

Writing Program Administrators and teacher trainers who want to support faculty efforts to apply contemporary composition pedagogy theory are the particular readers to whom I appeal in these pages. Equally targeted are those of us juggling administrative demands for general education curriculum that responds to corporate pleas for college graduates with critical thinking and collaborative working skills and to retention statistics' implicit argument for pedagogy that addresses what my institution refers to as "alternative learning styles." Generally, my argument is that these administrative efforts, as well as the pedagogies often collectively referred to as "liberatory teaching," are sabotaged by most institutions' teaching evaluation practices. Standard methods and instruments for evaluating instructors undermine or even negate the goals advocated by virtually all current theory and by most practice of writing instruction because those tools are informed only by conventional, performance-centered conceptions of teaching; in contrast, the social mission of most contemporary composition theories and the ethos of many practitioners are informed by more student-centered conceptions of teaching. This discrepancy undercuts not just the critical thinking and collaborative goals of liberatory pedagogy but also its counter-hegemonic intents and effects.

In the following pages, I'll appraise the assumptions that drive our standard

*Fostering
Liberatory
Teaching:
A Proposal
for Revising
Instructional
Assessment
Practices*

Jane E. Hindman

evaluative methods and compare them to those assumptions that undergird more critical approaches to teaching. In order to confront the contradiction between our theory and practice of teaching and our methods for evaluating teachers, I will also present an alternative evaluative instrument and explain how it more accurately measures what we say we believe to be effective teaching. Finally, I will offer the statistical evidence supporting the usefulness of the instrument, speculate on the implications of that evidence, and suggest further steps that will foster teaching practices that reflect the professed pedagogical theory and goals of composition studies.

What's at Stake: The Consequences of Evaluative Practices

To understand how the values driving a performance model of teaching may conflict with those central to the pedagogical theories and methods most favored in our discipline and others, consider what is probably the best-known exploration of innovative teaching in our discipline's last decade, namely Mary Louise Pratt's call for teachers to initiate "looking for the pedagogical arts of the contact zone" (455). With the clarity of hindsight, we can see how Pratt's descriptions of Stanford's revolutionary Western Civilization course not only portend the evaluative dilemma we now confront but also allude to the values intrinsic to effective teaching in the contact zone:

[t]he lecturer's traditional (imagined) task—unifying the world in the class's eyes by means of a monologue that rings equally coherent, revealing, and true for all, forging an ad hoc community, homogenous with respect to one's own words—this task became not only impossible but anomalous and unimaginable. (454)

To unify the world in the class's eyes is equally impossible and undesirable for pedagogues who not only reject monologic world views but also solicit students' participation in developing and organizing the content and practices of the course. Such goals are advocated by critical teachers (e.g. Ira Shor's *When Students Have the Power*) and feminist teachers alike (e.g. Kenway and Modra). Granted, feminist pedagogues often critique the patriarchal values manifested in some critical teaching (cf. Luke and Gore); granted, some of our scholars continue to challenge and refine Pratt's original suggestions about the role of oppositional discourse and critique and of homogenous definitions of culture in the college classroom (cf.

Miller, van Slyck). Nonetheless, and despite these disagreements about how to interpret and apply the theory and practice of critical pedagogy or of the pedagogical arts of the contact zone, most of us can probably agree on the answer to Pratt's question about assessing teachers' effectiveness:

Are teachers supposed to feel that their teaching has been most successful when they have eliminated such things [as students' "oppositional discourse, parody, resistance, critique"] and unified the social world, probably in their own image? (453)

"No!" is the reply to this obviously rhetorical question. This relative agreement about what does not constitute successful teaching sets the course for delineating the characteristics that do indicate effective teaching from the perspective of pedagogies emphasizing the values of student-centered, feminist, critical pedagogies. We can concede that these various pedagogies differ from teaching methods which position the teacher—her knowledge, her competence, her performance—at the center of the classroom; we can agree that they share some common goals and values: cooperative/collaborative learning, problem posing, critical thinking, challenging domination by appropriating social responsibility and authority, mediating conflicting points of view.

Pratt's rhetorical question regarding the advisability of teachers eliminating resistance, critique, parody and oppositional discourse in the classroom also helped us understand the importance of students and teacher(s) learning to mediate the inevitable conflicts between clashing cultures and viewpoints. But her questions regarding the stakes of traditional teaching and the possibility for reform in "contact zone" teaching evoke less obvious answers, less agreement: "Who wins when we do that [eliminate opposition and unify the social world in our own image]?" she asks, "Who loses?" (453).

