

7

WRITING ASSESSMENT PRACTICE

It might seem anticlimactic to conclude this book with a chapter on writing assessment practice, since the entire book has focused in some way or another on writing assessment practice, whether it be to emphasize its theoretical properties in chapter four, to talk about teaching assessment to student writers in chapter three, or to detail how to prepare new teachers to assess and respond to student writing in chapter five. By now, it should be clear I believe writing assessment theory is inextricably linked to its practice. Assessment is also linked to teaching, since it's impossible to teach writing or learn to write without constantly engaging in assessment. As I mention in chapter six, people often want ready-made assessment practices as if they were machines that could be purchased off the shelf; they'd rather not engage in a process of inquiry and research that requires the asking and answering of specific questions by members of a community of learners and researchers. Most of the time, I avoid providing examples of assessment practices, since I believe in their situated and contextual nature, and I think that administrators, faculty and students profit much from the process of developing their own assessments. Instead of offering examples of assessment practices, I try to give a working methodology that will help them to create their own assessments.

Throughout this volume, I've attempted to blur boundaries between theory and practice, noting from the beginning of chapter one that my aim is not to create a grand scheme or explanation, what I would call Theory with a capital T. Capital T theory usually means some sort of formal statement that details

a set of principles. It is necessarily abstract because it is meant to cover many different contexts, situations and locales. It is deductive because the principles it constructs are meant to be applied to specific situations. Small t theory refers to the beliefs and assumptions within a specific context, situation or locale. Small t theory is inductive, since it comes from a specific practice and is not abstract or applicable beyond that instance. The two kinds of theories are related, and my distinction between the two types of theory is in some ways rhetorical, since I write about theory in terms of its relationship to practice for a specific purpose, resisting the common theory/practice split. As far as I am concerned, James Zebroski's definition of theory best fits my purposes:

Theory is not the opposite of practice; theory is not even a supplement to practice. Theory is practice, a practice of a particular kind, and practice is always theoretical. The question then is not whether we have a theory of composition, that is, a view or better, a vision of ourselves and our activity, but whether we are going to become conscious of our theory. (1994, 15)

Zebroski's vision of theory in composition is applicable to many of the discussions within this book about writing assessment, since I have often worked to take tacit beliefs and assumptions about writing assessment practice and make them visible, hoping to make readers conscious of the theories driving their practices. In some ways, this book is about the relationship between theory and practice, with the underlying message that neither can be separated from the other without furnishing an impoverished version of either theory or practice. I focus on practice in this chapter not as a way separating it from theory but to show how its connection to theory creates a new understanding of writing assessment, in much the same way I've tried to create a new understanding for various aspects of writing assessment throughout the volume.

My decision to avoid promoting specific practices and to emphasize instead the theoretical aspects of all approaches to writing assessment has had some interesting repercussions. For

example, in her article on the history of writing assessment, Kathleen Yancey (1999) differentiates me as a “theorist” from several people in the same sentence whom she calls “theorist-practitioners” (495). This distinction would be interesting and perhaps puzzling in and of itself, considering my long, practical association with writing assessment in several contexts—as a reader, scoring leader, designer, consultant and director. What makes Yancey’s labeling of me as just a theorist even more interesting is that a few years ago in response to an essay I published, “Toward a New Theory of Writing Assessment,” in *College Composition and Communication* (a revised version of which appears here as chapter four), Alan Purves questioned the theoretical basis of my work. His response was addressed to the editor of *CCC*, but Professor Purves had sent a copy to me, and in his own handwriting had written across the top of the letter, “Great article except for the ‘T’ word, Al.” In the letter attached, Professor Purves (1996) commended me for my work and urged me to go even further. His only critique was my use of the word “theory,” the T word:

In general I am put off by articles that claim to establish or even move toward theories, since it means they are attempting to reach a depth of abstraction that is perilous. But in Huot’s case, the attempt is praiseworthy because I do not see a theory, but a really practical approach to writing assessment.

I was thrilled that someone like Alan Purves had even read my work, let alone liked it, and I was anxious to answer him in the pages of *CCC*. Unfortunately, Professor Purves passed away less than a month later, and we never had our conversation. I have reread his response several times, and as I prepared to write this chapter on writing assessment practice I read it again. His comments are interesting within the context of Yancey’s, since she sees me as a “theorist” (495), and Professor Purves sees my article on theory as practical.

This apparent contradiction is not only interesting as I muse on the different ways my work is received, but it also points to larger questions about theory and practice in writing assessment.

One of the problems that has permeated writing assessment is the notion that theory and practice are separate, and that because of practical considerations we haven't the time for theory right now (Gere 1980; Faigley, Cherry, Jolliffe, and Skinner 1985; Cherry and Witte 1998). In fact, as I detail in an essay with Michael Williamson (1998), I began my study of writing assessment after he and I had a discussion about the claim that holistic scoring was a practice without a theory. Of course, if we believe as Zebroski (1994) urges, that all of our practices are imbued with theoretical properties, then our goal is not to create a theory, so much as it is to understand the theory that is already driving our practices or to create new practices that are more consistent with the theories we hold or want to hold.

