Cindy Moore, Peggy O'Neill, and Brian Huot

Creating a Culture of Assessment in Writing Programs and Beyond

As writing-program administrators and faculty are being called upon more frequently to help design and facilitate large-scale assessments, it becomes increasingly important for us to see assessment as integral to our work as academics. This article provides a framework, based on current historical, theoretical, and rhetorical knowledge, to help writing specialists understand how to embrace assessment as a powerful mechanism for improved teaching and learning at their institutions.

> Can we have not simply writing-across-the-curriculum but also writing-assessment-across-the-curriculum? If the Department of Writing could model this for the rest of us, that would be great.

This question, asked in an email from a dean to a first-year-composition director, illustrates just how central writing and writing assessment have become to discussions about institutional-assessment goals and practices that are occurring at colleges and universities across the country (and around the globe). That a composition director would be included in such discussions is not surprising, given that more and more program-level administrators are being asked to provide information for campus-wide self-studies and accreditation reviews. This is especially true for writing program administrators (WPAs), who typically participate in cross-curricular general education efforts by way of coordinating introductory composition courses and supporting the instructors who teach them. What seems most interesting about the email query is the implied message about the potential role of the composition director in the broad-based assessment this dean is beginning to imagine. The dean seems not to be cajoling the WPA to fall in line with an assessment regimen that has already been envisioned (as higher-ed lore might encourage us to expect), but rather inviting her to take the lead in implementing what promises to become a significant campus-wide initiative.

Though potentially intimidating, the proposition embedded within this email is an important one, indeed. As research shows, writing assessments do much more than simply allow administrators to demonstrate that their institutions, departments, and programs are successful; they have the power to influence curriculum and pedagogy, to categorize teachers and writers, and, ultimately, to define "good writing" (e.g., Hillocks; O'Neill, Murphy, Huot, and Williamson). In fact, specific writing assessments, especially those perceived to have high stakes for students and teachers, function as what Deborah Brandt calls "literacy sponsors" because they encourage and support the development of certain types of writing and writing abilities over others. Thus, a department-level administrator who embraces assessment—especially the kind that extends beyond the boundaries of her specific program—is in a position not only to help set the agenda for campus-wide assessment but also to affect, even "transform," teaching and learning across the university community (Bean, Carrithers, and Earenfight).

Sadly, though the particular WPA in this scenario understood the positive aspects of involvement and was willing to help her dean think through how a college-wide writing initiative might be used, simultaneously, to evaluate learning across campus, many WPAs—in fact, writing faculty in general—are not inclined to assume an active role in assessment, even when deans or department chairs show confidence in their doing so. Undoubtedly, there are many complex reasons for disinterest. One reason is that, all too often, assessment initiatives *are* imposed from the top down, rather than invited or encouraged. When assessment is imposed (or perceived to be imposed), its relevance may not be apparent. This is especially the case when people outside of a program (a dean, provost, or institutional effectiveness office) dictate the parameters of the assessment—for example, the purpose(s), guiding question(s), and methods for data collection, analysis, reporting, and eventual use. An assessment that is not framed by questions important to the program administrators and faculty gathering the data and whose results, therefore, may not seem meaningful, likely will be viewed as pointless busywork, completed simply to help others fill in the blanks of reports that, if they are read at all by decision makers, will never be acted upon. Worse yet, if the purposes, audiences, and implications of externally initiated assessments are not made clear, program administrators and faculty may assume that results will be used in undesirable ways—for example, to exclude students, monitor faculty, and control curriculum—as has too often been the case at higher-ed institutions (e.g., Greenberg; Gleason; Agnew and McLaughlin).

Even when writing program administrators and faculty are invited or encouraged to design and facilitate assessments locally, they often hesitate because of unfamiliarity with possible approaches to large-scale assessment as well as the key concepts, documented history, and recorded beliefs associated with various approaches. This unfamiliarity is reflected in multiple ways—through urgent postings on disciplinary listservs asking for the "best way" to assess student work for course placement or curricular review; through assessment workshops in which program directors clamor for practical advice on how to confront administrative assessment mandates; and through the now-ubiquitous articles in the *Chronicle of Higher Education* and elsewhere about tensions between various constituencies (e.g., faculty, university administrators, legislators) over the presumed "validity" or "reliability" of particular assessment methods. Unfortunately, even the most informed responses to public pleas for assistance or reassurance do not magically solve the crises because, as assessment scholars argue, good assessments are those that are designed locally, for the needs of specific institutions, faculty, and students. As a result, well-intentioned suggestions often lead to poor assessments, which, in a circular way, can reinforce bad feelings about assessment generally.

Of course, many administrators and faculty do recognize the importance of assessment and know the assessment "basics"; they just have trouble translating their understanding into assessments that will be helpful to their programs and students. For some, the difficulty connecting theory with practice stems from resistance among colleagues who are not interested in assessment and do not want to participate in assessment efforts. For others, unawareness about the potential complexity of forging theory-practice connections inspires challenges even when there is no obvious resistance. Still others are faced with a lack of institutional support (e.g., money, release time, clerical assistance) for substantive, meaningful assessments.

In the discussion that follows, we seek to address the needs of our colleagues who are wary of large-scale writing assessments and those who acknowledge the importance of such assessments but have trouble making them work for their particular programs. Our aim is to provide an overview of how readers can participate actively in creating a culture of assessment, thus positioning themselves within their institutions to enhance teaching and learning across campus. Because assessment is central to teaching and learning in general (Johnston; Moss, "Shifting"; Shepard) and to writing in particular (Huot, (*Re)Articulating*; White, *Teaching*), and because the stakes are so high for us and our students, WPAs and other writing faculty must find ways to participate, wisely and well, in this powerful activity.

Current writing assessment theory, developed out of discussions and debates that have occurred over the last several decades, requires that administrators and faculty charged with assessment consider certain primary principles, five of which have been articulated by Huot ((*Re)Articulating*). According to Huot, assessment should be site based, locally controlled, context sensitive, rhetorically based, and accessible (i.e., transparent to all interested parties) (105). To these five basics, we add the need for writing assessments to be consistent with contemporary theories about language learning, literacy, and educational measurement as well as the histories out of which these theories evolved. Although these principles may seem simple on the surface, they are often difficult to enact because they require more than a willingness to be involved in designing and facilitating assessments and a familiarity with contemporary assessment methods (the focus of much of the writing-assessment literature). Our key contention is that creating local cultures that support meaningful assessment hinges on an understanding of how, when considered together, relevant historical trends, theoretical tenants, and contextual factors can influence assessment practice in truly transformative ways.