In an abstract sense and especially to those who have studied and internalized the tenets of critical teaching, it seems obvious that the "winners" in classrooms that present unified world views are those whose social world is reflected in that image and whose discourse conforms to rather than opposes the status quo. Other commonplace implications are that the socially dominant "win" and that the "dominated"—those who are not members of the more powerful social group controlling the terms of interaction in the contact zone—"lose" when the educational institutions successfully eliminate resistance and maintain the illusion of a unified,

coherent, harmonious society. But even those who don't buy into these "givens" concerning the outcomes of liberatory pedagogy will probably balk at the price we pay when our evaluative practices undermine our educational goals: if it's indeed accurate that student-centered teaching actually does or even could better instill in students the qualities required in professional contexts, then who are the "losers" when our teaching and/or our teaching evaluation practices undermine student-centered teaching? It may well be that employers lose potentially well-trained workers, for according to most accounts the essential characteristics that employers now seek in their workers are the ability to apply knowledge in a variety of contexts and to work in multiple locations, problem solving skills, an appreciation of others' perspectives, and the ability to work in collaborative, project-based teams. If university-trained students don't gain these skills—which are consistent with the disciplinary goals for effective teaching but oppose individualistic, competitive models for teaching and learning—then universities may well lose the support of employers. If the value system informing a performance model of teaching fosters teacher's intellectual responsibility and mastery rather than students', then evaluative methods that sustain it nurture teachers' advancement, not students', for they foster future citizens who abdicate rather than grapple with responsibility, who conform rather than question a "master," who recognize but don't develop competence. In some ways then, all students lose out when institutional methods of evaluating teaching reward only performance models, not just those students less likely to respond to the values of mastery, competition, domination, individualism.

Teachers can be "losers" too if they are working to incorporate the values and methods of feminist, critical, student-centered pedagogies, for students and colleagues react strongly and often negatively to aspects of critical pedagogies. In part, these reactions are fashioned by popular cultural stereotypes of heroic teachers taking center stage and by the traditional methods of evaluation that privilege such performance. However, at least to Pratt, contact zone teaching necessitates conflict. In fact, *challenge, tension, confrontation, insecurity* are responses that seem to hallmark instruction that requires students to problematize their everyday situations, to construct their own meanings, to transform reality. By analogy then, one could argue that complacency, agreement, and safety characterize responses to teaching that centers on students' ability to memorize and internalize an authority figure's pre-determined interpretation of materials. The short term effects of traditional teaching methods are clearly and by far

more comfortable, and thus by their training, by routine, and perhaps even by nature students often don't like and—in fact—usually resist a teacher who demands that they critique, evaluate, and synthesize information rather than passively bank it. But the long-term effects of critical pedagogy promise better dividends. Bell hooks explains that

...[i]f one primary function of such a pedagogy is to prepare students to live and act more fully in the world, then it is usually when they are in that context, outside the classroom, that they most feel and experience the value of what they have shared and learned. For me, this often means that most positive feedback I receive as a teacher comes after students have left the class and rarely during it. (103)

Though dividends for teachers investing in the long-term response of their students do accrue, their short-term payoffs can be meager. Thus, an “aspect of radical pedagogy that has been difficult,” hooks explains,

is learning to cope with not being seen positively by students. When one provides an experience of learning that is challenging, possibly threatening, it is not entertainment, or necessarily a fun experience, though it can be. (103)

If adjusting to this feature of critical pedagogy is difficult for a teacher with hooks' professional experience and repute, imagine what it requires of the untenured professor, the forward-thinking secondary school teacher, the newly-appointed graduate teaching assistant. And yet that distress is precisely what some of our professional practices perpetuate: composition studies inculcates student-centered and/or liberatory pedagogies in its literature as well as in most teacher-training programs; meanwhile, institutional reliance on conventional, performance-model teaching evaluation methods countermand that scholarship and training. Thus, teachers, trainers, even the profession to a certain extent—they all “lose.”

What can we do to change the stakes of this game, to cut our losses? We have seen how business needs, our professional literature and teacher training practices, and even social ethics urge us to teach students to analyze, synthesize, and evaluate the authority and relevance of information. Teaching critical thinking skills, insisting that students take active responsibility for their own education, fostering collaborative and

mediation skills—these are all pedagogical practices that more teachers could and would incorporate if they were required, inspired, supported by a system that rewarded them. The next logical questions, then, are these: How might the system do that? How might we revise our instruments for evaluating instructors in order to foster liberatory pedagogical practices?

Re-vision

Constructing An Alternative Evaluative Instrument

Closing the gap between what our theory contends is important in the classroom and what our assessment practices actually support must involve constructing reliable teacher-evaluation methods that validly measure the kind of teaching that our profession values. In the language of the principles governing measurement, this term “validity” refers to the correspondence between what we say we value and what our instruments actually measure. According to Ed White, expert in writing assessment,

[a]lthough 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. . . .The difficulty comes from the questions behind validity—that is, the questions about what we want to examine. (10-11)

In *Evaluating the Teaching of Writing*, editor Christine Hult also advocates “tak[ing] a careful look at our evaluation goals and match[ing] goals with methods”; she further explains that “the authors [of Part II, Evaluation Methods] advocate multiple measures of teaching performance, not single measures, to give a more complete picture of a teacher’s effectiveness” (9).

