

“But I Have to Have an A”: Probing the Cultural Meanings and Ethical Dilemmas of Grades in Teacher Education

By Margaret Placier

This is one chapter in a story of learning to teach in a college of education. It is also part of an ongoing collaborative study of the socialization of beginning teacher educators on four campuses. For three years, four of us have been sharing journals and conducting inquiries on aspects of those experiences that intrigue or trouble us. After three years, grading students was the aspect of my teaching that continued to concern me most. It was the source of almost all the conflicts with students I had documented in my journals and field notes over this period. Therefore, using a combination of participant observation and action research methods, I directed my inquiries for 1992-93 toward coming to a better understanding of grades and to improving my grading practices.

Margaret Placier is an assistant professor of education in the College of Education at the University of Missouri-Columbia, Columbia, Missouri.

My graduate education did not explicitly teach me anything about college teaching, much less the specifics of student evaluation and grading. Looking back, I find this omission strange in a college of education. I do not recall ever talking about college teaching with professors or student colleagues, even those who were teaching assistants for undergraduate classes.

“But I have to Have an A”

As a first year professor, not once did I consult the research on college teaching; in fact, it never occurred to me to do so. I also did not consult my senior colleagues at any length about their grading practices. There were no orientations or handbooks spelling out the “rules” for grading. My impression was that faculty had near-total autonomy to decide such things for themselves. It was only as I identified problems in my teaching as part of our collaborative study, and had to write about those problems for scholarly conference presentations, that I consulted the research on college grading and began talking with colleagues about their practices. The following review of the literature, then, is knowledge that I did not have (but wish I had had) at the time I collected the data.

Research on College Grading Practices

Milton, Pollio, and Eison (1986) have written the most complete review of college grading. According to these authors, belief in the “virtues and meaningfulness” of the A-B-C-D-F grading system is firmly rooted in the culture of academe in the United States (p. 1). Yet, a grade has little meaning outside the contexts of an historical era, disciplinary and institutional cultures, and/or the judgments of individual faculty. A grade is an inadequate, ultimately-reduced description of several weeks of performance, usually from a single instructor’s perspective. The GPA, made up of many grades assigned in different contexts, is meaningless without an attached transcript showing the student’s major and courses taken. Even then, transcripts from different institutions cannot be compared with certainty, although graduate school admissions committees and employers do make such comparisons regularly.

Grading has generally served more administrative than pedagogical purposes. In the historical records of United States colleges and universities that Milton *et al.* examined, the purpose of grades was almost always to rank students for sorting or gatekeeping functions, **not** to promote learning. Emphasis on grading waxes and wanes with the emphasis on scientific management, control of student enrollment, and trends in educational philosophy. In the 1960s and 1970s—a liberal, expansionist period—there was widespread experimentation with alternatives such as pass/fail, mastery learning, and grade contracts (Simon & Ballanca, 1976). These experiments, defended at the time on both political and pedagogical grounds, usually did not survive. In fact, some critics fault these experiments for contributing to “grade inflation,” an indicator of supposed educational decline (Goldman, 1985; Weller, 1986). With reference to Plato, Goldman (1985) argued that sorting **is** the social function of higher education, and to perform that function ineffectively by giving too many high grades, or refusing to grade at all, is unethical.

Today, in the wake of educational reforms of the 1980s stressing higher standards and increased accountability, there is more stress than ever on grades, particularly in colleges of education. For example, *A Nation at Risk* (National

Commission on Excellence in Education, 1983) claimed that many teachers came from the bottom quarter of their college classes. In response, many colleges of education now require higher grade point averages [GPAs] for student entry and retention.

What are the effects of grading policies on students and faculty? In *Making the Grade*, Becker, Geer, and Hughes (1968) reported a qualitative sociological study of college student perspectives on grades. According to these authors, most students share a “grade-point average perspective.” They must negotiate an environment that emphasizes the GPA as the basis for decisions about their lives, and grades as the “currency” of their academic economy. They are in a subordinate relationship with faculty and administrators who make these decisions and who distribute the rewards. Becker *et al.* contend that faculty are unrealistic to expect most students to do academic work out of sheer joy. They should realize that in this situation most students will devote considerable time to devising strategies for making grades. Rabow and Hernandez (1988) found that Becker *et al.*'s findings still held true in the late 1980s, that the GPA perspective persisted through the experiments of the 1970s.