Intersections among History, Theory, and Context

Disciplinary and Assessment History

Although most rhetoric and composition specialists acknowledge the history of their discipline as one marked by various struggles for legitimacy, they may not always situate assessment challenges within that larger history. Sometimes the challenges faced by WPAs and writing faculty as they propose assessments have less to do with the merits of the projects and more to do with lingering negative perceptions of the discipline. Recognizing the possibility of disciplinary bias can help us interpret local challenges, especially in a department without a strong cohort of writing experts. Such bias has been manifest since the early 1900s in discussions about who should teach within and administer composition programs—what their qualifications should be, what level of institutional status and support they should have, and what kind and degree of scholarship should be expected of them. While much has changed since the first composition courses appeared in the nineteenth century, it is important to recognize that the attitudes expressed in the familiar statement "anyone can teach writing" are still alive and well.

Of course, the temptation for WPAs and faculty who find themselves in less-than-ideal institutional situations is to explain local successes and setbacks solely in terms of a fraught disciplinary history, disconnected from the often-powerful influences of local context. Yet, researchers like Tim Peeples caution against such ready interpretations. In his analysis of a narrative written by Wendy Bishop and Gay Lynn Crossley about challenges Bishop experienced as a WPA, Peeples convincingly argues that her challenges were both multiple and "dynamic," reflecting not only larger disciplinary conversations about writing, but multifaceted, ever-evolving conversations about what it means to be an administrator, generally, and within a particular institutional framework. That is, though disciplinary differences and disparities offer ready explanatory evidence, program administrators and faculty benefit from not simply understanding disciplinary history but also from recognizing how this history is usually just one of many factors that can make their work challenging.

In addition to a knowledge of disciplinary history, WPAs and faculty benefit from understanding the history (or, perhaps more appropriately, the *histories*) of college writing assessment—accounts that highlight not only how certain kinds of approaches and methods evolved, but why. For example, historical scholarship, such as that published by Connors, Brereton, and others, shows how the composition entrance exam and the mandatory composition courses that followed from it were originally designed to foster educational and structural changes at both the secondary and postsecondary levels. As part of his goal to change the language of learning from Latin and Greek to English, Harvard president Charles Eliot added an exam in English composition to the battery of entrance exams in 1874. A key component of the formation of Harvard's writing program—considered to be the first modern composition program (Brereton 8)—was the testing of incoming students' performance in written English. These exams, which were designed to encourage the prep schools to teach writing so their students were ready for the entrance exam, were read and evaluated by Harvard faculty. Harvard's move to include written English composition in its entrance exams and curriculum quickly spread to other prestigious Northeastern schools and beyond. By 1900, these exams were commonplace features of higher education in America. Along with the exams and the composition courses came the underclass of writing instructors who, as Connors explains, were deluged with themes to read and grade (171–72). Acknowledging the role of assessment in composition's history, as illustrated through the Harvard program, demonstrates the intimate connection between composition as a university subject and writing assessment.

By the beginning of the twentieth century, the College Entrance Examination Board (CEEB) was established, and writing assessment as a funded, researched, and professional field and industry began. Individual colleges no longer had to create their own entrance exams, nor did the faculty assess them; they could use standardized tests created by CEEB and eventually its spinoff organization, Educational Testing Service.[1] Preparatory schools no longer had to prepare students for an array of different exams, designed by individual institutions, but just one that would be appropriate for a range of colleges. Writing assessment had been outsourced. Since then, CEEB's long association with the testing of writing has continued with the recently revised SAT that includes a scored writing sample. The move to the standardized entrance exam—and, at many schools, standardized placement testing—reflects an increasing desire for efficiency, objectivity, and quantification of achievement that can be traced back to World War I. The military's desire to efficiently sort and assign personnel during WW I contributed significantly to the rise in psychometrics, the science of educational and psychological measurement.[2]

Norbert Elliot, in his book-length treatment of the history of writing assessment, explains that the impetus for the CEEB came not from the colleges themselves but from the secondary schools that were preparing students to pass the individual exams to enter college. In 1895 Wilson Farrand, headmaster of Newark Academy, was elected president of the Schoolmasters Association of New York and Vicinity. Farrand outlined the rationale and structure of an organization based upon five guidelines that would eventually become the CEEB. The first of Farrand's guidelines included a certificate from the secondary institution documenting a student's progress and achievements that would guide the content and extent of the examination necessary for each student; the other four guidelines focused on the examination itself. In 1899 under the direction of a dean at Columbia University, the Association of Colleges and Preparatory Schools of the Middle States and Maryland proposed the establishment of the

CEEB, and Farrand was named the first secretary. Significantly, all but one of Farrand's guidelines were incorporated into the structure of the CEEB. The missing guideline was the certificate from the secondary school that provided contextual information about the academic achievement of each student.

In addition to attributing the founding of the CEEB to Farrand, then, Elliot's version is important for the way it draws attention to the omission of a guideline specifying that high school educators would participate in the process of college admission by certifying what students had learned in order to document academic achievement and direct the kinds of admission testing necessary. This kind of omission is important because it signals that over one hundred years ago, teacher judgments were found suspect, and a test was assumed to be better at helping admissions committees make consequential decisions about students than their secondary teachers. For over a century, teachers have had to struggle to be a part of key decisions made on the basis of the assessment of student writing. This undervaluing of students' performance in high school is bad practice at the very least, since high school performance has over the years remained the best single predictor of success in college[3] (Bard; Baron and Norman; Sacks). Nonetheless, faith in so-called objective measures of writing assessment has continued, so that while the ACT and SAT now include a writing portion, the exam allots only twenty-five minutes for students to write. The test, furthermore, is being marketed as an appropriate placement measure for incoming first-year writers, a fact that some (White, "Misuse") consider harmful not only to students and teachers but also to the communities and economies that support locally controlled placement programs.

The story of the CEEB is useful for composition faculty involved in assessment because it reminds us that writing assessment has a long history of important debate and discussion that illustrates its power to shape curricula, influence institutions and programs, and distribute rewards and sanctions. Learning to look for voices or stories of those not represented in the mainstream narrative of an assessment can provide useful insights for those who conduct assessments, just as looking beyond college composition and rhetoric's published literature is important for those who study writing assessment. Understanding that writing assessment has always been seen as a means to effect change reminds us to interrogate local calls for assessment and the potential wide-ranging consequences to students, faculty, programs, institutions, and disciplines that responses to these calls may have. As the history of Harvard's composition entrance exam demonstrates, sometimes assessments are used as a way for administrators to make curricular changes, or, as Agnew and

McLaughlin argue, they may serve as a mechanism for rewarding the "right" kind of students, reinforcing, for example, white privilege. In fact, debates that appear to be only local often not only reflect wider contemporary conversations but may also echo discussions heard over the last hundred years about who should teach and evaluate composition, what should be taught in composition courses, what "good" writing is, and how writing should be evaluated. Knowledge of the history of writing assessment can also help us understand faculty resistance, confront that resistance, and move beyond it to productive inquiry.