Unfortunately, however, most institutions—for example, all of the three large state universities and the two community colleges where I’ve taught composition—rely on single measures of teaching performance; further, many insist that those measures yield quantitative results. To make matters worse, the single score that such measures produce may well be the only evidence of teaching effectiveness that personnel committees consider. In these situations—ones in which institutional contexts demand quantitative measures of teaching effectiveness—an instrument that fosters

student-centered teaching and yields quantitative results could be indispensable to supporting disciplinary goals. To construct such an instrument, however, we must first determine the essential practices of an effective liberatory teacher and how to enumerate them.

In most current evaluative instruments the operative definition of an effective instructor seems to be “producer of authoritative texts” and/or “master of a particular body of knowledge.” I contend that a more apt definition of a professional teacher (as opposed to scholar) is “one capable of modeling and facilitating particular conceptual habits of mind.” A critical teacher’s professional charge is not necessarily and certainly not only or even primarily to demonstrate her command of a body of knowledge, i.e. the reified “canon” of her particular expertise. Rather she should model for her students her professional habit of mind, her disciplinary problem-solving abilities. Most importantly, she must make possible students’ own problem-solving abilities. In order to do that, it’s often necessary to engage students in dialogue, not just with her but with each other, not just with others’ texts but with their own conflicting ideas. The conceptual habit of mind most crucial to writing instruction, perhaps to any instruction, I call “contextual meaning-making.” This quality distinguishes a critical teacher as professional as well as effective.

“Contextual meaning-making” as a determiner of an instructor’s effectiveness in a student-centered classroom is best measured by her ability to make meaning out of the particulars of any specific educational moment. As I see it, two key qualities comprise this conceptual habit. One is the capacity to analyze effectively the particulars of the context within which one is making meaning, that is, to observe the “scene,” gather the information, and assess the available means and forms of argument appropriate to the particular context. The second key quality in “contextual meaning-making” is the ability to connect and synthesize the particulars of the context within which the teacher is making meaning.

Thus, within the specific context of the classroom and the course she’s teaching, the instructor should present to the class specific examples from students’ work (e.g. journals, assignments, previous papers, comments in class) and then be able to recall and refer often to them herself and encourage students to do the same. In this way, she facilitates dialogue and problematizing students’ existential situations at least with respect to their coursework and—if students’ assignments ask them to utilize their personal experiences—to their everyday lives as well. Furthermore, the instructor should be able to make connections between students’ earlier work and the

points being made in subsequent class discussions or assignments, and then to categorize and re-categorize that work as it demonstrates or exemplifies new points. Such teaching ability assists students in understanding how the same texts, the same perspectives, come to mean differently depending upon the contexts within which they are applied. After modeling this recontextualization, this re-situating and therefore re-constructing of meaning, the teacher asks students to do the same. Thus, she fosters students' ability to develop and evaluate multiple interpretations and uses of information, to perceive from a variety of perspectives.

In keeping with this capacity to create the conditions for dialogue and to promote students' abilities to develop various viewpoints on one topic, liberatory teaching also calls upon a teacher's ability to exhibit her own and foster students' cultural mediation skills. Such skills are crucial for a practitioner of any pedagogy that problematizes the unequal power relations between people of varying languages, classes, and genders who inhabit the same social space. In discussing her early experiences with the Western Civilization course at Stanford known for its revolutionary emphasis on such power relations, Mary Louise Pratt explains that "the very nature of the course put ideas and identities on the line" (454). In such a setting, students and teacher must perceive and negotiate myriad points of view, must mediate conflicts and contradictions if they are to be effective in that contact zone.

In summary then, we've seen that the values and goals that distinguish student-centered teaching highlight collaboration, problem-solving, dialogue, mediation of conflicting points of view, and active critique of dominant points of view. The skills that support these values are modeling one's own skills in contextual meaning-making and fostering students' development of their own. A teacher best realizes these skills by recalling and referring often to examples from students' own work; by modeling the re- and de- contextualization of texts; by engaging students in dialogue not just with her but with each other; and by requiring students to develop, evaluate, and mediate their own and others' diverse points of view on a single topic.

I've constructed a unique student evaluation form intended to measure these practices, values, and goals that support effective, student-centered teaching. Consisting of 27 Likert-style questions in two major categories, the instrument's purpose is to assess teaching effectiveness in terms general enough to apply across disciplinary boundaries and to yield the quantifiable and statistically reliable formats required by most

institutional procedures (See Appendix A). Most importantly, the evaluation form is designed to achieve two major revisions in institutional evaluative practices: 1) measure and reward student-centered teaching and 2) re-shape students', teachers', and administrators' attitudes about the qualities that define effective teaching.