Although I continue to resist any separation between theory and practice, it is helpful to acknowledge that working from practice to theory or from theory to practice causes us to do different kinds of work, both of which are necessary and important. In extrapolating beliefs and assumptions from practices, it is necessary that we work backward from the practices themselves to the ideas behind them. In creating practices, we need to be conscious of the principles we hold as we move outward toward new practice. Louise Phelps's (1989) PTP Arc is an example of how new practices are created through a theoretical consideration of older practices, moving us towards newer and better practices. On the other hand, we should not ignore another kind of movement in Phelps's model, since according to her, we cannot move toward new and better practices unless we explore more substantively the theoretical implications of these practices. In Phelps's model, this movement toward theoretical sophistication is represented by going deeper as the Arc pushes us out further toward new practice. Working just on the practical or theoretical level requires only one kind of movement; whereas, being a theorist-practitioner or engaging in what is otherwise known as reflective practice (Schön 1983) requires a two-way movement that can also be called dialectic.

Whether we call it reflective practice, the PTP Arc, or a dialectic, what's important to remember is that without the dual theoretical and practical action, real change in writing assessment and its teaching will not occur. Without this deepening theoretical component, our practices cannot be substantively altered. For example, even though in the last thirty years we have been reading student writing as part of formal assessment, the theory behind holistic, analytic, primary trait and other procedures developed to produce consistent scores among raters is still based on the ideas that produce multiple-choice or computerized editing tests like COMPASS; we are reading to produce reliable scores. When William L. Smith (1993) constructed his placement system upon the expertise of his readers and not their ability to agree (even though they agreed at a higher rate than he could train them to agree), then psychometric theory that holds that reliability is a necessary but insufficient condition for validity, no longer supported writing assessment practice. Teachers were now reading the writing of students about whom they were most knowledgeable in order to make an appropriate, contextual judgment. In this way, current theories about the ways we read and theories of how literacy is taught and learned became possible beliefs and assumptions for a whole new kind of writing assessment practice. My purpose in this chapter, then, is to explore the idea of practice as a specific component of writing assessment. As I have in several places throughout the volume, I offer some examples of practice, though I hope to emphasize the theoretical and procedural activities that help us arrive at such practices.

Although my work focuses on the relationship between theory and practice in writing assessment, I think the creation of a more formal theory for writing assessment based upon the principles from validity theory could be valuable for writing assessment as a field of study. Working out such a theory would be an important way for us to learn how to talk about assessment and to understand the ways in which our practices are limited and can be improved. As I outline in chapter two, however, such a

theory would need to recognize the important contributions of both college writing assessment and the educational measurement community. In fact, I think a theory for writing assessment would be a good way to link the different groups of people who work on writing assessment. As Pamela Moss has noted (1998), we also need to address the ways in which we talk about assessment and about how our discourse about assessment has created problems and limitations for the very students and programs assessment it is designed to help. Writing assessment theory could address these problems, linking together those who work in writing assessment and providing them with a coherent vocabulary for their joint venture. More importantly, a more formal notion of writing assessment theory would provide a constant reminder of the inextricable bond between theory and practice, ensuring that more and more assessment practices are held accountable to a theory that promotes teaching and learning.

THE NEED TO ASSESS

Edward White has warned the college composition community repeatedly that if we do not assess our students, teaching and programs, then it will be done for us from the outside (1994). Over the last several years, I have heard White's admonition played out in stories posted on the WPA listserv by writing program administrators (WPAs) in need of assistance with assessment. The story is always a little different in terms of the specific situation of the WPA, the program that is the object of the proposed assessment, and the person(s) or agency requiring the assessment. What remains constant in all of these stories is that a person who feels quite qualified and confident about her ability to run a writing program feels inadequate, beleaguered, and put out by the need to assess the program. The call for the assessment is always from the outside, from people who are not qualified to teach writing or administer a writing program. Often their notions of assessment are quite different from those of the WPA and others administering the writing program who are familiar with current teaching practices in writing. The usual culprits

calling for assessment are deans, provosts, other administrators, or outside accreditation agencies. The responses from people on the list are remarkably similar, depending on the type of program being assessed and the individual situation. What's remarkable as well is that even though people who are experienced and qualified in assessment respond to the posting requiring help, there appears to be no cumulative culture about assessment practice, since similar requests are made over and over. In fact, one active list member well known for expertise on assessment has even referred people back to earlier strands on the list via the archives, since the territory had been so thoroughly covered earlier. These stories, although common enough on the WPA listserv are even more prevalent if we consider the number of times we hear similar tales from colleagues across the country. In fact, such stories of writing specialists who become involved with assessment are also a part of the assessment literature (Elbow and Belanoff 1986; Haswell and Wyche-Smith 1994). Examining some of the basic tenets of these stories can provide some insight into the many factors that can influence writing assessment practice.