Assessment Theory

In developing thoughtful methods of inquiry—and confronting possible resistance to them—WPAs and faculty charged with assessment can also benefit from a knowledge of ongoing discussions about key assessment constructs, theories, and practices, which appear not only in recent composition and rhetoric literature but also in the psychometric and educational measurement scholarship. An understanding of debates surrounding validity and reliability, in particular, is crucial for facilitating productive collaborations between writing faculty and other entities, such as campus testing offices, institutional research personnel, or representatives from testing companies.

Validity, the critical theoretical concept in measurement theory, is often misrepresented in composition literature. For example, in the 1994 edition of Edward White's still-popular book, *Teaching and Assessing Writing*, validity is explained this way: "Although validity is a complex issue—colleges offer advanced courses in it—one simple concept lies behind the complexity: honesty. Validity in measurement means that you are measuring what you say you are measuring, not something else, and that you have really thought through the importance of your measurement in considerable detail" (10). Although we appreciate the attempt to make validity understandable to those not trained in psychometrics, this definition neglects essential elements of the concept. Ensuring validity requires construction of a sound argument to support the interpretation and use of assessment results with both theoretical and empirical evidence.

According to educational measurement scholars such as Lee Cronbach, validity "must link concepts, evidence, social and personal consequences and values" (4). Samuel Messick, another influential voice, argues that validity uses "integrated evaluative judgment," supported by empirical evidence and theoretical rationales, "to support the adequacy and appropriateness of inferences and actions based on test scores and modes of assessment" (5). In other words, validation arguments are rhetorical constructs that draw from all the available means of support. Validation is also ongoing and should include a feedback loop: "As validation proceeds, and new evidence about the meaning of the test's scores becomes available, revisions may be needed in the test, in the conceptual framework that shapes it, and even in the construct underlying the test" (AERA 9). The validation process starts with explicit statements about the proposed interpretation of the test, including the concepts or constructs being sampled, the conceptual framework of the test, the knowledge, skills, abilities, processes or characteristics being assessed as well as how the construct is to be distinguished from other constructs and how it relates to other variables (9). The conceptual framework also takes into consideration the use of the test results. *Standards for Educational and Psychological Testing*, considered to be the professional and academic code of conduct for work in assessment, makes clear that evidence used in validity arguments should address the content, response process, internal structure, relationship to external variables, and consequences of the assessment and its results.

Although some composition scholarship (e.g., Broad; Huot, *(Re)Articulating*; and Haswell, *Beyond*, "Multiple") does address validity in more nuanced ways, many composition scholars still resist the complexity and power inherent in psychometric theory and its key constructs. For example, Patricia Lynne argues that composition and rhetoric as a field should reject educational measurement concepts of validity and reliability, which she claims are inexorably linked to positivism, and develop our own terms that are compatible with social constructionist theory. In her argument, Lynne contends that educational measurement, with its desire for objective measurements, fails to support assessments that are compatible with composition studies. However, contemporary psychometric theory—as represented by the *Standards* as well as by measurement scholars such as Cronbach and Messick—is compatible with social constructionism. Because some people and agencies (including policymakers, testing corporations, and school districts) do not adequately or accurately apply the theory does not mean the theory is at fault, nor does changing the terms and developing our own guarantee that writing assessment practices will be any better. Furthermore, the roots of contemporary composition grew in part from educational research, including assessment. Unlike Lynne, we argue that compositionists need to become informed assessment practitioners who have a better understanding of the professional standards and guidelines for educational assessment.

Current psychometric theory, especially validity theory, is critical for those of us who need to develop useful, ethical assessments and to interpret the results responsibly and accurately. Besides, teaching and assessing writing are not the exclusive domain of college composition professionals; educational researchers have made and continue to make critical theoretical and pedagogical contributions to composition studies, and we need to value their contributions. College composition students bring with them all of the experiences with literacy learning and assessment that they have accumulated prior to attending college. If anything, we need to engage in more conversation with the PreK–12 community—including those who specialize in assessment—rather than less. There is, after all, a nationwide movement to make the transition from high school to college more seamless as evidenced by programs such as the American Diploma Project (Achieve, Inc.). Refusing to engage in theoretical and practical validity scholarship will make college composition vulnerable to the very kind of reductive, oversimplified approach to assessment that Lynne critiques. With the Spelling Commission's Report calling for accountability of higher education institutions through assessment, we all need to be more astute about assessment theory and practice. Simply put, rejecting psychometrics is not a viable option—practically or theoretically.

We agree with Pamela Moss, an educational measurement scholar, who suggests "that all of us who work within the field of educational research"—which, we argue, includes all of us who develop and conduct writing assessments regardless of our disciplinary training or institutional locations—"have a role that we can play in the dialogue on validity theory and . . . a responsibility to participate in the dialogue within our own research communities in critical and generative ways" ("Joining" 92). Moss explains that validity theory is "socially constructed" (92), reflecting changing and often conflicting points of view. As educators charged with assessment, we need to understand validity theory not only as it has developed within the measurement community but also as part of the social sciences more generally. According to Moss, an understanding of the complex—and often contentious—development of current validity theory can open up the assessment process to alternative practices and interpretations. These alternatives, Moss contends, can broaden the assessment possibilities and can include different disciplinary research traditions, such as ethnography, hermeneutics, and postmodernism.

For those of us in composition and rhetoric responsible for writing assessment, Moss's position means that we should not simply adopt a test or assessment plan that someone else has developed but rather see the need for assessment as a research opportunity (see also Huot, *(Re)Articulating*; O'Neill, Schendel, and Huot). Working in writing assessment requires some familiarity with validity and educational research to ensure ethical, thoughtful, appropriate assessments. It also helps in communicating with other interested parties, such as the university testing office, advising personnel, and administrative staff about procedures, rationales, and results. We can't afford to leave the debates about validity to others because chances are good that they won't share our knowledge, theoretical frameworks, and values.

Although validity is the central concept in measurement theory, reliability is another critical construct. Often considered the twin to validity, reliability is in fact a much more restricted concept that refers to test scoring, which enables the quantification, evaluation, and interpretation of the behavior or work samples (AERA 25) and is critical to the generalizability of the test scores. In determining reliability, test users need to identify potential sources of measurement error, which are divided between those that originate in the examinee and those that are external to the examinee (26). Because measurement errors are random and unpredictable, "they cannot be removed from observed scores"; however, they can be summarized and reported in various ways (27). In addressing reliability, we need to think about the generalizability of the results and possible sources of errors that may influence students' scores (e.g., time of day, administration procedure, or format of the exam).