Reshaping these attitudes is crucial to sustaining the philosophical shift central to the distinctions between performance-centered and student-centered teaching. As David Bleich points out in his discussion of the questions of ideology associated with the evaluation of teaching, even in recent studies committed to improving teaching in postsecondary education, "[Ernest L.] Boyer and others who have addressed the matter of teaching in a constructive spirit nevertheless remain tacitly tied to the language and assumptions of competitive individualism" (13). Thus, Bleich argues, with respect to the teaching and evaluating of writing, we must conceive of a "friendlier ideology—friendlier to all as well as one more responsive to feminist thought" (11). First-time results of the instrument I've developed could conceivably satisfy the "friendlier" goal of supporting student-centered teaching. The goal of revising institutional attitudes about teaching effectiveness is more slowly realized, however, since it necessitates internalizing the values reflected in the instrument. Nonetheless, first-time results of the instrument may be able to predict the likelihood of students (and others) adopting a "friendlier ideology" of evaluation of teaching.

Probably most crucial of the innovations in the evaluative instrument is its inclusion of an entire section entitled "Questions about the Student's Experience." This set of eleven questions asks a student to consider how the course has improved her own knowledge and understanding, her ability to express her thoughts, and her capacity to perceive differences in her ideas and others. She's also asked to judge the extent to which the course required her to think for herself, interact with her classmates, engage with the course materials and with others' points of view, and change her mind about her opinions. Clearly, this section of the instrument is meant to improve the validity of our measuring effective student-centered teaching by directing students' attention to relevant aspect of their own performance in the course, not just the teacher's.

This intent is sustained even in the instrument's section entitled "Questions about the Teacher." Like its traditional predecessors, the revised instrument asks students to judge certain aspects of a teacher's performance such as her ability to present clearly the goals for the assignments and to explain the inter-relation of individual units of the course. However, the

alternative evaluation form's questions also ask whether or not the teacher included students in decisions about course goals, assignments, and reading materials. In further attempts to support the practices of a student-centered teaching, this section requires students to rank the teacher's familiarity with individual students' work and her ability to relate course material to students' interests and experience; her response to students' input in classroom interactions; her capacity to give students feedback that fosters their improvement and supports their taking risks in their coursework; her encouragement of opposing and divergent viewpoints; her expectations of students to act responsibly. To an extent then, these questions—like those in the first section of the revised instrument—also assess students' experience, for they rank a teacher's ability to facilitate student learning rather than to demonstrate her level of knowledge and/or preparation. In fact, other than its focus on an instructor's ability to explain the goals and structure of the course, the only attention this section gives to teacher performance as it exists (more or less) independently of students' participation is in its concern for whether or not the teacher uses her authority and her humor to promote learning and for her capacity to admit her mistakes and accept constructive criticism.

Thus, the validity of this alternative evaluation form resides in its focus on student performance as well as on the conceptual habits and practical behaviors that characterize critical teachers. By requiring students to consider their own functioning in the classroom as well as their teacher's capacity to facilitate their learning, this evaluative instrument should yield teaching-effectiveness scores that prize student-centered teaching. In addition, its original and provocative questions initiate efforts to re-shape students' attitudes about their instructors and their educational experiences.

Possibility: Testing the Revised Instrument

With the support of a Ford Foundation seed grant administered through Southwest Institute for Research on Women (SIROW), my colleague Kari McBride and I were able to test the instrument at our respective home institutions, the University of Arizona (UA) and San Diego State University (SDSU). Responses were gathered from 1420 students enrolled in 79 different sections, six of which were Professor McBride's AY 96-97 courses in Women's Studies (five lower division, one upper) and the remaining 73 of which were from SDSU's Department of Rhetoric and Writing Studies.

Responses from SDSU students were collected at the end of the Spring '97 semester within 22 developmental writing courses, 42 lower and

7 upper division writing courses, and one graduate course. The 47 total instructors of these various sections ranged in their background and training from first and second year TAs (26) to seasoned lecturers and tenure-track professors with multiple years of teaching experience. Forms were distributed to students at the end of a class session during the last two weeks of the semester and at the same time as the traditional evaluation form. All student respondents were anonymous.

Individual instructors of sections utilizing the new instrument were not anonymous: for the purposes of including statistics regarding individual teachers' professional status and experience, it was necessary to track the identity of each individual instructor. This tracking also made it possible to compare and isolate an individual's score on the revised and the traditional evaluation instruments. However, all participating instructors were assured that—for the purposes of any publication or administrative review of the results of the study—their identities would not be revealed.