A set of assumptions about assessing, teaching, and administering writing programs emerge from these stories to give us a sense of the current culture and climate within which much writing assessment practice takes place. First of all, expertise in assessment is not considered important for those who administer writing programs. Second, assessment is not even considered a necessary part of administering a writing program.¹ This separation of teaching and administration from assessment has created a stratified approach to the teaching and assessing of student writing that mirrors the split in writing assessment scholarship I discuss in chapter two or in the ways we use assessment in the classroom that I discuss in chapter three. The problem of the separation of the two has long been a focus of writing assessment scholars. (White's germinal text is titled *Teaching and Assessing Writing*.) Third, writing assessment is often a reaction to outside pressures. People like White have urged writing teachers and program administrators to become proactive rather than

reactive, so that instead of being in a defensive position when it comes to assessment, we can take the offensive. In particular, White (1996) has urged writing teachers to become more sophisticated about statistical matters and the other technical specifics common to conventional writing assessment practices like holistic scoring. Fourth, people who have little or no knowledge about the teaching of writing or the administration of writing programs are often in positions of power to decide how writing and writing programs should be assessed. In all of these stories, any kind of assessment culture or link between assessment and teaching is missing, replaced by a sense of urgency and crisis.

I'm sure there are more assumptions that can be extrapolated from these stories, but these four seem daunting enough to give us an entry into talking about writing assessment practice. And, clearly the four assumptions are related to each other. Writing program administrators do not receive adequate education in writing assessment, because assessment is not commonly considered part of their jobs. Assessment and teaching, as I outline in chapter three, are considered separate and distinct from each other. Thus, the impetus for assessment necessarily comes from the outside; in a world where assessment and teaching are distinct, there would be no need to assess or to involve those responsible for instruction in assessment. Consequently, people with little expert knowledge about literacy and its teaching find themselves in the position of making decisions about assessment that ultimately shape curriculum and instruction.

The response from those in college writing assessment to these continuing scenarios has been to urge writing teachers and administrators to become more involved in assessment issues. While I agree with White's basic tenets about the need of those who teach writing to become more involved in its assessment, I disagree about the way it should be done. In addition to the need to work at the theoretical and epistemological levels I detail earlier in this chapter and in chapter four, I think it's important for those who teach writing to work toward altering

the power relationships inherent in most calls for assessment. Because WPAs typically do not have even a rudimentary understanding of assessment and calls to assess originate outside of the program, WPAs often find themselves in a relationship where they have to be accountable to a higher authority for something they really don't understand. Accountability is often constructed as an integral component in assessment practice. In this way, assessment is seen as calling teachers and administrators to task, so that they can account for their programs and students to a higher authority defined by the assessment itself. Often however, calls for assessment and insistence on certain assessment practices are part of a larger political agenda to achieve and maintain power and control over educational programs (Huot and Williamson 1997). As Michael Williamson and I have argued (1997), however, it is possible to understand assessment as responsibility rather than accountability, as a necessary and vital part of administering an educational program (Beason 2000). Writing program administrators should be responsible for providing persuasive evidence that their teachers and curricula are providing students with ample opportunities to learn according to recognizable and articulated goals. Being responsible rather than accountable alters power relationships, so that the responsible person has control and ownership over the programs and practices for which she provides evidence. She decides how the evidence is generated and analyzed rather than being accountable to some higher authority who chooses the assessment regardless of the programs or people being assessed. I understand that my use of *responsible* over the more common term *accountable* won't change many of the outside and often unreasonable calls for assessment that those who administer writing programs receive. I do think, however, that if we were to become more interested in and responsible for assessment, we would ultimately have better control over the fate of our courses, teachers and programs. Rather than advocate a proactive stance toward assessment rather than a reactive one, I think it's important for us to recognize that assessment, like education and literacy itself, can have profound

social implications. An often neglected but important fact is that assessment is a social action that can be used toward positive or negative ends.

A (RE)ARTICULATION FOR THE NEED TO ASSESS

Originally, assessment was designed to be a kind of social action, since it was supposed to disrupt existing social order and class systems. For example, the first formal programs for written examinations initiated in Ancient China (Hanson 1993), in nineteenth century America (Witte, Trachsel, and Walters 1986), and during the twentieth century through instruments like the SAT and organizations like the ETS (Lemann, 1999) were partly intended to interrupt the then current practice of awarding civil service appointments and educational opportunities based upon social position, family connections or other priorities unrelated to personal merit, achievement or ability. Unfortunately, as Michel Foucault details (1977), the concept of the examination is closely related to acts of punishment and hegemony toward those in society who hold positions of vulnerability. Tests and testing are constructed from specific social positions and therefore promote a particular social order designed to furnish the more powerful in society with a disproportionate number of resources and opportunities. Although there are many, here are two prime examples of testing as a negative, interested agent for social action. One example is Herrnstein and Murray's book *The Bell Curve: Intelligence and Class Structure in American Life* (1994) in which African-Americans' lack of access to education and other important cultural institutions is defended based upon their lower test scores. The second example comes from earlier in the century (circa 1930) when Louis Terman, the primary developer of the Stanford-Binet I.Q. tests, renormed the instrument after initial results showed that girls had outperformed boys (Darling-Hammond 1994). Tests as a pervasive, negative, shaping force on individuals by institutions should not be underestimated. F. Alan Hanson (1993), a cultural anthropologist details in his book, *Testing Testing: Social*

Consequences of the Examined Life, the myriad ways people are constrained, labeled and identified through a range of physical, psychological and educational tests. Unfortunately, there are many good reasons why tests and testing are regularly viewed as largely hegemonic exercises invested in reinscribing current power relations in American society, and why many writing teachers and writing program administrators would resist working in assessment at all.