A series of psychological studies has categorized college students as primarily GO (grade oriented) or LO (learning oriented) (Janzow & Eison, 1990; Milton *et al.*, 1986.). GO and LO students have different expectations of classes, and they evaluate their instructors differently. Faculty may also be categorized as GO or LO, and conflicts in orientations between students and faculty may lead to dissatisfaction and miscommunication. Moreover, because faculty disagree on the philosophical or psychological bases of grading, policy decisions about grades are fraught with conflict (Janzow & Eison, 1990).

The theoretical framework for LO/GO studies derives from individual psychology, while Becker *et al.*'s sociological framework focuses on collective beliefs and behaviors. Yet the conclusions are similar. Psychologist Lowman (1990) pointed out that the LO/GO dichotomy is very similar to extrinsic/intrinsic motivation, and argues that instructors who emphasize grades and their power to grade students encourage external motivation. He interpreted this as a negative personality characteristic, to be discouraged through strategies such as reducing the emphasis on grades. Milton *et al.* (1986) have concurred.

This is also the point reached by Becker *et al.* (1968): if faculty want to create a learning-oriented student culture, they must change grading policies and perhaps even abolish grades. However, they note that this position may not be supported by students who succeed within the current system. In fact, most students probably bring a grade-orientation with them to college (Rabow & Hernandez, 1988). Jackson (1968) and others have argued that socializing children to conform to external demands is a fundamental **aim** of schools in the United States. Can higher educators undo the deep socialization of our students?

Further, the assumption that grades are negative is not shared by all faculty.

“But I have to Have an A”

Pollio and Humphreys (1988) contend that many professors believe good grades must be significant because **they** received good grades as students. Faculty may even believe that student performance reflects innate ability (Becker *et al.*, 1968; Rau & Baker, 1989). In Fong’s (1987) view, though faculty may at times lament the “evils” of grades, a rejection of grades would be too drastic a break with expectations and conventions of schooling. According to McDermott (1987), American schools are rooted in the necessity of comparing students, of measuring success and failure. Would higher educators challenge the basis of their own success? Would they challenge the deep culture of American education?

The LO/GO studies have particular relevance for teacher educators. If colleges of education reinforce attitudes that correlate with grade-orientation or external motivation, we may be incapable of preparing so-called reflective practitioners or critical intellectuals. Studies have found that teachers in the United States believe they have lost substantial control over their work in recent years, as a result of the intensified bureaucratization of schools (Carlson, 1992; Frymier, 1987; Johnson, 1990). Was this disempowerment promoted by teacher education programs based on external motivation? In studies of teacher education students, Korthagen (1988) described them as either externally- or internally-oriented. He was concerned, however, that innovative programs favoring internally-oriented students may lose externally-oriented students, who do not identify with the new emphasis on reflection and critical thinking. The change from an external to internal orientation, Korthagen says, may be slow.

It is particularly disturbing to consider how policy decisions about grades affect a college of education, which many believe should be primarily concerned with preparing teachers who are learning-oriented in the broadest sense and who, in turn, will nurture a learning orientation in their students. Moreover, the undemocratic faculty/student relationships reinforced by the GPA perspective cannot be a model for teaching students about democratic schooling. Through our grading policies, we may create a contradictory culture that undermines our most important purposes.

In short, conflicts about grades reflect fundamental conflicts about the individual, culture, society, the purposes of education, and the instructor’s role in socializing students. The research on college grading raises fundamental issues but it cannot provide an instructor with a definitive guide to grading decisions, which will depend upon personal beliefs (Frisbie & Waltman, 1992) and upon the pressures exerted by the political and institutional context.