The method for recording the data collected during instrument testing was as follows: values were entered for each student response to each question on the revised evaluation instrument, with "5" assigned to a "strongly agree" response and "1" to a "strongly disagree." (These values were reversed for the negatively worded questions, e.g. 14, 18, 23.) With respect to students' responses on traditional evaluation instruments used in the same courses, the only raw data available were SDSU instructors' composite scores on the institutional form used at that time. The composite score on this form reflects the overall mean of the mean scores for each of the ten Likert-style questions on the traditional form. (Appendix B itemizes these ten questions. Appendix C contains a sample printout of evaluation results distributed to instructors and administrators.)

As you'll recall, the primary goal of this study was to construct an instrument that would validly and reliably measure and reward effective student-centered teaching. Secondly, the evaluative instrument was intended to alter attitudes about what constitutes quality instruction. Results of McBride's and my test of the revised evaluation form indicate that it does indeed meet these two goals.

The particularly high Cronbach's alpha coefficient of the revised form (.92) speaks to its degree of reliability: compare it to, for instance, the .94 coefficient that Daly and Miller found for their Writing Apprehension Test (Lauer and Asher 137). Such a score suggests that the instrument's 27 questions have been well-conceived and that they assure respondents' interpretations of the single ability or achievement isolated by each question

as well as the inter-relatedness of the constructs being measured.

However, and as Lauer and Asher explain, “[w]hile the questions of reliability can be rather definitively answered,” “the question of validity can never be fully resolved” (141). We’ve seen that validity in assessment generally depends on the ability of the instrument to measure what it purports to measure and consequently on a careful consideration of what is being examined. Indeed, current instruments’ dubious capacity to measure teaching effectiveness—rather than teacher performance—is the problem which prompted construction of an alternative instrument in the first place. With respect to that revised instrument, the type of validity most at issue is “construct validity,” which is “the measure’s congruence with the theoretical concepts of a field” (Lauer and Asher 143). To that end, the questions on the new instrument are based on my understanding of the values and goals our discipline most commonly attributes to student-centered teaching. Because Lauer and Asher tell us that “validity depends in important ways on social consensus,” readers’ validation of my understanding is crucial (141). My exploration of the specific behaviors and conceptual habits likely to manifest student-centered teachings’ goals and values and my application of those concepts are meant to earn readers’ validation of these constructs.

Achieving the goal of privileging student-centered teaching methods depends in part on the revised instrument’s capacity to measure qualities different from those measured by the traditional instrument. The analysis of statistical significance of the difference in the scores generated by the two forms suggests that the notions of teaching effectiveness implicit in each form do indeed differ (cf. Appendix D, Figure 1). It appears that requiring students to rate particular aspects of their experiences in the course itself results in evaluations of teaching effectiveness less likely to center on teacher performance alone and more likely to foreground student-centered teaching.

However, I cannot be entirely certain that this difference is the one being measured. The lack of any traditional-evaluation-form raw data other than individual instructors’ composite scores by section precludes verifiable conclusions about the nature of the statistically significant difference being measured by the revised instrument and about any possible trends related to the significance of difference in individual instructor’s scores on the two instruments. Thus, comparative analysis can be done only on the basis of means, but not with respect to variance.

However, anecdotal evidence (based on my review of the list of individual instructors’ composite scores received on the traditional form and

on review of the total as well as the individual sections of the form) indicates that writing instructors are less likely to receive exceedingly low scores on my instrument. This conclusion is supported by a comparison of the range of mean scores generated by the revised instrument (3.2-4.4) to those generated by the traditional form (2.5-4.6) This informal evidence further suggests that the higher an individual teacher's score is on either form, the less likely it is that her scores on the two will differ. These possible trends, however, may in part or in whole be explained by the fact that many of the instructors who participated in this study are trained in and experienced with the techniques of student-centered teaching.

In addition to the goal of constructing a reliable and valid alternative instrument for measuring and rewarding student-centered teaching effectiveness, a proposed goal for the revised instrument was to alter students' attitudes about what constitutes quality instruction and about their responsibility in their own education. That long-range goal will require consistent and repeated use of the revised instrument so that students are exposed to and then internalize the characteristics of effective teaching valued by their teachers and by administrators (as those values are reflected in institutional evaluative practices). Nonetheless, the probability of the instrument's attaining that goal is suggested by the results of students' responses to three questions about the evaluation form itself. Those three questions asked students if the revised form was "significantly different from others used at this university," if it was "appropriate to this course," and if they "would like to use this evaluation form" in other classes. These figures, as well as the summary of mean and mode responses to each question about the evaluation form (cf. Appendix D, Figure 2), seem to indicate students' willingness to shift their attitudes about teaching effectiveness and their support for revising evaluative instruments.