In writing assessment, reliability has received a lot of attention, although it is usually confined to interrater reliability, the rate at which independent raters agree with each other when scoring a writing sample. Cherry and Meyer explain that many different facets of reliability are at issue in the scoring of writing samples, including how interrater reliability is calculated, but that in most of the composition literature, there is little attention to it beyond reporting interrater reliability co-efficients. Considering reliability only in reference to the raters' agreement, they argue, narrows the term too much because it doesn't address aspects of the concept that are important in testing. While reliability may be an important issue in testing, then, it means much more than the agreement rate of two or more independent readers. As assessors of writing, we need to realize this, so we can begin to address reliability in more sophisticated ways. Too often, testing officers and representatives of testing companies use reliability as a way to stymie more thoughtful and useful writing assessments and promote standardized exams.

In discussing alternatives to traditional emphases on reliability, Moss concludes that "although growing attention to the consequences of assessment use in validity research provides theoretical support for the move toward less standardized assessment, continued reliance on reliability, defined as quantification of consistency among independent observations, requires a significant level of standardization" ("Can" 6). However, these less standardized forms of assessment are often preferable "because certain intellectual activities" cannot be documented through standardized assessments (6). Writing, we argue, is one of those intellectual activities that cannot be adequately documented through standardized tests. Standardization may make it easier to address reliability, but for writing assessment it can also compromise validity too much. Moss suggests that we look beyond psychometric theories and practices in cases where acceptable reliability rates are difficult or impossible to achieve. Moss reminds readers that reliability and objectivity are no guarantors of truth and that they can, in fact, work against "critical dialogue" and can lead "to procedures that attempt to exclude, to the extent possible, the values and contextualized knowledge of the reader and that foreclose on dialogue among readers about specific performances being evaluated" ("Can" 9). Approaching assessment as Moss describes makes more sense given the complexity of writing and also is more aligned with the disciplinary values and theories of most composition and rhetoric faculty.

William L. Smith's work on writing placement at the University of Pittsburgh illustrates how a more nuanced understanding of reliability can provide insight for developing more effective writing assessments. Smith conducted a variety of studies on his placement program to determine if the testing conditions, the structure of the topic, and other variables influenced students' performance on the placement essay. He also tracked raters' agreement—and disagreement—not just between one another but also with themselves. After several years of study, Smith developed a placement process that depended on the teaching expertise of the rater ("Assessing"; "Importance"). His ability to understand reliability and validity as complex, theoretical concepts—within his disciplinary and institutional contexts—allowed him to move beyond the narrow concept of interrater reliability and develop an effective placement method theoretically consistent with language learning and supported with empirical evidence (O'Neill). Smith's work also illustrates that writing assessment has the ability to influence and change the way reliability can be conceived. His expert reader system produced an assessment that delivered more accurate and appropriate decisions (validity) while at the same time increasing the interrater reliability of independent readers.

More recently, Jay Parkes proposed that reliability be seen as an argument, much as validity is conceived. Instead of emphasizing methods for attaining acceptable reliability rates, Parkes focuses on developing methods that support the values—such as accuracy, dependability, stability, and consistency—that reliability represents. This shift in focus, according to Parkes, de-emphasizes the methods, which are most often associated with quantification and statistics, that typically dominate reliability discussions. Parkes argues that methods for determining reliability should "serve as evidence of broader social and scientific values that are critically important in assessment" (2). In Parkes's approach to reliability, a "coefficient is a piece of evidence that operationalizes the values of accuracy, dependability, stability, consistency, or precision. In practice and in rhetoric, however, the methodologies for evidence of reliability are often conflated with the social and scientific values of reliability" (2). If methods don't currently exist for effective assessments, according to Parkes, then new methods need to be developed.

Parkes's approach to reliability fits with the values and theories supported by most composition and rhetoric scholars because it acknowledges the need to understand reliability in terms of specific contexts and purposes. Like the approach to validity endorsed by scholars such as Moss ("Joining"), Messick, and Cronbach, it demonstrates that psychometric theories are evolving and compatible with the social constructivist approach to literacy favored by composition studies. However, we can only use the theory and concepts from the measurement literature to create richer, more nuanced understandings of reliability and validity that support richer and more nuanced understandings of what it means to write by engaging thoughtfully with the psychometric community.

Assessment Contexts

Because we are used to thinking about contexts as scholars and teachers, it is not difficult for us to consider the possible impact of, for example, student experience, faculty attitudes, and institutional priorities on assessment design and implementation. What may be new to readers is the range of ways that such impact can be observed or felt—and the reasons behind it. Approaching assessment rhetorically, as we might a published text or classroom lesson, can help facilitate both an appreciation of relevant contextual factors and an understanding of how they might be addressed, in theoretically consistent ways, to ensure useful assessments.

While assessment-related texts, such as prompts, criteria, and responses, are subject to rhetorical analysis, the assessment itself also needs to be considered within a rhetorical framework (Huot, *(Re)Articulating*). Those who experience success with assessment understand it as a rhetorical act, requiring articulation of exigency, purpose, and audience. From a rhetorical standpoint, it becomes important to ask questions such as "What is motivating assessment at a particular moment?" and "What is the ultimate purpose or purposes, in terms of teaching and learning?" For example, there may be a pressing assessment need, such as accreditation review. Or there may be less formal reasons for assessment, such as curiosity on the department's part to discover the effect of first-year composition on students' attitudes toward writing, or a practical desire to determine if the extra resources devoted to the basic writing courses are a worthwhile use of the department's limited resources. Assessment methods are then chosen, first and foremost, according to their ability to help answer the specific assessment question(s).

At the root of questions that help highlight the purpose of a large-scale writing assessment are more basic questions about who the students are and what they need to succeed in a particular environment—and this is where the contextual terrain begins to appear complicated. Much of the assessment literature (e.g., Broad; Harrington; Haswell, *Beyond*; Smith, "Assessing") has provided ideas about how to gather information to answer questions about students. Recommended data-gathering methods include archival research (e.g., studying memos, reports, syllabi, and assignments for how they characterize students and student-faculty relationships); interviews (with students, faculty, and relevant administrators); and empirical research (such as teaching the courses in which particular groups of students enroll, observing student behavior, and talking to students about their writing). What the literature doesn't always make clear is that although such research can lead to better, more useful assessments, the process can be unexpectedly perplexing. For instance, most administrators and faculty who have posed the simple question "Who are the students?" quickly find that the answers are multiple and depend on the backgrounds, experiences, and status of the people being asked.