Political, Cultural, and Ethical Contexts of My Grading Decisions

If, as Milton *et al.* (1986) have argued, grades are context-bound, it is important to describe the context in which my grading decisions are made, based on field notes

and documents collected over three years. I teach in a college of education on the main campus of a state university with relatively high entry requirements and tuition. Under pressure from the Board of Curators and State Coordinating Board for Higher Education, our campus recently became even more selective in enrollment and retention of students. Faculty are under pressure to adopt tougher grade distribution standards. In fall 1992, data on percentages of As and Bs assigned in all programs were published as a basis for criticizing "grade inflation" in some programs.

Education majors must maintain a 2.75 GPA (on a 4-point scale, a low B) to retain "professional standing" in teacher education. This requirement, along with figures on our students' standardized test scores, are evidence that our college has students who compare favorably with those in other colleges. Each semester, faculty receive lists of students who are on probation or ineligible to register because their GPAs have dropped below 2.75 (even if only by .001). It is clear that our grading practices have gatekeeping consequences, and specifically that low grades remove students from our program.

Moreover, a C is a "low grade." If education professors made C the average grade (as traditionally defined) in our classes, we would soon have few students to teach. The meaning of "C" has changed in this policy context. While Goldman (1985) deplored such "grade inflation" in response to policy mandates and job insecurity, he did not take into account that the criterion for "failure" has shifted upward. I can now perform my "gatekeeping function" by assigning Cs, not Ds or Fs. Milton *et al.* (1986) argued that, given the meaninglessness of the GPA, carrying it to two decimal places, especially as a basis for decision-making, is indefensible. However, it has been politically defensible as a strategy for improving the historically negative image of the college of education. Summerville, Ridley, and Maris (1990) found that education deans claimed that their students' higher GPAs reflected student quality (although the data did not confirm this). I do not know if my students are "better" than the ones before the higher GPA rule went into effect, but I am skeptical.

I teach required Foundations of Education courses to juniors, seniors, and masters-level education majors, teamed with another colleague and one or two teaching assistants. The courses examine philosophical, historical, political, and sociological aspects of education. In my first two years, I was not involved in team-teaching and was uncertain about how to evaluate students' knowledge of such topics. The use of the A-B-C-D-F grading system is taken for granted on my campus, and I adopted it initially without question. But being unsure about my goals and standards, and developing rather "chummy" relationships with students, I was extremely generous in my grading. As Natriello and Dornbusch (1984) might put it, I had not developed an optimum balance of "warmth" and "standards." In fact, when the Curators' list was published, "Foundations" was listed as assigning a relatively high percentage of As and Bs.

“But I have to Have an A”

Over time, through practice and collaboration with colleagues, I have become more certain that I can compare students' work and be more selective about assigning As. At the same time, my teaching evaluations have gone from very positive to mediocre, even though Gleason's (1986) reading of the research questioned this relationship. I have also had more conflicts with students about grades. At times, I am tempted to revert to my initial leniency. Johnson and Beck (1988) divided students into classes that were “strictly” and “leniently” graded. Students in the lenient classes were more responsive in class discussions, complained less about unfair tests and grading, and gave their instructors better teaching evaluations. These findings are consistent with my personal experiences. But “strict” grading was more likely to motivate “low ability” students (labeled on the basis of SAT scores) to achieve. Strict grading, therefore, seems to be the more ethical and more pedagogically sound strategy. But then the authors theorize that the low ability students probably worked harder because they were more grade-oriented, and strict grading was an external motivator. The dilemma returns in full force.

The following sections report my findings on the grading dilemma in three areas: everyday manifestations of grade orientation, two “fiascoes” or critical incidents that brought students' beliefs about grades to the surface, and student beliefs about grades.

Everyday Manifestations of “Grade-Orientation”

Analyzing my field notes in light of the research, it is clear that many of my students exemplify the attitudes described by Becker *et al.* (1968) and Milton *et al.* (1986). For example, most students can readily cite their GPAs to two or even three decimal places. They consider a C disastrous because of its effects on the GPA. Even a B (3.0) is not far enough above the minimum to make a difference to some. Despite my relatively lenient grading, too often my interactions with students concern grades. I have often questioned why my most intense discussions with students, in and out of class, concern grades rather than course content. Now I know that this is an everyday part of college culture (Becker *et al.*, 1968; Eble, 1976; Jedrey, 1984).