It seems clear then, that this instrument is definitely a step in the right direction. Statistical verification of the internal reliability of the revised instrument and of the significant difference in its measurement from that of the traditional instrument certainly warrants further implementation and further development of this alternative method for assessing teaching effectiveness.

However, and as I explained above, a limitation of this study was its lack of raw data reflecting students' responses to individual items on traditional teaching evaluation forms (SDSU instructors' composite scores on that form were the only data available). Thus, comparative analysis of scores yielded by the two evaluative instruments could only include tests

using the mean of the mean student responses. Future tests of the instrument should collect more complete interval measurement data for students' responses to Likert questions on both, not just the revised, instruments.

In particular, analysis of variance, which measures variances in scores, could analyze the difference in students' responses to particular questions on the traditional as opposed to the revised instrument. More complete data on students' responses to individual items on traditional evaluations could also enable a multivariate analysis of variance, which in turn could help us assess the relationships between the two instruments' responses given to individual questions. Further, with regression we could predict if and when positive responses to the revised form's questions, such as "My teacher included student participation in deciding on the course goals, assignments, and/or reading material," are associated with negative responses on a traditional form question, such as one regarding the effectiveness of the instructor's selection of course materials. In short, more analysis of variance and regression techniques can help us better determine the nature of the relationship of the two instruments' measurements. Such analyses could thus ensure that the revised instrument does indeed reward effective student-centered teaching by generating high scores for the practices supporting those methods and lower scores for teaching practices centered more on teachers' performance.

In addition, further testing of the revised instrument should include instructors from a wider variety of departments and disciplines as well as a larger sample size. This expansion will increase the likelihood of diversity in teaching styles and training reflected in the instructors who participated and therefore improve the quality of our explanations for the statistically significant difference in scores generated by the two evaluation instruments.

Conclusion

WPAs' Roles in Revising Assessment Practices

Like most administrators in higher education, WPAs have been greatly involved in the systemic revision of general education curriculum. This institutional trend responds to public demands for "better" graduates and "more qualified" employees who can demonstrate critical thinking and collaborative working skills and to upper administration's call for teaching that will improve the retention and success of minority students. For decades, we in writing studies have been theorizing, developing, implementing into our own classrooms and departments the very curriculum and pedagogy that our colleagues across campus and around town now insist upon. Thus, and as many of us have already witnessed, our disciplinary knowledge and experience can facilitate our assuming leadership roles in curricular trends.

In this strategic position, we must not lose sight of the social mission informing our pedagogical theory and practice. Our politics are in our methods for evaluating and responding to what we and others are doing in the classroom. No matter what we may profess about the values and effectiveness of student-centered, liberatory, critical, radical, feminist and/or even banking pedagogies, our practices prove what we are for or against in the long run. We must recognize the habits and values we indoctrinate in our practices.

Like many other WPAs, Ed White focuses scholarly attention on the history and consequences of our practices. Within a profession notoriously resistant to change, he contextualizes the remarkable national and institutional shift in only one decade from the practice of multiple choice testing of writing proficiency to essay testing:

Something happened to the essay test on the way to the millennium, and I think it was a matter of power as well as naivete. By calling for essay testing and more recently for portfolio assessment, teachers have hoped to gain power over assessment and hence over the definition of what is to be valued in education. (292)

Clearly, that instance of writing teachers asserting power made a difference not just to the quality of and opportunity for students' education, but also to the ways that the institution, indeed the nation, conceives of instruction. In our current context, we are again presented with an

opportunity to change institutional practice, in this case the practice of relying on students' responses to teacher-centered performance as an accurate indicator of teaching effectiveness. Predicting the future politics of assessment, White gives us the hope that "we as a profession can affect the way our society defines and measures its goals, and the evaluation of writing has an important, perhaps central, role to play in that process" (297).

Likewise evaluation of teaching can and should play a central role in the way we define and measure educational goals. We have seen here a specific method for revising institutional practices that define and measure teaching effectiveness, a method WPAs are in a position to institute and refine. Because we in writing studies devote so much of our scholarly, creative, and classroom energies to improving the quality of instruction, we can and should play crucial leadership roles in the process of assessing instruction. If WPAs were to capitalize on this role, as well as on writing instructors' curricular and classroom skills and on the public's current attention to critical thinking and communication skills, perhaps their efforts would support effective student-centered teaching. As an essential result, WPAs' efforts could also be instrumental in reversing current practices' tendency to subvert the empowering and counter-hegemonic objectives of education for liberation.