Although writing instructors are a great source of information about students, they often perceive students, and define their needs, in different ways. With respect to students labeled "developmental," some faculty will emphasize a need for academic preparedness (e.g., the writing skills needed to succeed in the "regular" first-year course); others will highlight a need for social acclimation (e.g., improved organizational or study habits). The first group of faculty can often be further separated into those faculty who view essential skills as including mostly grammar and punctuation, and those who see overall rhetorical and analytical abilities as especially crucial. Students, too, will define themselves—and their needs—in different ways. In her research on students identified as basic writers, Linda Adler-Kassner found that while some students spoke of their writing weaknesses in terms of grammar and punctuation, most perceived their key writing "problem" as unsuccessful "information transfer," or the inability to put on to paper the ideas they have in their heads (79–80).

In terms of assessment, competing definitions of writing and writers can have important consequences. Faculty who define the "developmental writing problem" as a matter of correct grammar, for instance, will likely support assessments that focus on errors in writing; those who define it in terms of critical thinking will want assessments that highlight analytical abilities. Faculty who privilege acculturation may resist any kind of assessment that focuses solely on improved writing ability without considering improvements in students' dispositions toward school, attendance, participation, and so on. Similarly, problems can arise when there is a gap between what one "stakeholder" (student, faculty member, administrator) believes an assessment is, or should be, evaluating and what another interested party believes. As Adler-Kassner discovered at her institution, neither the placement mechanism nor the required basic-writing course accounted for students' own sense of their writing needs. The unfortunate results of disjunctions like these are frustration and even anger on the part of students who often don't know why they have been placed into a particular class or feel that the class does not address their presumed deficiencies. As Susanmarie Harrington puts it, as a result of not "understand[ing] their audience," many writing programs end up not being as "rhetorically effective" as they could be if they took the time to actually talk with students (105). Because of the power of assessment to sort students and distribute rewards and sanctions, rhetorical sensitivity is critical when designing tests, implementing them, or reporting results. It does little good to "accurately" place students into classes if they see no need for it themselves, which is part of the rationale for directed self-placement (Royer and Gilles).

Of course, even if faculty agree on basic constructs and how to appropriately assess them, they may not agree on the interpretation of the results of a given assessment, which creates other challenges. As Smith ("Assessing") acknowledges in his discussion of the pros and cons of holistic scoring of placement exams, for example, there is a host of contextual influences on how the exams are read and evaluated. These influences include how long the readers can read and "efficiently" score; how many essays are being read; how the rating scale is designed; the raters' understanding of the scale; and whether students clearly fit the categories designated in the scale (150–55). Other scholars (e.g., Williamson, "Introduction"; Pula and Huot; Ball; and Broad) have found that despite attempts at "norming" raters so that they read consistently, and in line with a published scoring rubric or guide, assessment decisions are nonetheless influenced by such environmental factors as raters' reading and writing experiences, disciplinary training, teaching experience, scoring experience, ethnic identity, and attitudes toward a particular writing assessment or use of results.

Faculty differences can also influence the ultimate use of assessment data. This is especially true when instructors most involved with assessment have one idea about what constitutes good writing and how such an ability should be evaluated, while those actually teaching the courses associated with the assessment have very different ideas. If those in charge of conducting assessment see the data as supporting the need for improved analytical-writing ability, for instance, but classroom faculty feel that students can't write analytically until they have mastered punctuation conventions, it is reasonable to expect that many students won't be doing much analytical writing (which will, in turn, influence the work that is assessed).

In short, faculty, like students, both reflect and in many ways, are defined by individual histories, cultural identities, and disciplinary knowledge that affect not only their general approach to teaching but also specific teaching practices. Because many of the competing definitions about who students are and what they need come from faculty, it is also important to understand who the faculty are, what their values and beliefs about writing are, and how these may affect assessment. The methods used to understand faculty are similar to those used with students: reviewing program materials (such as annual reports, course syllabi, and textbooks); conducting informal interviews (over lunch or through brown-bag forums); listening carefully to discussions about writing and assessment in department meetings; and observing classes.

Ultimately, though, understanding the faculty who teach the courses—including their assumptions about the students, teaching, writing, and evaluation—requires an understanding of the institutional environment in which they work. For example, institutional practices surrounding staffing, particularly of general education courses, convey tacit messages about faculty status and security that can greatly influence an assessment's design, implementation, results, and consequences. Generally speaking, faculty who feel they have more power within a program will participate differently in assessment initiatives than faculty who feel they have little or no power. Cooper, Sipe, Dewey, and Hunt illustrate this point in their discussion of an innovative portfolio-based assessment designed by composition specialists to replace a problematic "minimal competency exam." During the assessment pilot, tensions among two key groups of portfolio readers (graduate TAs and adjunct faculty) arose not just due to differences in experience but because of disparities in perceived institutional standing. The TAs, who had recently "quadrupled" in number, saw the assessment as a way to learn more about composition theory and, thus, build professional credentials; adjunct faculty, who had recently been described as "a dime a dozen" by the university president, perceived the assessment as a vehicle for exposing their weaknesses as teachers and getting them fired. Importantly, the comments on and about the student writing evaluated during the assessment reflected the degree of perceived threat as much as the particular teaching philosophy or approach.

While faculty positioning can create special challenges, where the WPA is located within an institutional structure can also have an impact on large-scale assessment. As anyone who has tried to secure funding for an academic initiative knows, organizational structure influences the flow of resources. Typically, a department-based WPA looking for money must confront many institutional layers as a request for financial support goes to the chair, the dean, and then sometimes on to the provost or university assessment coordinator. A WPA who works more directly with upper-level administrators often enjoys more direct access to decision makers, making it possible to not only bypass several levels of bureaucracy but informally pave the way for moral and financial support of assessment projects.

Another important institutional consideration, which has been covered extensively in the professional literature, is whether or not the WPA position is tenure line, and if it is, whether the WPA has received tenure and/or promotion. Typically, a tenured WPA will command more respect than a program administrator who is untenured or holds a staff position. Of course, sometimes theoretical status can be at odds with how a position is actually perceived. At some schools, for instance, the institutional position of the WPA is low, yet the everyday, practical status is quite high—especially if a school is emphasizing writing, is committed to hiring writing faculty with appropriate experience and credentials, or if the WPA is respected by faculty and administrators *because* of her demonstrated expertise.

