

WAC Directors and the Politics of Grading

Robert W. McEachern, Southern Connecticut State University

As director of my university's Writing Across the Curriculum program, I have had the opportunity to invite some of the field's best known scholars and program directors to co-facilitate faculty workshops. I have valued the insights they have shared with my school's faculty, but even more so, the advice they have shared with me about running a sometimes politically-charged program.

A couple of years ago, I had a conversation with one such scholar as we drove from her hotel to campus for the workshop. I told her about some of the work I had planned for the semester, including an invitation from my own department to facilitate a "norming" session for our sophomore literature course faculty. I would be responsible, I told her, for collecting and distributing sample papers, and then guiding the discussion that would help these twenty or so full-time faculty members reach some consensus on a grading scheme for the course.

"Hmm," she said, with a single, experienced syllable. "Wait until you're tenured."

Her comment underscores something that most WAC directors already know from running workshops for faculty: grading is a sensitive topic. It is also consistently the topic for which faculty at my school want additional workshops, and the topic that engenders the most discussion at our two-day workshops. Faculty are fine with grading, and appreciate topics like grading rubrics, which help them articulate evaluation criteria for students. But it is during those times when they are called upon to justify their grades (such as during the

workshops' group grading discussions and exercises) that they get a little uncomfortable.

In her forward to the collection *The Theory and Practice of Grading: Problems and Possibilities*, edited by Frances Zak and Christopher Weaver, Pat Belanoff discusses some of this discomfort among composition specialists, and urges them to "open up channels of communication with all segments of the public and with our colleagues in other disciplines and share with them what we know about grading and assessment" (xi). While not all WAC specialists are compositionists, we are (or should be) campus leaders in conversations about all aspects of teaching. I want to take up Belanoff's challenge and suggest a way for WAC directors to initiate and facilitate this necessary conversation about grading, by advocating we volunteer our services by leading grading-discussion sessions in our workshops, and norming sessions within departments in our schools.

Such sessions are not difficult to design and lead, but they are, as indicated by my workshop co-leader's comments, potentially sensitive. I will begin by discussing why we as faculty members are often reluctant to talk about and share our grades. I will then describe the norming session I led for my department, and review some lessons I learned from the process.

Faculty Grading: All Action, No Talk

As a number of contributors to Zak and Weaver's collection point out, few of us enjoy grading. By that, they don't mean the sometimes tedious process of reading and responding to student writers; rather, they mean the process of evaluating work and assigning to it a letter or number or some other symbol of its quality and success. Our dislike comes from a number of potential sources: our discomfort with the power we have over students' lives (Belanoff, "What"; Boyd); our understanding of the inadequacy of a symbol in summing up a student's knowledge and/or learning (Elbow); and our belief that grades make students complacent, more focused on the symbol than on our comments (Walvoord and Anderson). In short, we don't much enjoy label-

ing our students. However, what we enjoy less (and maybe what we fear most) is how our grading practices label *us*.

While these issues are important, especially if we are to facilitate conversations about grading with our colleagues, they do not get at the political issues surrounding the open discussion of grading. It is in the comparison of standards that my workshop participants get uncomfortable; very few complain about having to assign grades. But having their grades compared to others', even when given the opportunity to justify them, is a different matter. In my experience, there are a number of related reasons for faculty members' reluctance. Recognizing these reasons is necessary in helping to facilitate conversations on grading.

Fear of Grade Inflation

Richard Boyd, in his history of grading practices, shows that there has always been a "moral" dimension to grading; in fact, antebellum grading included an assessment of moral character (7–8). The moral dimension of grading, though, has shifted, so that the measure of morality is of the teacher, not the student. Few of us are worried that our colleagues will think that we are too hard on our students. We don't want to be seen as unreasonable, but we do want to be seen as upholding some kind of "standards." Exposing ourselves as soft on grading opens us up to scorn.

Perhaps the most prominent recent example of this phenomenon is taking place at Harvard University. After a newspaper reported that 91% of Harvard undergraduates had earned honors, faculty members approved changes to lessen grade inflation, including moving to a 4.0 scale and capping the percentage of students who may receive honors. The university's president, Lawrence Summers, praised the changes as an adoption of "higher standards." While one reason that grade inflation is a problem is its effects on students, just as problematic for Summers is its reflection on Harvard's teachers. His well-publicized feud with Cornel West of the university's African-American Studies department was at least partly about West's too-high grades. West left for

Princeton, and “standards” were upheld (“Harvard”).