Works Cited

- Bleich, David. "Evaluating the Teaching of Writing: Questions of Ideology." *Evaluating Teachers of Writing*. Ed. Christine A. Hult. Urbana: NCTE, 1994. 11-29.
- Hooks, bell. *Talking Back: Thinking Feminist, Thinking Black*. Boston, MA: South End Press, 1989.
- Hult, Christine A., ed. *Evaluating Teachers of Writing*. Urbana: NCTE, 1994.
- Kenway, Jane and Helen Modra. "Feminist Pedagogy and Emancipatory Possibilities." Ed. Carmen Luke and Jennifer Gore. *Feminisms and Critical Pedagogy*. New York: SUNY, 1992. 138-166.
- Lauer, Janice M. and J. William Asher. *Composition Research: Empirical Designs*. New York: Oxford UP, 1988.
- Luke, Carmen and Jennifer Gore, eds. *Feminisms and Critical Pedagogy*. New York: SUNY, 1992.
- Miller, Richard E. "Fault Lines in the Contact Zone." *College English* 56 (1994): 389-408.
- Pratt, Mary Louise. "Arts of the Contact Zone." *Profession* (1991): 33-40.

Shor, Ira. *When Students Have Power*. U Chicago Press, 1997.

Van Slyck, Phyllis. "Repositioning Ourselves in the Contact Zone." *College English* 59 (1997): 149-170.

White, Edward M. *Teaching and Assessing Writing*. 2nd ed. San Francisco: Jossey-Bass, 1994.

Appendix A
Revised Questionnaire For Evaluating Faculty

Students were instructed to circle the one response that best described their answer to each of the following questions. Available responses to each question were these:

strongly agree agree unsure disagree strongly disagree

Questions About the Student's Experience

The work I have done for this class has paid off in increased knowledge and understanding of the material.
This course improved my ability to express what I am thinking.
This course has value in my life outside of school.
This course required me to see concepts/arguments from multiple perspectives.
The course enabled me to realize how my thinking differs from the authors we read.
This course changed my mind about at least one opinion I had.
This course challenged me to think for myself.
This course required me to interact with my fellow students.
I value constructive criticism from my teacher and peers.
My fellow students were interested in the subject materials and assignments of this course.
My fellow students treated my teacher and me with respect.

Questions About The Teacher

My teacher clearly stated the goal or objective for the assignments of the course.
My teacher clearly explained how the individual units or readings related to each other.
My teacher is NOT familiar with my work.
My teacher gave me enough feedback to improve my work.
My teacher supported my taking risks in my thinking and writing.
My teacher encouraged me to develop opinions different from hers (or his).
My teacher did NOT listen and respond to students' questions, comments, and complaints.
My teacher included student participation in decisions about the course goals, assignments, and/or the reading materials of the course.
My teacher responded to signs of puzzlement, boredom, curiosity from the students.
My teacher presented and/or encouraged divergent viewpoints on topics or issues.
My teacher related the subject material of this class to students' interests and experiences.
My teacher did NOT treat students as adults.
My teacher used her authority in the classroom to create an environment conducive to learning.
My teacher was willing to admit her mistakes and/or insufficient knowledge.
My teacher respects constructive criticism.
My teacher's use of humor was positive and appropriate.

Questions About The Form

This evaluation form seems significantly different from others used at this university.
This evaluation is NOT appropriate to this course.
I would like to use this evaluation form in my other classes.

Appendix B
Traditional Faculty Evaluation Form

This particular "Evaluation of Faculty Questionnaire" was used in all San Diego State University Department of Rhetoric and Writing Studies courses during the Spring 1997 semester, the time when the study was conducted. Though students' narrative comments are maintained on file, few students respond in the "comments" sections; fewer still actually review those comments. While comments are made available when faculty submit their materials for annual review and/or promotion and tenure purposes, students' narrative remarks are not reflected in scores reported to instructors, departments, deans, and review committees.

Students were instructed to fill in a circle corresponding to the response that best described their answer to each of the following questions. They were advised to focus on the course rather than the instructor's personality. Available responses to each question were these:

1 Very low 2 Below average 3 Average 4 Above average 5 Outstanding

How would you evaluate this instructor's teaching?

Given the nature and objectives of the course, the instructor's organization appears to be . . .

The instructor's selection of reading and other outside class activities in relation to the objectives of the class was . . .

The instructor's statement of the specific objectives of the class was . . .

The evaluation process (e.g. tests, papers, etc.) in relation to the material covered in the class was . . .

Interest created by in-class activities, such as lectures, discussions, and group work was . . .

Clarity of assignment requirements and purpose of writing and reading assignments was . . .

The instructor's knowledge of the subject appears to be . . .

The instructor's ability to communicate the subject matter of the course clearly and effectively, whether by lecture, discussion or other means was . . .

As a personally enriching educational experience, this course has been . . .