This overall impression of assessment is exacerbated for those who teach writing, since one of the driving impulses in the formulation of new procedures for teaching writing that began in the 1970s was against current-traditional rhetorical practices that emphasized correctness and the assessment to enforce it. For example, because the COMPASS test used for college writing placement, which we discussed in the previous chapter, contains a passage of predetermined deviations from a prestige dialect a student must identify and correct, it defines writing and the teaching of writing in terms of the linguistic features that approximate the rhetorics and dialects of powerful groups. Regardless of the writing program into which students are placed, such a test sends a powerful message about the value of writing in the courses for which students are enrolled, when placement into the course is based upon her ability to proofread and edit a specific dialect.

Seeing writing assessment as social action helps us to recognize the power and potential for writing assessment to shape instruction, possibly enabling certain students while limiting others. It also helps to make clear that often assessments imposed from the outside have specific political agendas that are designed to profit certain groups of people. If we believe in the fundamental right and power of literacy for all students, promoting choice and social mobility for those who can complete a specific academic goal, then we need to design and implement assessments that will promote such objectives.

Recognizing assessment as social action requires a new understanding of our need to assess. Instead of assessment being a call

from the outside for us to be accountable for our programs or for an opportunity to be proactive or assess before it's done to us, assessment becomes the way by which we ensure that writing instruction provides successful educational opportunities for all of our students. It allows us to recognize not only the importance and ramifications of teaching literacy, but it also alerts us to the crucial nature of what we value in our students' and programs' performance. How we value these performances are as important as what we value. An assessment is always a representation (Hanson 1993) and as such it has the ability to take on a life of its own. Assessments are powerful cultural markers, whose influence ranges far past the limited purposes for which they might originally be intended. The systems we create to assess ourselves and our students can have much power over the ways we do our jobs, the kinds of learning our students will attain and how we and others will come to judge us.

A NEW ARTICULATION FOR WRITING ASSESSMENT PRACTICE

Ultimately, being able to understand writing assessment in new ways comes down to being able to change the way writing assessment is practiced. Since I believe that practice and theory are linked, real change in writing assessment means more than the number of samples we read or whether or not we write or edit on a computer. It is not easy to make any substantive changes in writing assessment practices because we must do more than just change practice, we must be able to disrupt the theoretical and epistemological foundations upon which the assessments are developed and implemented. Not only should we address practice on a theoretical or epistemological level, we must, as I outline in chapter six, also learn to look past the technological orientation of assessment and begin to see it as research that requires a community to ask and answer questions about value and judgment in order to make appropriate educational decisions. All of the chapters in this book advance a new articulation of writing assessment practice, whether it be the integral relationship of

assessment in writing and teaching others to write in chapter three, reconceiving the field of writing assessment in chapter two, or the way we envision our ability to respond to student writing in chapter five.

In addition to the process of re-conceiving the various sites and features of assessment, we must also, as I've attempted to describe in this chapter, begin to articulate the role that writing teachers, writing program administrators and educational assessment experts see for themselves. As Larry Beason (2000) notes in a recent collection on the ethics of writing instruction, it is our ethical obligation to determine how well our teaching and programs are helping students learn to write. Preparation for a career in writing program administration should also include instruction in the rudiments of writing assessment, and administrators should count the assessment of their programs as an ongoing part of the job. In other words, instead of seeing ourselves as proactive rather than reactive in writing assessment—where we assess before being asked to—understanding writing assessment as a vital and important site for social action can support teachers and protect students from political agendas and other outside pressures that can strip the importance and vitality of effective instruction in literate communication. We need to understand that assessment can be an important means for ensuring the values and practices that promote meaningful literacy experiences and instruction for all students.

In chapter four, I looked at some emergent writing assessment programs that are based upon the theoretical and epistemological bases that drive much current practice in writing instruction. In figure two of that chapter, I note the beliefs and assumptions that drive these practices: that assessment should be site-based, locally controlled and based upon the explicit teaching goals of the program being assessed. In formulating principles upon which writing assessment practice might be based., I would like to add to these principles lessons we have learned from other parts of the volume. In chapter six, we differentiated between technological and research-based