Certainly, such arguments are made outside of the Ivy League, including at my own school, a comprehensive, regional public university. I recall being part of a weekly faculty discussion group my first year as a professor. During one session, members of the admissions department spoke with us about the university’s plans to increase enrollment by 2% per year. There was a vigorous argument against the practice by one veteran faculty member, who argued that we were already letting in too many students who didn’t belong in college to begin with. No one argued with him. I certainly wasn’t going to, as a rookie faculty member. It didn’t surprise me then that no one else did, given the kind of message any counterargument would potentially send: you’re soft on students.

Kathleen Blake Yancey and Brian Huot posit that “our grading patterns construct us. . . . The smarter the professor, the higher the standards, the harder the grader. In our departments, such faculty earn begrudging respect from others. Beyond our departments, our grading patterns are often used to warrant promotion and tenure: they provide a check against student evaluations that are a wee bit too high” (49). Given that kind of pressure, it is no surprise that faculty members are uncomfortable with discussing their grading practices.

Objectivity and Subjectivity

In the workshops I conduct for writing-intensive faculty, participants are anxious for some kind of “magic potion” for grading: a way of objectively measuring the worth of a piece of writing. Our grading-discussion sessions are not actually about coming to some kind of agreement about a grade, but as opportunities for the participants to explore their beliefs about what makes writing good, and where those beliefs have come from (previous experiences with teachers? disciplinary standards? perhaps some universal agreement?). What they come to realize (or admit that they know) is that there is no true objectivity in grading: to some extent, we each must make judgments about what is important.

One workshop topic that seems to help ease some fears about grading is a discussion of rubrics, detailed matrices of grading criteria and their values for an assignment or entire course. Thus, for example, “organization” may be worth 20 of 100 points for a paper assignment; perfect organization will yield 20 points, a lack of a transitions may result in 15, etc. (For excellent introductions to grading rubrics, see Goodrich; Popham.) Many instructors like the objective feel of rubrics; students can see exactly where and why they received the grade that they received. Of course, rubrics are not even close to objective: the criteria and their relative values are determined by the instructor. But the rhetorical effect of assigning a number to a piece of writing is very comforting.

That isn’t to say there is not some agreement about “good writing.” Belanoff, in an essay written for students called “What is a Grade?” argues that there is much subjectivity to grading writing, but there is some objectivity, or at least general agreement, among teachers of writing. Within certain discourse communities, Belanoff argues, there will be agreement about what counts as “good.” Certainly, there are some agreements that all of us can make about the worth of academic writing. But much of our disagreements comes from disciplinary standards, an issue that is obviously important to WAC directors who are interested in facilitating conversations about grading. Certain features of writing are easily judged “good” by certain disciplines, and “bad” by others. This idea of disciplinarity was underscored for me in a recent graduate class on the teaching of writing. One student, a former adjunct in the English department, was struggling with assigning grades to student papers. She eventually dropped the class, and didn’t come back to teaching writing. She explained to me, “I’m going to go back to teaching computer science. It’s [the grading is] easier. If the computer program works, it’s right. If it doesn’t, it’s wrong.”

Certainly, this instructor, and most writing-intensive instructors, understand that grading writing is not a black-and-white issue. And while they crave objective standards for grading writing, they know better than to expect them. But the desire for objectivity, and the

frustration at the lack of it, is another obstacle that must be overcome if a conversation is to take place.

Knowledge of Writing

A final and related issue is a problem as old as WAC itself. Art Young calls it one of “the enemies of WAC”: the belief by certain professors that, since they are not English teachers, they don’t know anything about writing and its teaching.

Unfortunately, sometimes they’re right. Anyone who has participated in a WAC workshop knows the kinds of myths about writing and its teaching that are ingrained in the heads of non-English faculty members (and in some within English as well). Young described and overturns these myths, and while there is some comfort in knowing that people outside one’s own school are experiencing the same kinds of resistance, that doesn’t solve the problem.

My own program made it a point to include in its guidelines for developing and proposing writing-intensive courses a statement that “reminded” faculty that they are the experts in the writing in their own fields, and I have repeated this reminder during workshops, whenever possible. With individual faculty members, I have pointed to research in writing-in-the-disciplines to back up my point (beginning with Bazerman’s and McLeod’s ideas of a “second stage” for WAC, and moving on to more specific studies of writing in particular disciplines). Still, whether used sincerely or as a convenient defense, many faculty, particularly those just beginning to think about how they can use writing in their classes, are reluctant to admit their own expertise with writing. General WAC workshops have been successful for me in getting them started on this recognition. But it is yet another obstacle to overcome.

