate the type of responses that are posted to the list. The unidentified posts were selected from a long list of posts that resulted from several searches of the archives conducted 11/12/99 and 5/31/00 with search terms that included *placement*, *assessment*, and *test*.

² Two of the authors participated in this program as members of the Composition Committee at the University of Louisville from 1995 to 1996. The overview presented here draws on their experiences as well as the committee’s minutes, reports, and correspondence.

WORKS CITED

- Agnew, Eleanor, and Margaret McLaughlin. "Those Crazy Gates and How They Swing: Tracking the System that Tracks African-American Students." *Mainstreaming Basic Writers: Politics and Pedagogies of Access*. Ed. Gerri McNenny and Sallyanne Fitzgerald. Mahwah, NJ: Erlbaum, 2001. 85-100.
- Berlin, James. "Rhetoric and Ideology in the Writing Class." *College English* 50.5 (1988): 477-94.
- Brereton, John C., ed. *The Origins of Composition Studies in the American College, 1875-1925: A Documentary History*. Pittsburgh: U of Pittsburgh P, 1995.
- Cherry, Roger D., and Paul R. Meyer. "Reliability Issues in Holistic Assessment." Williamson and Huot, 109-41.
- Connors, Robert J. *Composition-Rhetoric: Backgrounds, Theory and Pedagogy*. Pittsburgh: U of Pittsburgh P, 1997.
- Haswell, Richard, ed. *Beyond Outcomes: Assessment and Instruction within a University Writing Program*. Westport, CT: Ablex Publishing, 2001.
- . "RE: doin' more than thinkin' 'bout quittin.'" WPA-L. <http://lists.asu.edu/archives/wpa-l.html> (8 Mar. 2000).
- . "Multiple Inquiry in the Validation of Writing Tests." *Assessing Writing* 5.1 (1998): 89-110.
- Hillocks, George, Jr. *The Testing Trap: How State Writing Assessments Control Learning*. New York: Teachers College Press, 2002.
- Huot, Brian. "A Survey of College and University Writing Placement Practices." *WPA: Writing Program Administration* 17.3 (1994): 49-67.
- . "Toward a New Theory of Writing Assessment." *College Composition and Communication* 47.4 (1996): 549-66.
- Madaus, George F. "The Influence of Testing on the Curriculum." *Critical Issues in Curriculum: Eighty-Seventh Yearbook of the National Society for the Study of Education*. Ed. Laurel Tanner. Chicago: U of Chicago P, 1988. 83-121.
- Mortenson, Peter. "Going Public." *College Composition and Communication* 50.2 (1998): 182-205.
- Moss, Pamela. "Can There Be Validity Without Reliability?" *Educational Researcher* 23.2 (1994) 5-12.
- Mountford, Roxanne. "Let Them Experiment: Accommodating Diverse Discourse Practices in Large-Scale Writing Assessment." *Evaluating Writing: The Role of Teachers' Knowledge about Text, Learning, and Culture*. Ed. Charles R. Cooper and Lee Odell. Urbana, IL: NCTE, 1999. 366-96.
- Murphy, Sandra, Sybil Carlson, and Paul Rooney, with the CCCC Committee on Assessment. Report to the CCCC Executive Committee: Survey of Post-secondary Writing Assessment Practices. Unpublished manuscript, 1993.

- Smith, William. L. "Assessing the Reliability and Adequacy of Placement Using Holistic Scoring of Essays as a College Composition Placement Test." Williamson and Huot, 142-205.
- . "The Importance of Teacher Knowledge in College Composition Placement Testing." *Reading Empirical Research Studies: The Rhetoric of Research*. Ed. J. R. Hayes, et al. Hillsdale, NJ: Erlbaum, 1992. 289-316.
- Starch, Daniel, and Edward C. Elliott. "Reliability of the Grading of High-School Work in English." *School Review* 20 (1912): 442-57.
- Sternglass, Marilyn S. *Time to Know Them: A Longitudinal Study of Writing and Learning at the College Level*. Mahwah, NJ: Erlbaum, 1997.
- White, Edward M. "Holistic Scoring: Past Triumphs, Future Challenges." Williamson and Huot, 79-108.
- Williamson, Michael M. "An Introduction to Holistic Scoring: The Social, Historical, and Theoretical Context for Writing Assessment." Williamson and Huot, 1-44.
- Williamson, Michael M., and Brian Huot, eds. *Validating Holistic Scoring for Writing Assessment: Theoretical and Empirical Foundations*. Cresskill, NJ: Hampton, 1993.