approaches to assessment. Technological approaches involve an application of a set of methods developed by others and used across sites and contexts. Research-based assessment, on the other hand, requires that the community of teachers, students and administrators come together to articulate a set of research questions about student performance, teaching, curriculum or whatever they are interested in knowing more about. In an attempt to answer these questions, research methods are employed in much the same way that we might approach doing research on any other issue or set of questions. In chapters two, four and six, we discussed the importance of understanding newer and more encompassing notions of validity. We've seen that validity pertains not just to whether a test measures what it purports to measure, but also that it scrutinizes the decisions that are based on a test—how they impact students, teachers and educational programs. In addition to seeing assessment as research, we are also responsible for validating these procedures—providing theoretical rationales and empirical evidence that make the argument that the decisions based upon our assessment have real educational value for our programs teachers and students. The latter set of principles should involve ideas developed earlier in this chapter, which suggest that writing teachers and administrators should see writing assessment as part of their responsibility and should initiate assessment efforts in the same way as they might revise curriculum, supervise instruction, or attend to other tasks important to effective educational programs. Following is a list of some guiding principles for writing assessment practice, as articulated in this chapter.

- Site-based
- Locally controlled
- Research-based
- Questions developed by whole community
- Writing teachers and administrators initiate and lead assessment
- Build validation arguments for all assessments
- Practicing writing assessment

The initial step in any writing assessment should involve the people interested in and affected by the assessment. I'm trying to avoid using the word *stakeholder*, even though it is a common term in educational measurement, because I don't accept the implication that different people with various motives all have a stake in an assessment, and that all of these stakes and the holders who represent them have an equal claim on the assessment in question. For example, how can we design a writing assessment that satisfies a politician's need for evidence that he is tough on education and supports strong standards, while at the same time tailoring the assessment to measure the strengths and weaknesses of individual students. In other words, if we really believe that assessment is a necessary part of making strong, appropriate decisions, then we must treat assessment the same way we would other educational decisions. This is not to say that we should exclude anyone from initial conversations designed to articulate the research questions that will drive the assessment. On the other hand, if decisions based on an assessment must promote teaching and learning, as current validity theory dictates, then we must be accountable to those people who are most expert about teaching and learning—students and teachers. In most specific instances, disagreements among constituents in an assessment will be limited to a couple of issues—for example, whether or not they can afford multiple sample assessments like portfolios. It is important to note that although I advocate favoring teacher and student concerns, the process of validating an assessment acts as a strong check on allowing them to “do whatever they want,” since ultimately the people conducting the assessment have to make an argument that any decisions made on the basis of the assessment can be supported by theoretical rationale and empirical evidence.

Although my way of “practicing assessment” favors the local development of writing assessment measures that privilege teachers and students, I recommend that any serious writing assessment initiative involve the hiring of an outside expert who can mediate disagreements, help design the assessment, and in

general guide local participants through the process of articulating questions, designing procedures to answer the questions and implementing a strong program to validate the assessment. As well, a consultant can draft proposals for assessment that situate the local project within the larger theoretical frameworks available in the assessment literature and help write any needed final reports. Involving an outside consultant is an important step in answering concerns of some that English teachers or writing program administrators might lack the technical knowledge to design and implement valid writing assessment procedures (Scharton 1996; Camp 1996).

In addition to formulating questions, the assessment should answer, it is also necessary in the initial stage to make decisions about the scope of the assessment, what decisions will be made on its behalf, who will be given the information obtained from the assessment and how this information might be shaped and disseminated. In leading several workshops on program assessment, I have developed an activity that I call “Design An Assessment Worksheet” that contains several questions that are important in the beginning stages of designing an assessment. Several of the questions on this worksheet (see figure 1) are a good place to begin. Questions one and two, which ask what it is we want to know about an assessment and where we would go to know about it, reflect the importance of assessment as research and underscore the relationship between what we want to know and how we will go about finding out. In other words, the methods we use depend upon the questions we have. Points 1 a), b), and c) are also good targeting devices, since they allow us to look at different groups of people, though I also think we might expand our targets for information and include such things as curriculum and instruction—perhaps even sub categories under a), b) and c) that would permit us to ask questions about students’ abilities in certain areas or different kinds of teacher or administrator activity. It’s important to understand that figure one is a generic version, and that it can and should be tailored to a specific set of needs.

Figure 1
Writing Assessment Worksheet

1. What do you want to know?
 - a) about or from students
 - b) about or from teachers
 - c) about or from administrators
2. How will you go about getting this information?
3. Who is this information for?
4. What use will be made of this information?
5. What form will the information take? Will there be a report? Who will write it?
6. What verification or corroboration will ensure the accuracy and consistency of the information?
7. How will the information be collected and how will the way it is collected help to improve the program being assessed?
8. Who will be affected most by the assessment, and what say will they have in the decisions made on behalf of the assessment?
9. What are the major constraints in resources, time, institution, politics, etc?