Some students come in as soon as they sense that they may not receive the grade they want. Their message is “I have to have an A in this course. What can I do?” Students can actually find themselves in a situation in which they must receive As in **all** their courses in one semester in order to salvage their GPAs. But since their performance in my class may be typical of the performance in other classes that brought them to this desperate situation, I am at a loss to suggest any simple strategies for “getting an A.” After discussions with such students, I ask myself questions such as the following: “How much one-on-one coaching should we provide?” “Do we encourage the students' grade-orientation by focusing on graded assignments?” “Should we refocus students on the everyday reading, critical

analysis, questioning, discussion, etc., that lead up to success on those assignments?" All this is assuming, perhaps naively, that the student is genuinely asking for help. Becker *et al.* (1968) found that some students visit professors to draw positive attention to themselves, in the hope that this alone will influence their grades.

Others take a more legalistic approach, appealing the grading of a particular assignment. I now ask them to put these appeals in writing, rather than expecting me to decide immediately based on their oral arguments. Or they may propose mitigating circumstances that account for their poor performance. For example, one student argued that an in-class short essay exam was really a "time test," and as a slow writer, he was unable to finish in the time available. A student with very large handwriting said she was penalized for lack of content, because I had not allowed enough space for her answer. Another student had a genuine learning disability, but had not informed me at the start of the class. Death, divorce, illness—am I also to take them into account?

Other students wait until after final grades are issued. If their grade has had negative effects on the GPA, they try to negotiate grade changes. These are the most difficult interactions, because some students feel they must invalidate a teacher's grading decisions to force a change. They may be angry, tearful, or both. I check the ineligible list after such visits and often find their names; this immediately explains the strong emotions. The gate to a teaching career has just slammed in their faces, and I was one of the gatekeepers. Some take a negotiative approach. They want to know if I will accept extra credit assignments in exchange for a higher grade. Of course, that is both extra work for me and potentially unfair to less assertive students (Jedrey, 1984).

Statham, Richardson, and Cook (1991) reported that both male and female professors encounter such situations, but female professors try to avoid them through "bolstering strategies," which are elaborate explanations of grading practices. Their research helped me understand why I, and not my male co-instructor, was the one who developed the specific grading guidelines listed on the syllabus—the very explicit directions for assignments and the checklists of grading criteria. Given women's tendency to "personalize" teaching more than men, these authors suggest that we try to prevent situations in which evaluation and personal attachments will be entwined. They speculate that women may also find the emotional intensity of interactions over grades to be threatening. Certainly, I have been quite shaken by some encounters over grading. Although I have tried all the prevention strategies recommended by Pulich (1983), such encounters continue to occur.

Two Grading Fiascoes

Elbow (1983) vividly describes the "contraries" of the college teaching role: professors have obligations to students, but also to society and to their disciplines.

“But I have to Have an A”

He proposes that the effective professor finds a way to be **both** nurturer and gatekeeper. I am wary of the gatekeeper role because gatekeepers in education are notorious for sorting **certain** students out based on class, race, language, gender, disability, and similar factors. Sociologists Rau and Baker (1989) urge college faculty to consider that their teaching practices perpetuate inequalities among students (p. 172). My first grading fiasco shows how an experiment with grading, to find out which students I was sorting out and why, placed me in conflict with the student culture.

Fiasco #1: The Bottom of the Curve

In my first year of team-teaching, I worked with a senior colleague who gave multiple-choice examinations. He thought I was wasting my time, as an untenured instructor, grading essay exams. Although I did not tell him this, I thought that using multiple-choice examinations to evaluate students' knowledge of the philosophy of education was a travesty. Rather than mediating our differences, we divided the course into two separate eight-week sections.

When this colleague left suddenly before the start of the next school year, I was faced with the entire Foundations teaching load. We hired one of his advanced Ph.D. students to take his place temporarily, and she adopted the same examination model as her mentor. I became convinced that, under the circumstances, my senior colleague may have been right: if I were ever to achieve tenure, I should give up my ideal of essay evaluations.