What was most valuable about this course? What recommendations would you make for improvement?
(please comment)

Appendix C
Report Of Responses To Evaluation Of Faculty Questionnaire

In the interest of anonymity, this particular report represents scores for all instructors for the Department of Rhetoric and Writing Studies' course 92A during the Fall 1997 semester. Though no individual instructor's course is represented here, the report demonstrates the format of the report that an instructor receives and then references as evidence of her teaching effectiveness. The "score" the instructor reports is the mean score in the last line "Composite of Questions 1-10." In addition to this score, the only other items of significance to committees reviewing faculty's teaching effectiveness are the "department mean" score (on the same line) and the difference between the instructor's mean and the department's. In this particular report, for instance, we see that the mean score of instructors of 92A in Fall 1997 is higher than the mean score for all instructors of all courses within the department that semester.

**Course Evaluation Department of Rhetoric and Writing Studies,
 San Diego State University Fall 1997**

	Response Marked					Mean	Median No.	Resp.
	1	2	3	4	5			
1. This Instructor is knowledgeable about writing and the writing process.	2	7	57	239	396	4.46	4.61	701
	%	1%	8%	34%	56%			
2. The Instructor communicated the goals of this rhetoric and writing course.	4	14	80	240	363	4.35	4.53	701
	1%	2%	11%	34%	52%			
3. The course materials (texts, handouts, etc.) helped me achieve these course goals.	22	55	180	231	211	3.79	3.90	699
	3%	8%	26%	33%	30%			
4. The Instructor came to class prepared.	6	18	72	208	398	4.39	4.61	702
	1%	3%	10%	30%	57%			
5. The Instructor communicated the subject matter clearly and effectively.	8	33	105	225	331	4.20	4.41	701
	1%	5%	15%	32%	47%			
6. The assignments (reading, discussion and writing) helped me learn.	15	38	139	242	268	4.01	4.15	702
	2%	5%	20%	34%	38%			
7. The Instructor's comments on my work helped me to evaluate my progress and to improve my reading, writing, and thinking skills.	14	25	87	237	338	4.23	4.44	701
	2%	4%	12%	34%	48%			
8. The Instructor reliably met with the class, kept office hours, and kept scheduled appointments.	9	18	47	180	447	4.48	4.71	701
	1%	3%	7%	26%	64%			
9. The Instructor encouraged students to think, ask questions and express their views.	7	15	81	194	405	4.39	4.63	702
	1%	2%	12%	28%	58%			
10. Overall, how would you evaluate the Instructor's teaching?	7	16	89	243	348	4.29	4.48	703
	1%	2%	13%	35%	50%			
Composite of questions 1-10	94	238	937	2239	3505	4.26	4.49	7013
	1%	3%	13%	32%	50%			

Appendix D
Results Of Testing Of Evaluative Instrument

Analysis of the reliability of the revised evaluative instrument involved determining the internal consistency of the questions on the form. The Cronbach's alpha test, considered appropriate for assessing interval data such as the one through five responses students choose on the questionnaire, was applied to students' responses to the original 30 questions in five scales. After dropping the three items with the lowest reliability, I found a Cronbach's alpha coefficient of .92 for the remaining 27 items.

In order to determine the statistical significance of scores generated by the revised and the traditional evaluative instruments, a paired samples *t*-test was used. A comparison of the "Questions about the Student's Experience" section of the revised instrument to the traditional form yielded a *t* value of 6.47 ($p < .001$). The total revised instrument compared to the traditional form at a *t* value of 2.08 ($p < .05$). A comparison of the "Questions about the Teacher" section of the revised instrument to the traditional form was not statistically significant. These results are summarized in Figure 1 below.

Figure 1
T-test comparison of scores on traditional and alternative evaluative instruments

Evaluation Scores Compared	t	Sig. (2-tailed)
Revised form & traditional form	-2.07	.038
Revised form's "Questions about Student's Experience" section & traditional form	-6.46	.000
Revised form's "Questions about Teacher" section & traditional form	1.30	.193

We can predict the source of this difference in teachers' scores on the two different instruments by noting where it is greatest: no statistically significant difference was found in a comparison of the responses to the "Questions about the Teacher" section of the revised form with responses on the traditional form. However, a high significance level (.000 where $p < .05$) was found in the difference between responses to "Questions about the Student's Experience" and those on the traditional form. A significant difference (.038 where $p < .05$) was also found between total responses to each instrument.

Frequency statistics regarding the mean answer for each question in the "Questions about the Form" section and the valid percent of particular responses to each questions indicate that even though 70% of students agree or strongly agree that the revised instrument is significantly different from others, 77% of students agree or strongly agree that the form is appropriate to the course and 59% agree or strongly agree that they would like to use it in other courses.

Figure 2
Frequency Statistics for responses to questions about Revised Instrument

	Eval Diff	Eval Appro	Other Uses
Mean	3.81	3.91	3.57
Mode	4.00	4.00	4.00