The Politics of Norming

On my campus, there are few sustained, systematic, school-wide discussions about teaching. Our semi-annual WAC workshops and

follow-up “mini-workshops” are one such means of discussion. And even these are not “sustained”: given limited program resources, and limited time for an overworked faculty, most participants come to a two-day workshop, perhaps a mini-workshop once a year, but cannot commit to any more.

Still, commentators (Fulwiler; Soven) on WAC suggest that WAC directors are ideal campus figures in initiating and sustaining the kinds of conversations about teaching that Belanoff advocates about grading in particular. For me, those conversations have begun in the grading-discussion sessions we conduct in our two-day workshops.¹

During the workshop, I call this activity a norming session, but that is really a deliberate misnomer. Traditional norming or inter-rater reliability sessions require each participant to assign a grade to a paper and then discuss criteria until the group reaches consensus on a grade. But in the workshop, because the focus is on individual instructors, I am less interested in the grades they give than in the discussion that follows, when participants articulate for themselves (sometimes for the first time) their criteria for good writing. Participants are given three papers written in response to the same assignment. The writing will necessarily have some disciplinary slant, but is understandable to all, no matter the discipline. Participants read silently and grade each paper on the ETS scale of 1-6². We then post and compare grades and discuss particular features of the writing. At the end of the session, participants engage in some reflective writing about how they design

¹ I am indebted to Martha Townsend and Martha Patton of the University of Missouri for modeling a norming session for me at several of the workshops we held early in the life of our WAC program. While I have adapted the model to meet the needs of our faculty (at the Martyrs’ suggestion), the basic structure remains theirs.

² With the ETS scale, raters assign a whole number between 1 and 6 to each paper. The 1-3 grades are “below average,” and the 4-6 are “above average.” The idea is that participants are forced to make a choice; there are no “average” or “Gentleperson’s C” grades. I use this scale in our WAC workshops, and it is successful, though it takes some explaining (some participants, for example, want a 1 to be an F, a 2 to be a D, etc., until a 6 is an A+. Others will divide the scale into 100, so that a 1 is a zero, a 6 is a 100, a 5 is an 83.33, etc.). For an excellent discussion of the uses and value of this scale, see Edward White.

assignments and evaluate writing, and whether or not that is in line with the values they articulated during the exercise.

In evaluating the workshops, participants routinely point to these grading–discussion sessions as among their most valuable experiences. One reason is probably because of the individual focus: I don't want agreement among participants, and I try to find papers that will provoke some disagreement because of style, argument strategies, surface error, etc. Full agreement makes for lousy conversation. And conversation is the point of the exercise.

The trick has been to carry forth that conversation to other venues. Because the grading–discussion sessions are so popular, I have had several requests from faculty members to conduct similar norming sessions for their departments. I want to encourage WAC directors to be explicit about offering their services in this capacity. Of course, these sessions will likely be true norming or inter-rater reliability sessions, if my own experience holds: requests for my services have not been about getting faculty to consider their individual likes and dislikes about writing (after all, they can come to a two-day workshop for that). Rather, the requests have focused more on getting all faculty to agree on grading standards for some commonly-taught course or group of courses.

Such was my experience with my own department, English – the situation I was warned against by the veteran WAC scholar, as I described at the beginning of this essay. I was asked to lead a norming session of faculty who taught our sophomore literature course, a general education requirement for the university. Nearly all full-time members of the department teach this course, with a few exceptions (including me).

The session was similar to those I conduct in the two-day workshops. The difference was that the participants were trying to come to a broad consensus about grades. The chair provided an hour for the session, and she and I agreed that two papers were all that we could cover in that time. More than two would have provided a better range of writing. The papers' assignment prompt was to argue whether

Hamlet was sane or mad. The assignment was not one that anyone in the room had used before, but it was, they agreed, representative of the kind of argument that students in a general education, 200-level literature course would be asked to make. The participants were asked to grade each paper on the 1-6 ETS scale, and to make notes on why they would grade as they did (though not to provide full written comments as they would for a student in one of their classes). I then asked each to state their grades aloud, and recorded them on a chalk board for all to see.

The results surprised everyone: on the first paper, each participant gave a below-average grade of 1 or 2. On the second, each gave a below-average grade of 2 or 3. While I won't get into the statistical analysis, I can say that someone leading an inter-rater reliability session would have likely been very pleased at the result. I personally was further surprised at the low grades. I had chosen papers that I thought were, respectively, above and below average.