My first objective exam seemed quite successful. The scores fell into a lovely curve. Sixteen (out of 120) received Ds and Fs, according to a standard percentage distribution (90-100=A). Before handing back the exams, I had a teaching brainstorm: Why not find out more about those people and try to discern why they failed? I decided to offer them the option of talking with me about their performance and re-taking the exam.

The TA wrote the grade distribution on the chalkboard, and I told the class I was happy, because someone had actually scored 100 and there were many As and Bs. But I was worried about the people at the bottom, the D and F people. I wanted to talk with them about improving their performance. Perhaps their experiences could teach me how to better design this type of examination, since this was my first attempt.

At this point the class erupted. “Why should I help people who were too lazy to **work**? They had probably been in a bar the night before the exam.” One student said that she would be lucky to get a C out of the class, and **she** would not get extra help. “It wasn’t fair.” Many students were talking and scowling. I argued back: “How would I know who those people were unless I found out?” “How could I assume they had not worked? How could they?” “Is it that simple, that students who **work** get high grades? Is that the **only** factor in achievement?” “Every teacher should be concerned about low-achieving students,” I said. “That was one of the

points of the course.” My argument was not working.

Between classes, a visibly angry student came up to me and complained about her grade. I pointed out (much more harshly than usual) that she had received a B, and a B was a good grade. By that point, I was physically drained and emotionally spent. This was my first open conflict with students. **What was happening to me?**

When the next section met, they had already heard the news. One student confronted me: “Are you really going to give the people who got bad grades another chance?” The same arguments came up, but not in such an aggressive way. I told the students that I had never seen such anger from students, and that this showed that grades were more important to them than learning. An A student in the front of the class said, “If this doesn’t affect us [those with good grades] why should we care?” I replied that I had no idea, and the discussion ended.

Later that day I received a note in my mailbox from two students, arguing that helping people with low grades was unfair to those who had worked hard and received Bs and Cs. Maybe they wanted a chance to make an A. I wrote back, explaining my reasoning, which I readily admit had not been well-developed at the time of the decision. One of the note-writers visited my office. She said she was afraid I would be angry with her about the note. We had a fairly friendly talk. The other note-writer came to me in class, thanked me for my reply, and said she wanted me to know that she was not “mad at me.” They were both having second thoughts about having disagreed with a professor.

My co-instructor said that students in her discussion group were mad at me, but they knew students in my group who “liked” me, so I might not be all bad. A student told me that everyone was discussing my “controversial” decision outside of class. It seemed that I had become notorious as a result of one well-intended, spur-of-the-moment decision that affected the grades of a small number of students. The incident brought home to me the **intensity of students’ attitudes about grades**, and the meanings they assign to grades (as reflections of “hard work”). If I had read Becker *et al.* (1968), I would not have been surprised.

In an attempt to learn from this “teaching disaster,” I turned this first grading fiasco into a research project. When the D and F students came into my office, I interviewed them about what had led up to their performance on the exam, including their previous educational experiences. Frankly, they seemed indistinguishable from the rest of the class. None attributed their poor performance to gender or class (all were white), or critiqued the system. They were prepared to accept a D or F without complaint. Most of them attributed their grades to individual effort. They honestly said they did not know the answers due to inadequate studying. Some had had several examinations in one week and had short-changed their effort on this one, or had run out of time and energy by the time it was given at the end of the week. These comments expressed the belief that students control their grades through the effort they expend (Becker *et al.*, 1968). In contrast, others thought that they had studied just as “hard” as friends who had received higher grades, and they had no

“But I have to Have an A”

idea what had gone wrong. A few critiqued the wording of items as confusing, or said they had trouble with multiple choice tests or testing situations in general.

For extra credit, I asked them to annotate the items they had missed, explaining why they had chosen the wrong answer and why the correct one was correct (if they agreed it was). The most interesting comments were about multiple-choice items for which they had chosen the answer they thought **should** be correct, not the one that **was** correct according to the authors in the text. This was despite the fact that the items were worded in the form, “According to [specific author]...”