Questions three through five remind us that writing assessment is inherently rhetorical, since what we are trying to do is to create a document that makes a specific point about writing and its learning to effect some kind of action. In deciding what questions to ask and what data to gather to answer the questions, we also need to think about whom the eventual audience for the assessment might be. Richard Haswell and Susan McLeod (1997) have an interesting book chapter in which they script a dialogue that details the different kinds of reports that specific audiences might be interested in reading. Haswell and McLeod also note that effective assessment programs require different roles played by administrators, faculty and researchers. Important in these questions, as well, is the notion that certain decisions eventually will be based on the assessment. This is often a crucial component for accreditation agencies. It's not enough merely to collect and analyze data, we must also have a plan for how this information will be used. An assessment should result in a written document, or perhaps, as Haswell and McLeod note, a series of documents. It's important, going into an assessment, to have plans about what documents will be prepared and who will prepare them. I have often written, reviewed, or helped to write documents for institutions for which I served as a consultant. When I

authored these documents, they were reviewed by various local faculty and administrators, and then revised according to their feedback.

Questions six through eight refer to the ways in which the assessment itself will be reviewed and validated. The previous questions on the worksheet look outward to the questions we can articulate and the data we can gather and analyze to answer the questions. On the other hand, questions six through eight also look inward toward the process of assessment, asking how this information will be used, how we can be sure that the process will profit teaching and learning, and how we can ensure that those affected by the assessment will not be harmed. Simply, these questions point us toward how to craft an argument that an assessment can and should be used to make important educational decisions. These kinds of questions are the beginning of the validation process, and they affirm that validation is an important component of the process not only after an assessment has been conducted but in the planning stages as well.

Question number nine is an important consideration for many reasons. Local, site-based assessment needs the support of local administrators who control the purse strings of their institutions. As White (1995) warned us several years ago, assessment done on the cheap is often bad assessment. Quality assessment requires a serious investment of time and energy from those who design and implement it. And, as I advocate a larger role for English teachers and writing program administrators in assessing their programs, teaching, and students, it's also important for the institution to provide necessary support for faculty who lead assessment programs in terms of release time and extra pay. As well, it is necessary to compensate all those who work on an assessment project, whether they read student writing or are involved in other types of data generation, collection, or analysis. Before an institution can ask for assessment, they need to provide financial support, and any English teacher or writing program administrator who proposes an assessment plan should include a viable budget.

TWO MODELS FOR ASSESSMENT

I'm going to conclude this chapter by discussing two models for assessing student writing and writing programs that are supported by the principles we have developed over the seven chapters of the book. The first model is based on an ongoing assessment of the first-year writing program I currently direct. Because this model comes from a specific site and program, there will be problems with applying it unaltered to any other program, though it is possible to adapt it to most college writing programs. The second model comes from a presentation I heard a few years ago, and that has been reported in the scholarly literature (Cheville, Murphy, Price and Underwood 2000). This second model describes an assessment of the Iowa Writing Project that was conducted to satisfy a funding agency who wished to have an assessment of the project site before it would continue funding it. This model accomplished its goal of garnering continuing funding while at the same time providing very useful information for actually improving the work of the writing project itself.

The first model takes place in the University of Louisville Composition Program, which staffs around two hundred first-year writing courses (English 101 and English 102) per year. These courses are taught by part time instructors (PTLs), Graduate Teaching Assistants (GTAs), or full time professorial faculty. PTLs teach for a stipend by the course. GTAs, who can be enrolled in either an M.A. literature program or a Ph.D. program in rhetoric and composition, receive a yearly stipend plus tuition remission. Full-time tenure-track faculty are required to teach one first-year writing course per year. While the composition program goals stipulate the number of formal papers for each class and differentiate between a writing-process orientation for 101 and a research focus for 102, they do not dictate a specific curriculum or text. All non-faculty instructors are required to participate in a doctoral-level seminar on composition theory and practice that is designed to prepare them to teach in the program. There is great variety in the courses and approaches that individual instructors take, and

we encourage this diversity. Instructors have used thematic approaches as varied as basketball and surrealism.

This background information about the freedom that instructors in the program have in developing their own curriculum is important, because one assessment procedure we use focuses on whether or not individual instructors actually meet program goals. Every semester, all course syllabi are read to make sure they conform to the general guidelines we expect from each course. In this way, we are providing a mechanism for ensuring that curriculum does focus on the teaching of writing. Instructors whose syllabi do not seem to meet the general guidelines have an opportunity to make the point that they are in compliance or to revise their syllabi. Another way we assess curriculum is that we require all non-faculty instructors to compile and maintain a teaching portfolio in which they keep current copies of their syllabi, assignments, and all other instructional materials. We also ask that instructors write a reflection after each semester about how they feel their courses went. These portfolios allow the Composition Program to know what's going on in various classrooms, while at the same time providing instructors with freedom in course design and curriculum. The teaching portfolios also come in handy when instructors are applying for jobs or for admission to graduate schools. This kind of information about what people are doing in their courses can be collected without a lot of programmatic effort or expense.