The conversation that took place after the grades were listed, during which participants justified their grades, also showed much agreement from participants. They spoke of common writing issues like lack of a clear thesis, organization problems, thin evidence, and an absence of direct quotes from the play. The agreement bears out the argument that Belanoff makes in "What is a Grade?": while there were some writing issues that were mentioned by only one participant, most that were mentioned would elicit nods and verbal agreements from the group. Though some grading is subjective, much is quasi-objective in that they are tacitly agreed upon by members of a discourse community, in this case, the group of literature faculty who teach 200-level courses at a particular school, though I wouldn't be surprised if the community extended beyond such a local community. I should note again that I do not teach this course, and thus am not part of this community, which is probably why I was surprised at the low grades that both papers received.

There are probably a number of reasons for the surprisingly agreeable outcome that have nothing to do with me (which should be an

encouragement to WAC personnel who are considering taking on the kind of leadership role that I am advocating here). However, I also believe that the norming session ran very smoothly because of some of the administrative steps that I took:

I understood how norming/inter-rater reliability sessions work. This seems intuitive, if not obvious. But there are small details and ways of behaving (when listing criteria for all to see, use the exact words of the participant; remain neutral but encouraging) that I became comfortable with because I had already run a number of grading-discussion sessions in another context. Anyone wishing to learn more about norming and writing, including advice for setting up such sessions, would be wise to read works by White; Willa Wolcott and Sue Legg; and Leo Ruth and Sandra Murphy.

I had the support of the department chair. This, too, seems obvious. But if an initial invitation comes from a department member, move quickly to discussions with the chair. She will have a better sense of the overall politics of the department, may suggest ways to handle difficult situations, and can generally provide advice. I made sure that all potential participants were given, in writing, a statement about the purpose of the session, a description of the activities, and the role of all participants, including the session leader. I also made an announcement of the session at the preceding department meeting, with the chair's blessing. I didn't want to spend too much time answering questions about the process or purpose, and I wanted any objections to be aired beforehand. At the session, I provided a brief reminder about how and why the session would work, and we were able to get right to business. One change I would insist on for the next session: make sure the chair reads and approves the sample papers.

I considered carefully the source of the papers to be discussed. In some ways, having actual papers from an actual class would be the best choice for this kind of activity, since they provide a realistic and accurate sense of the work of the students. There are, however, some potential pitfalls. The first is obtaining the papers; many faculty members are reluctant to show others their assignments and the kinds of responses

they get from them. Second, there is the danger that the assignment, and not the grades, becomes the issue. I expected, and heard, comments such as, “Well, before I give my grade, I just want to say that I never, ever would have assigned a paper like this.” My solution for that session was to provide real papers, but not from any of my colleagues. I found several papers on the same topic at free papermill sites *schoolsucks.com* and *bignerds.com*. They fit my criterion of being “real,” but didn’t present the kind of potential hurt-feelings and political problems that internal papers would have presented. I devised an assignment myself, retroactively, that the sample papers fit. In fact, the “I wouldn’t have assigned this” issue was productive: with no one to get defensive over the assignment (which asked students to argue whether Hamlet was mad or sane, a fairly popular topic in the free papermills), the session expanded into a productive discussion of assignment topics.³

I considered carefully the grading scale to be used. The consensus reached in my session might call into question the value of the ETS scale. The participants told me afterward that a more traditional A–F scale would have allowed them to more easily compare grades. But I was concerned about the possibility of getting bogged down in discussions about the difference between a B and a B+, for example. The chair and I had discussed the possibility of continuing the discussion, at which time we could have refined the grading scale. (The follow-up session never materialized, since there was a general consensus among the participants.) However, I could see a situation in which the unfamiliarity of the ETS scale would be a help. Perhaps in a politically-

³ There are, no doubt, readers who would consider my using such papers to be ethically suspect. That is a subject for another essay (one which I have already begun). But I will say that, at the beginning of the session, I announced that the papers were real, but were not written for a class at the university (though they may very well have been), and that at the end of the session, I would reveal the source, which I did do. I also revealed that the assignment was made up retroactively. I will grant that I had doubts myself about the ethics of the practice, but decided that eventual full disclosure was the right thing to do, and that no actual grades were given, and thus my conscience was clear.

charged department, the difference between a 4 and a 5 may seem smaller than the difference between a B and an A-.

Conclusion: Onward into the Conversation

Ultimately, my workshop co-leader's fears were unfounded. The department did not have the knock-down, drag-out fight that some had predicted. And, I'm happy to report, I was granted tenure the following spring.