Complicating the influence of institutional positioning are the multiple effects of how the role of program administrator is “read” or constructed by departmental or program colleagues who will be assisting with assessment. In programs where writing is mainly taught by adjunct faculty and graduate TAs, the WPA will be positioned “above” most classroom instructors as a supervisor or manager (Sledd; Bousquet). The challenge in this arrangement is that any administrative initiative will be seen as an imposition—even when the WPA attempts to be collaborative. Many former and current WPAs have described frustrated attempts to challenge institutionalized hierarchies through such means as co-directorships and collaborative mentoring (e.g., Schell; Harrington, Fox, and Hogue; Miller). If most of the courses are taught by full-time faculty, as is often the case at small liberal arts schools, the WPA will be regarded more as a facilitator and less as a manager or “boss”—especially if he or she does not have tenure. As Rebecca Taylor explains, full-time tenure-track faculty who support a writing program (in her case a well-established WAC program) can be fiercely independent—both in terms of their teaching and the assessment of teaching efforts (59–60). She underscores the importance of avoiding approaches that might threaten faculty members’ “academic freedom or pedagogical decision making” (60). In a program where a mix of faculty (including TAs, part-timers, full-time lecturers, and tenured professors) teach the writing courses, the WPA’s role can be even trickier to negotiate.

Broader social and political factors can also influence how the program administrator is perceived within a specific academic environment. Feminist compositionists (e.g., Barr-Ebest; Miller; Schell) have written convincingly about how larger cultural factors, like gender, can affect how administrative abilities and efforts are perceived. If, as Ebest-Barr suggests, for example, a female WPA is more likely than her male counterpart to be viewed by colleagues and other administrators as a service practitioner (rather than a serious scholar), her administrative work (including assessment) may not be as highly valued. Though not addressed much in the literature on writing program administration, other sociopolitical issues that affect institutional identity and positioning—issues such as race, class, and sexuality, which often are discussed in relation to scholarly productivity and teaching success—certainly play a role in how a WPA is read by students, faculty, and upper-level administrators.

The Practice of Assessment: Creating the Culture and Determining What It Will Tolerate

In an ideal world, writing faculty, including program administrators, would embrace large-scale assessment as a powerful means to positively affect teaching and learning, and they would have the historical, theoretical, and contextual knowledge needed to argue for and design meaningful approaches. While this article presumes less-than-ideal assessment realities, it also acknowledges the willingness and readiness of many readers to change their own perspectives on assessment and encourage colleagues to do the same. In fact, our use of the term *creating* to capture our vision of the future for assessment assumes that change—change in current attitudes, understandings, approaches, and, ultimately, perceived outcomes—will be the first step toward a new culture of assessment. We use the phrase *culture of assessment*, instead of something more prosaic such as *writing assessment program*, deliberately, to evoke the dynamic approach to assessment that is supported by research and theory. A culture is not static, or finite; it is ever-evolving and web-like, encompassing many interconnected values, practices, and people. Because the best assessments are ongoing and necessarily involve every component of a program, including students, faculty, administrators, curriculum, and resources, and because they both reflect and are affected by individual and institutional values and beliefs, culture is, we believe, a fitting metaphor for the approach that we endorse. Since our position requires a fundamental paradigm shift for many writing faculty and administrators, we conclude by acknowledging the difficulties associated with such a shift and providing suggestions for minimizing these difficulties so that theoretically consistent assessments can be implemented.

First, WPAs who have experienced success with large-scale assessments understand that if there is no "buy in" on the part of key stakeholders, there will likely be scant moral and financial support and little long-term, sustainable success. Furthermore, they realize that all views of teaching and assessment are, implicitly or explicitly, theory based, and that because of this, successful assessment initiatives often require wholesale changes in how faculty view student learning, the role of the teacher in such learning, and so on. Since differences in teaching philosophies often run deep, touching "founding insights, beliefs, [and] axioms" (Harris 42), it can be very dangerous for a WPA or group of well-meaning faculty to impose a paradigm based on a conflicting philosophy or, worse yet, to suggest that colleagues with differing philosophies are uneducated or "out of it." What seems like a commonsense effort on the part of some to be more theoretically consistent may feel like a threat to the teaching identity, self-concept, and even worldview of others. At the same time, while it is important to acknowledge the experience and expertise of all participants, not all participants should exert the same influence on assessment design. Ultimately, writing assessments need to be based on solid research, theory, and "best practices" in composition and rhetoric. The challenge is to respect the experience and values of others while moving discussions of assessment in a historically informed, theoretically sound direction.

One way to foster a positive climate for assessment without alienating colleagues is to investigate participants' experiences, interests, and values (as we suggested earlier) and also recognize similarities and differences at all stages in the assessment process. This approach allows faculty, staff, and administrators to feel included and to see their interests and ideas acknowledged, if not actually represented, in assessments. In fact, differences in experience, if acknowledged and discussed, can be used to create more effective writing assessments, as Smith found through his student-placement research. Smith discovered that the accuracy of first-year student placement into composition courses improved at his school when teachers who themselves regularly taught the courses placed students based on their own local knowledge of students and the curriculum rather than according to scoring rubrics that are meant to minimize the impact of teacher differences. While an assessment like this one may be based on philosophies that differ from those held by faculty, the fact that faculty experiences are acknowledged through their involvement can make a big difference in terms of attitude and openness to changes based on assessment results. Of course, as mentioned earlier in our discussion, when participants are graduate students (as they were in Smith's situation) and part-time instructors, it can be easier for assessment administrators to include their perspectives and ideas while implementing theoretically sound decisions; when participants are tenured and tenure-track faculty, the negotiations may be trickier. In the latter situation, using an outside consultant from another institution can be especially helpful for ensuring that assessments are grounded in sound research and theory and all participants are involved.

The same approach used with faculty—acknowledging their experience, perspectives, and expertise and including them in discussions and decision making throughout—can help a WPA develop the kind of working relationship with upper-level administrators that makes requests for resources easier to imagine and submit. In fact, good relationships, built on mutual respect and trust, can result in deans and provosts providing greater-than-expected leeway in how an assessment is designed and conducted. Sometimes, if simply shown how a theoretically valid writing assessment will fit university requirements, many upper-level administrators will be glad to let experienced and knowledgeable faculty spearhead assessment initiatives. As Richard Haswell and his colleagues discovered when their university's faculty senate mandated the development of both a required writing placement exam and a "rising junior" test, interested administrators, such as the dean and vice provost, "listened" to their concerns about how plans for these exams were "shaping up" and considered their proposals for more theoretically sound alternatives (*Beyond* 14–15). As Haswell and Susan Wyche explain, "We discovered the first truth about institutional assessment of writing: no one feels competent to do it . . . .[T]he administration was delighted to find writing teachers who said they had an innovative plan and were eager to get a problematic task out of their own hands" (15).