While keeping syllabi and teaching portfolios on file provides an ongoing record of the curriculum in the various courses, we also have all non-faculty instructors observed on a regular basis. We provide observation forms that include a section for pre- and post-observation consultation in case the observer and instructor wish to meet before and after observations. The form also includes spaces for observers to comment on what they've seen and what they think about what they've seen. Observers are also asked whether the observation was acceptable. If they deem a class unacceptable (this has only happened a handful of times in the six years we have been conducting the observations), the

class is visited again by a member of the Composition Program staff. Observations help us to satisfy our accreditation agency's mandate that non-full-time faculty be observed on a regular basis. In addition, many full-time faculty find the experience enjoyable and informative, since all faculty teach some writing courses each year. Observations also help to de-privatize the classroom, making it a more public space. As well, observations provide information on teaching that can be used to satisfy mandates for program assessment, and this information is collected on a regular basis without great effort or expense. Like their teaching portfolios, some instructors put a lot of time into their observations, using the pre- and post-conferences as a way to reflect upon their teaching.

The last component of our program assessment focuses on student writing. Because evaluating student writing is something that requires some effort and expense, we do not assess student work every year.² In addition, because we are looking to evaluate the program and not individual students, it's not necessary that we assess every student's writing. We choose to look at about ten percent of the students' writing in each of the courses that constitute the two-course sequence required of most students. Because we are looking at a limited amount of student writing, we can choose to look at it in some depth. This depth consists of three separate tiers.

The first tier of evaluation is comprised of teachers who meet in three-teacher teams to read each other's students' writing. In addition to the portfolios or collections³ of student writing, each team also considers five students' high school portfolios,⁴ and the writing done for their first-year writing courses, since this gives us an opportunity to compare the kind of work students did in high school with what they are doing in college. Teachers read each other's students' writing and discuss their reading with each other. Although teachers record grades for all the student writing they read, only the grade by the student's own teacher will count for the student. Teachers also characterize the writing for each of the grades. For each of the three teacher

teams, we receive a list of grades given by each instructor and the qualities of writing for each of the grade levels. In addition, we receive grades for each of the five portfolio sets of both college and high school writing and the grades accorded for those portfolios. We also ask the teams to provide us with a discussion of the similarities and differences they see between high school and college writing.

In the second tier, we assemble a campus-wide committee, representing all the schools and colleges of the university, and we ask them to read the fifteen sets of high school and college writing, giving grades for each of the college collections and characterizing the qualities of writing for each grade. In addition, we ask them to discuss the similarities and differences they see between high school and college writing. The committee conducts its business via a listserv in which individuals can single out particular portfolios and papers for discussion. All discussions are archived to be part of the possible information to be used in analyzing the status of the program.

In the third tier, we assemble writing assessment and program professionals across the country on a listserv. We send all participants the fifteen sets of high school and college writing, and ask them to give grades to the college writing, indicating the rationale they use for grading. In addition, we also ask them to compare high school and college writing. We ask participants to single out any individual papers or sections of papers for discussion. In using faculty from across the country to talk about student writing over the internet, we are following the pioneering work of Michael Allen (1995) and others as they have used electronic communication to conduct program assessment (Allen, Frick, Sommers and Yancey 1997).

The three-tier evaluation of student writing provides us with a wealth of information about the ways in which our students' writing is read. It allows us to conduct an assessment that includes the voices of people who teach in our program, people who teach our students after their first-year of writing instruction, and people from beyond our campus who can give us a sense of how

our students would be perceived by those who teach writing, administer writing programs, and evaluate both writing and programs from across the country. Comparing grades and the characteristics for grades for each of the different groups provides us with important information about standards and outcomes. However, instead of imposing either outcomes or standards from the outside, ours come from the writing of our own students.

All of the information generated by the three tiers is reviewed by the Composition Director, the Assistant Director of Composition (an advanced doctoral student) and an outside consultant chosen from the third tier readers. These three people compile a report that is shared with all three tier participants before being revised. Generating such a wealth of information about student writing that is read by such a diverse group of people provides an opportunity for real discovery about the kinds of writing students are doing in the program and how successful different groups of readers consider this writing to be. Comparing high school and college writing should provide some useful information about the writing experiences students have as college students and how this experience builds on what they have done before college. This comparative information should be useful for both high school and college teachers, as it allows them to get a better sense of who their students are and what is expected of them. As well, the final report on the status of student writing should be used to revise appropriately course goals, faculty development opportunities, grading procedures, and other program guidelines and policies that the assessment shows needs revising or improving.

The second model was developed to assess the Iowa Writing Project (IWP), which had received a grant that stipulated that the success of the project should be independently reviewed. Although project administrators wanted to comply with the need to assess, they were also concerned with providing an assessment that helped teaching and learning—the overall goal of the IWP. “Project planners believed the best structure for the review was one in which assessors could act as

colleagues, offering participants feedback about a small sample of their work and the work of their students” (Cheville, Murphy, Price, and Underwood 2000, 149).

Teachers were invited to submit portfolios of their teaching and of their own students’ work. The IWP hired an outside consultant team to help design the assessment, read all the portfolios and provide feedback for both the teachers in the project and the outside funding agency. The assessment team read all portfolios, providing for teachers a description of the teaching practices available in their portfolios and a description of the writing of their students. They also compiled an inventory of effective practices available from their reading of teacher and student work. All teachers in the IWP received a detailed description of their portfolio and the portfolios of their students along with an inventory of teaching practices from across the project. The result was that teachers got an opportunity to look with new eyes at the assignments, curriculum, and instruction and at the output of their students based upon this curriculum and instruction. In addition, teachers were able to see what other teachers were doing and to make for themselves a comparison if they so wished. In this way, teachers received useful feedback about their instructional practices and their students’ work, while at the same time the funding agency received a detailed description of the kinds of teaching and learning that were ongoing at the IWP.