Still, not all situations will be so smooth. I credit my departmental colleagues for being reflective teachers who are open to discussion and willing to learn from one another. That might not be the case in all departments, for a variety of reasons.

Despite the success of this session, though, I must confess some failures: While I was invited to conduct three norming sessions for other departments, none of them actually took place. One in a health sciences department and one in a humanities department involved an initial invitation, but never materialized. Another in a social science was scheduled, and then cancelled because of an emergency department meeting, and was never rescheduled. Some of the problems resulted from the kind of poor planning that I give myself credit for avoiding above: not having the support of the department chair, for example. But I think for the most part, the departments backed away because of objections from their members. I'd heard whispered rumors from some department members about fears of exposure of grade inflation, paranoia about lack of knowledge of writing expertise, etc.

Those failures, though, should not dissuade WAC directors from initiating the conversations. If Belanoff is correct, and my experience is generalizable, then there will likely be more agreement than not, once a session is conducted. The difficult part is getting the members of departments to agree that the conversation should happen in the first place. The costs are far outweighed by the benefits: increased exposure for the program, and a strengthening of the mission of WAC—to engage teachers in discussions about their teaching.

Works Cited

- Bazerman, Charles. "The Second Stage in Writing Across the Curriculum." *College English* 53 (1991): 209-22.
- Belanoff, Pat. "Forward." *The Theory and Practice of Grading Writing: Problems and Possibilities*. Ed. Frances Zak and Christopher C. Weaver. Albany: SUNY Press, 1998. ix-xii.
- . "What is a Grade?" *The Subject is Writing*. Ed. Wendy Bishop. Portsmouth, NH: Boynton/Cook, 1993.
- Boyd, Maylath. "The Origins and Evolution of Grading Student Writing: Pedagogical Implications and Cultural Anxieties." *The Theory and Practice of Grading Writing: Problems and Possibilities*. Ed. Frances Zak and Christopher C. Weaver. Albany: SUNY Press, 1998. 3-16.
- Elbow, Peter. "Ranking, Evaluating, and Liking: Sorting Out Three Forms of Judgement." *College English* 55 (1993): 187-206.
- Fulwiler, Toby. "The Quiet and Insistent Revolution: Writing Across the Curriculum." *The Politics of Writing Instruction: Postsecondary*. Ed. Richard Bullock and John Trimbur. Portsmouth, NH: Boynton/Cook, 1991. 179-188.
- Goodrich, Heidi. "Understanding Rubrics." *Educational Leadership* 54:4 (January 1997): 14-18.
- "Harvard Raises the Bar to Curtail Grade Inflation." *Chronicle of Higher Education* (7 June 2002): A39.
- McLeod, Susan H. "Writing in the Disciplines: The Second Stage and Beyond." *College Composition and Communication* 40 (1989): 337-42.
- Popham, James W. "What's Wrong—and What's Right—with Rubrics." *Educational Leadership* 55:2 (October 1997): 72-77.
- Ruth, Leo and Sandra Murphy. *Designing Writing Tasks for the Assessment of Writing*. Norwood, NJ: Ablex, 1988.
- Soven, Margot. "Conclusion: Sustaining Writing Across the Curriculum Programs." *Writing Across the Curriculum: A Guide to Developing Programs*. Ed. Susan H. McLeod and Margot Soven. Newbury Park: Sage, 1992. 189-197.

- Walvoord, Barbara and Virginia Johnson Anderson. *Effective Grading: A Tool for Learning and Assessment*. San Francisco: Jossey-Bass, 1998.
- White, Edward M. *Teaching and Assessing Writing: Recent Advances in Understanding, Evaluating, and Improving Student Performance*. 2nd Ed. San Francisco: Jossey-Bass, 1994.
- Wolcott, Willa with Susan M. Legg. *An Overview of Writing Assessment*. Urbana, IL: NCTE, 1998.
- Yancey, Kathleen Blake and Brian Huot. "Construction, Deconstruction, and (Over)Determination: A Foucaultian Analysis of Grades." *The Theory and Practice of Grading Writing: Problems and Possibilities*. Ed. Frances Zak and Christopher C. Weaver. Albany: SUNY Press, 1998. 39-52.
- Young, Art. "The Wonder of Writing Across the Curriculum." *Language and Learning Across the Disciplines* 1:1 (Jan 1994) 58-71.
- Zak, Frances and Christopher C. Weaver, eds. *The Theory and Practice of Grading Writing: Problems and Possibilities*. Albany: SUNY Press, 1998.