Good working relationships inspire the kind of mindset that allows members of different constituencies to assume the best (or at least not the worst) of one another. Sometimes, what looks like an upper-administration assessment mandate that ignores values and beliefs shared by composition and rhetoric specialists may be an expression of a highly complex and specific set of circumstances, some of which may no longer exist. When, as a new WPA, Ruth Mirtz researched the genesis of the placement exam being used at her school ("a multiple choice grammar and usage test"), she discovered that the exam didn't so much reflect an individual or even small-group agenda as much as ambivalence among English faculty about having to create a placement process that would correspond with "the literal politics and the tacit ideologies" of a state legislature and educational system existing during the late 1970s and early 1980s, when the test was adopted (127). Other times, upper-level administrators may present otherwise-flexible assessment approaches, using specialized, unfamiliar terminology (e.g., "outcomes" versus "objectives," or "reporting group" versus "department/program") that suggests methodological constraints where none exist. Additionally, administrators, knowing little about accreditation themselves, sometimes interpret accrediting agency language more narrowly than necessary. All that may be required here is a program director or faculty member who is knowledgeable about what terms such as *validity* and *reliability* really mean to explain why interpretations can be broader.

While we offer various suggestions for creating new cultures of assessment in writing programs, and departments and colleges as a whole, we must acknowledge that this is easier said than done. The reality is that sometimes the values and beliefs of upper-level administrators or institutional-assessment personnel can override a WPA's knowledge of assessment history, current theory, and local environment. In these cases, the WPA and other writing faculty should not give up, but instead they should do what they can with the support and resources they have and should continue arguing for responsible, theory-based, context-sensitive assessments, based on this article, the other articles and chapters we cite, and professional standards and guidelines for assessment, available from WPA, NCTE, and elsewhere. As they present their informed cases for expert-reader placement methods or portfolio-based competency testing, they should ask those with contrasting ideas to back up their arguments based upon current research and theory on educational assessment. While it is not always possible to "change the world" of assessment, it is possible to push colleagues and administrators to publicly acknowledge their commitment to efficiency or to articulate their assumptions about good writing, teaching, and learning and, by so doing, spark a gradual shift toward a more ideal assessment culture.

Notes

This article draws on the authors' new book, *A Guide to College Writing Assessment* (O'Neill, Moore, and Huot).

1. While the CEEB, like ETS, is a nonprofit corporation, it is nonetheless a corporation employing extensive marketing and management strategies to generate revenue. Since the formation of the CEEB, testing has grown into a multi-billion dollar industry, with a multitude of for-profit companies as well as nonprofits such as ETS and ACT.

2. Psychometrics, the science of educational and psychological measurement, is primarily concerned with the development of instruments and procedures for measuring personality, ability, and achievement. These instruments and procedures include various types of tests for gathering data and statistical formulas and mathematical techniques for analyzing data.

3. Developers of admissions tests like the ACT and the SAT caution against using one measure (explicitly referring to their own tests) to make admissions decisions. They do recommend combining ACT/SAT scores with high school grades and other measures to increase the numerical probability of accurately predicting first-year student success in college.

Works Cited

Achieve, Inc. "What Is the American Diploma Project Network?" 1 Aug. 2007 <http://www.achieve.org/node/604>.

Adler-Kassner, Linda. "Just Writing, Basically: Basic Writers on Basic Writing." *Journal of Basic Writing* 18.2 (1999): 69–90.

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 H. Fitzgerald. Mahwah, NJ: Lawrence Erlbaum, 2001. 85–100.

American Educational Research Association, American Psychological Association, and National Council on Measurement in Education. *Standards for Educational and Psychological Testing*. Washington DC: AERA, 1999.

Ball, Arnetha. "Expanding the Dialogue on Culture as a Critical Component When Assessing Writing." *Assessing Writing* 4: 169–202.

Bard, Leonard L. "Do Grades and Tests Predict Adult Accomplishment?" *Research in Higher Education* 23 (1985): 3–85.

Baron, Jonathan, and M. Frank Norman. "SAT Achievement Tests and High School Class Rank as Predictors of College Performance." *Educational and Psychological Measurement* 52 (1992): 1047–55.

Barr-Ebest, Sally. "Gender Differences in Writing Program Administration." *WPA: WritingProgram Administration* 18.3 (1995): 53–73.

Bean, John C., David Carrithers, and Theresa Earenfight. "Transforming WAC through a Discourse-Based Approach to University Outcomes Assessment." *WAC Journal* 16 (Sept. 2005): 5–21.

Bousquet, Marc. "Composition as Management Science: Toward a University without a WPA." *JAC* 22 (2002): 493–526.

Brandt, Deborah. "Sponsors of Literacy." *College Composition and Communication* 49 (1998): 165–85.

Brereton, John C., ed. *The Origins of Composition Studies in the American College, 1875–1925: A Documentary History.* Pittsburgh: U of Pittsburgh P, 1995.

Broad, Bob. *What We Really Value: Beyond Rubrics in Teaching and Assessing Writing*. Logan: Utah State UP, 2003.

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.

Cooper, Allene, Martha Sipe, Teresa Dewey, and Stephanie Hunt. "What Happens When Discourse Communities Collide? Portfolio Assessment and Non-Tenure-Track Faculty." *Administrative Problem-Solving for Writing Programs and Writing Centers: Scenarios in Effective Program Management.* Ed. Linda Myers-Breslin. Urbana, IL: National Council of Teachers of English, 1999. 44–52.

Cronbach, Lee J. "Five Perspectives on Validity Argument." *Test Validity*. Ed. Howard Wainer and Henry Braun. Hillsdale, NJ: Erlbaum, 1988. 3–17.

Elliot, Norbert. *On a Scale: A Social History of Writing Assessment in America.* New York: Peter Lang, 2005.

Gleason, Barbara. "Evaluating Writing Programs in Real Time: The Politics of Remediation." *College Composition and Communication.* 51 (2000): 560–88.

Greenberg, Karen L. "Grading, Evaluating, Assessing: Power and Politics in College Composition." *College Composition and Communication* 49 (1998): 275–84.

Harrington, Susanmarie. "The Representation of Basic Writers in Basic Writing Scholarship, or Who Is Quentin Pierce?" *Journal of Basic Writing* 18.2 (1999): 91–107.

Harrington, Susanmarie, Steve Fox, and Tere Molinder Hogue. "Power, Partnership, and Negotiations: The Limits of Collaboration." *WPA*: *Writing Program Administration* 21. 2–3 (1998): 52–64.

Harris, Joseph. *A Teaching Subject: Composition since 1966*. Upper Saddle River, NJ: Prentice Hall, 1997.

Haswell, Richard H., ed. *Beyond Outcomes: Assessment and Instruction within a University Writing Program*. Westport, CT: Ablex, 2001.

———. "Multiple Inquiry in the Validation of Writing Tests." *Assessing Writing* 5 (1998): 89–109.