This program assessment model does a good job of providing feedback for teachers about their work with students. While the model stops short of giving explicit evaluative commentary of teachers, it does give them enough descriptive information about their own work and the valued practices of others to allow teachers to take the next step, if they so wish, in looking for new and better ways to teach students writing. Although this model does involve the use of outside evaluators, the effective teaching practices themselves come from the teachers in the program and are not imported by outsiders. There is also a strong attempt to provide teachers with the opportunity to improve

their own practices by allowing them a reflective pause through which they can see their work and the work of their students in the eyes of others without any explicit pressure or judgment. This model would work well for something like a writing-across-the-curriculum-program where content-area teachers could provide teaching portfolios and samples of student work that could be described for them, since the model provides such strong feedback for teachers. It might also be a good alternative for model number one described above, so that instead of looking exclusively at student writing, teachers could receive detailed feedback about their instructional practices.

Both of these models honor the principles for writing assessment that are the ultimate focus of this book. While both of these schemes involve outside participants, the participants themselves do not set the focus for the assessment or decide standards or outcomes for the programs being assessed. On the other hand, involving people from outside of a writing program or a specific institution in local assessment programs answers mandates for local assessments to be sensitive to standards beyond a specific locale or institution. In this way, the assessment is controlled by those who teach and administer the programs being assessed and maintains the site itself as the focus of the assessment, while at the same time answering outside calls for standards and accountability. Both of these models also offer legitimate inquiry into the programs where important questions are asked and answered, and the answers to these questions can be used to improve the program.

CONCLUSION

Clearly, this entire volume has been about writing assessment practice, since essentially I am interested in helping others create assessments that can advance informed decisions about the teaching of writing. For example, in chapter two I map the field of writing assessment, arguing for a unified field of scholars who recognize and respect each other's work and positions, while always maintaining that the main thrust of decisions based on

assessment must be for the promotion of teaching and learning. Chapter three defines keywords in classroom writing assessment, differentiating between assessment, grading and testing, so that we begin to understand and teach our students the importance of assessment in writing well. Chapter four unpacks the theoretical assumptions that inform many assessment practices, arguing (as I have argued in this chapter) that to substantively change assessment practices we must move to change the beliefs and assumptions that guide the practices. Chapter five considers the connections between the way we read student writing and the way we respond to student writers, attempting to provide the same kind of “practical” theory Alan Purves contended I had advanced for assessment outside the classroom. Chapter six characterizes the ways in which writing assessment can be both technology and research, emphasizing the benefits of using opportunities for assessment as chances to ask and answer questions about our students, teaching and programs. In some ways, this final chapter has been about all of those subjects, as we have focused specifically on practice and have realized that it is impossible to talk about writing assessment practice without recognizing a myriad of issues, some of which were the focus of the “theoretical” chapters in this volume.

One of my favorite quotes comes from the scene in *The Wizard of Oz* in which the Wicked Witch rubs her hands together with a pensive look on her face cackling, “These things must be done delicately, or you’ll hurt the spell.” Although the Wicked Witch’s words can be applied to many things, they are especially relevant to a discussion of writing assessment practice. The “spell,” of course, is teaching and learning, and unless we apply assessment in specific ways, it can be irreparably harmed. (This potential harm of assessment for teaching and learning is a topic that I have deliberately tried to avoid in this volume, although it informs my treatment of the subject, especially in chapter three, when I labor against certain connotations inherent in conflating assessment with testing and grading.) In this volume, I attempt to downplay the negative side of writing assessment because in

order to (re)articulate assessment as something controlled by teachers to promote teaching and learning, teachers must learn not to avoid it or to leave it in the hands of professional testers or administrators.

Just as I have argued in this chapter that we cannot consider assessment practice outside of our considerations of theory or our jobs as writing teachers or program administrators, ultimately to own assessment, we must learn to see it as a necessary part of understanding how to write and how to teach writing. In some ways, then, this book is about deprogramming a certain understanding of assessment—or perhaps “decodifying” it would be more accurate. I hope we can come to understand the necessary and important role assessment plays on all levels of learning to write and of documenting that learning for students, teachers, administrators, parents and public stakeholders. This type of conclusion, I’m afraid, leaves as many questions as it supplies answers—the main question being *is what I’m proposing possible?* Although what I call for is ambitious and far-reaching, it is possible. The kinds of changes I envision and advocate will not come easily. Unlike what Dorothy finds, the answers have not been with us all along. In this volume, I offer no definitive answers. In fact, I am certain that the practices I do advocate can and will be revised in a continuing process of validation and reflection. This volume then is just a beginning, a challenge for all of us who are dissatisfied with past and current writing assessment to create a new future.