Haswell, Richard, and Susan Wyche. "Authoring an Exam: Adventuring into Large-Scale Writing Assessment." *Beyond Outcomes*. Ed. Richard Haswell. Westport, CT: Ablex. 13–24.

Hillocks, George. *The Testing Trap: How State Writing Assessments Control Learning.* New York: Teachers College P, 2002.

Huot, Brian. *(Re) Articulating Writing Assessment for Teaching and Learning*. Logan: Utah State UP, 2002.

———. "Toward a New Theory of Writing Assessment." *College Composition and Communication* 47 (1996): 549–66.

Johnston, Peter. "Constructive Evaluation and the Improvement of Teaching and Learning." *Teachers College Record* 90: 509–28.

Lynne, Patricia. *Coming to Terms: Theorizing Writing Assessment in Composition Studies.* Logan: Utah State UP, 2004.

Messick, Samuel. "Meaning and Value in Test Validation: The Science and Ethics of Assessment." *Educational Researcher* 18.2 (1989): 5–11.

Miller, Hildy. "Postmasculinist Directions in Writing Program Administration." *WPA: Writing Program Administration* 20.1–2 (1996): 49–61.

Mirtz, Ruth. "WPAs as Historians: Discovering a First-Year Writing Program by Researching Its Past." *The Writing Program Administrator as Researcher*. Ed. Shirley K. Rose and Irwin Weiser. Portsmouth, NH: Boynton/Cook, 1999. 119–30.

Moss, Pamela. "Can There Be Validity without Reliability?" *Educational Researcher* 23.2 (1994): 5–12.

———. "Joining the Dialogue in Validity Theory in Educational Research." *Blurring Boundaries: Developing Writers, Researchers and Teachers.* Ed. Peggy O'Neill. Cresskill, NJ: Hampton, 2007. 91–100.

———. "Shifting Conceptions of Validity in Educational Measurement: Implications for Performance Assessment." *Review of Educational Research* 62.3 (1992): 229–58.

O'Neill, Peggy. "Moving beyond Holistic Scoring through Validity Inquiry." *Journal of Writing Assessment* 1 (2003): 47–65.

O'Neill, Peggy, Cindy Moore, and Brian Huot. *A Guide to College Writing Assessment.* Logan: Utah State UP, 2009.

O'Neill, Peggy, Sandra Murphy, Brian Huot, and Michael Williamson. "What Teachers Say about Different Types of State Mandated Writing Tests." *Journal of Writing Assessment* 2.2 (2005): 81–108.

O'Neill, Peggy, Ellen Schendel, and Brian Huot. "Defining Assessment as Research: Moving from Obligations to Opportunities." *WPA: Writing Program Administration* 26.1–2 (2002): 10–26.

Parkes, Jay. "Reliability as Argument." *Educational Measurement: Issues and Practice* 26.4 (2007): 2–10.

Peeples, Tim. "'Seeing' the WPA with/through Postmodern Mapping." *The Writing Program Administrator as Researcher*. Ed. Shirley K. Rose and Irwin Weiser. Portsmouth, NH: Boynton/Cook, 1999. 153–67.

Pula, Judith J., and Brian A. Huot. "A Model of Background Influences on Holistic Raters." Williamson and Huot 237–65.

Royer, Daniel J., and Roger Gilles. "Directed Self Placement: An Attitude of Orientation." *College Composition and Communication* 50 (1998): 54–70.

Sacks, Peter. *Standardized Minds: The High Price of America's Testing Culture and What We Can Do to Change It*. Cambridge, MA: Perseus, 1999.

Schell, Eileen E. "Who's the Boss? The Possibilities and Pitfalls of Collaborative Administration for Untenured WPAs." *WPA: Writing Program Administration* 21.2–3 (1998): 65–80.

Shepard, Lorrie A. "The Role of Assessment in a Learning Culture." *Educational Researcher* 29.7 (2000): 4–14.

Sledd, James. "A Return to Service." *Composition Studies* 28.2 (2000): 11–32.

Smith, William L. "Assessing the Reliability and Adequacy of Using Holistic Scoring of Essays as a College Composition Placement Technique." Williamson and Huot 142–205.

———. "The Importance of Teacher Knowledge in College Composition Placement Testing." *Reading Empirical Research Studies: The Rhetoric of Research*. Ed. R. John Hayes. Norwood, NJ: Ablex, 1992. 289–316.

Taylor, Rebecca G. "Preparing WPAs for the Small College Context." *Composition Studies* 32.2 (2004): 53–73.

White, Edward M. "The Misuse of Writing Assessment for Political Purposes." *Journal of Writing Assessment* 2 (2005): 21–36.

———. *Teaching and Assessing Writing*. 2nd ed. Portland, ME: Calendar Islands, 1994.

Williamson, Michael M. "An Introduction to Holistic Scoring: The Social, Historical, and Theoretical Context for Writing Assessment." Williamson and Huot 1–44.

———. "The Worship of Efficiency: Untangling Theoretical and Practical Considerations in Writing Assessment." *Assessing Writing* 1.2 (1994): 147–73.

Williamson, Michael M., and Brian Huot, eds. *Validating Holistic Scoring for Writing Assessment: Theoretical and Empirical Foundations*. Creskill, NJ: Hampton, 1993.

Cindy Moore

Cindy Moore serves as associate professor and chair of the Department of Writing at Loyola University Maryland. Over the past ten years, she has directed writing programs and taught a wide range of writing courses at universities in Indiana, Minnesota, and Kentucky. Her current scholarly interests include professional mentoring and program assessment. She has co-edited or co-authored several books, including *The Dissertation and the Discipline: Reinventing Composition Studies; Practice in Context: Situating the Work of Writing Teachers;* and *A Guide to Professional Development for Graduate Students in English*.

Peggy O'Neill

Peggy O'Neill focuses on writing assessment theory and practice as well as writing program administration and the disciplinarity of composition and rhetoric. Her work has appeared in such journals as *College Composition and Communication*, *Composition Studies*, and the *Journal of Writing Assessment*. She has edited or co-edited four books, the most recent being *Assessing Writing: A Critical Sourcebook* and co-authored *A Guide to College Writing Assessment* with Cindy Moore and Brian Huot. Currently she is an associate professor and director of composition in the Writing Department at Loyola University Maryland.

Brian Huot

Brian Huot is one of the founding editors of *Assessing Writing* and more recently the *Journal of Writing Assessment*, which he continues to edit. He has authored and co-authored numerous articles and book chapters and co-edited five scholarly books, the most recent being *Assessing Writing: A Critical Sourcebook* (edited with Peggy O'Neill). In 2002 he published the monograph *(Re)Articulating Writing Assessment for Teaching and Learning*. He is professor of English and coordinator of the Writing Program at Kent State University.