These comments presented a student critique of the readings that the “objective” exam had not invited. Granted, it is naive for upper division college students to assume that a multiple choice examination could be based on their choice of the “best” answer on some normative basis. But these responses showed that this kind of thinking, of which even the lowest-scoring students had been capable, was not elicited by the examination. After realizing this, and listening to students describe their struggles to “cram in the material,” I determined never to use multiple choice exams again. But this is self-sacrificing, because the reward system does not account for more time-consuming methods of student evaluation (cf. Rau & Baker, 1989).

Interestingly, all the D/F students improved their grades on the final exam. But the other students soon had “their revenge” on me for breaking the grading rules by “rewarding” the “lazy” people at the bottom. My teaching evaluations plummeted, and some comments referred back to the incident. A colleague nominated me for a teaching award the following semester, but I knew that this was futile because I would have to submit my teaching evaluations to the awards committee.

Fiasco #2: Grade Appeals

The year after Fiasco #1, we hired a new Foundations faculty member. We worked as a team with two teaching assistants. The grade for our course was based on three or four “take-home” essays that required students to develop positions on fundamental conflicts in U.S. education. A new source of tension under this arrangement was student complaints that grading was “too subjective.” They also complained that essays were evaluated differently by different readers. Some said that they were simply not “good writers” and therefore had no chance for an A. These are reasonable complaints that reflect beliefs in fairness and equal treatment. The following are illustrative:

I have had three different grades from three different people in this class. I put nearly the same amount of time in each paper and do not understand how the grades could fluctuate so greatly.

The grading system was terrible. Different people grading everyone’s paper—how consistent and fair can that be? If our entire grade is to be based on papers, we should have either a review process or a better system of grading.

Unreliability in marking of essays has long been cited as their major drawback. According to Beard and Hartley (1984), several strategies can be applied to make essay marking more reliable: (1) coming to firmer agreement among graders about standards to be applied; (2) limiting or specifying essay topics very carefully; and (3) having more than one grader for each paper. We had tried the first two strategies, becoming increasingly specific about evaluation criteria. The campus writing program provided advice on grading rubrics. Yet the complaints persisted. It was our attempt to apply the third strategy that led to Fiasco #2.

In a conversation about evaluating students' writing, a colleague from the School of Law told me that he asks students who are dissatisfied with their grades to write appeals stating their specific objections to his decisions. He thought that requiring students to put their objections in writing probably discouraged all but the most serious complaints and decreased the number of office visits to haggle over grades. We decided to try this. On returning the mid-term essay, we informed students that in the spirit of fairness, since only one person had graded each essay and we did not have time to double-grade every one, we would accept appeals or requests for a second grader.

My law school colleague reported that only rarely did students file appeals, but we underestimated the differences between our two contexts. In most U. S. law schools, students are socialized to be intimidated by their professors and to accept severe and sometimes arbitrary grading practices. For instance, in some classes in our law school, only one A is awarded, regardless of the quality of the work or potential of the students. Law students also know that the appeal will be heard by the same "judge" who issued the original opinion, so that their chances of reversal are slim. We promised that a different grader would look at the appeal, to respond to the charge of unreliability. We received 47 grade appeals!

Appellants were required to write their arguments in favor of a grade change and to attach their written appeal to their original, marked paper. In most cases, I upheld the decisions of the first grader in the section I reviewed. In contrast, "Jim," my co-instructor, made many changes in the grades assigned by one of the teaching assistants. For example, the teaching assistant had given one student a C for what Jim thought was an outstanding, original paper. Her decision had apparently been based on a very rigid interpretation of the assignment. Jim agonized over his stack of appeals for days, and said it was "depressing" to read the papers again and to read the students' often weak justifications for a grade change. Some of their justifications were as follows:

Grades did not reflect effort. "I put a great deal of time and effort towards this assignment, which I do not mind doing, yet I was not awarded the grade I feel I deserve... I could not have worked harder on this paper. Yet I have friends who threw their paper together the night before, some of whom used my class notes because they do not attend, and were awarded the same grade as me. I think something is wrong with this system when hard work doesn't pay off."

“But I have to Have an A”

Grades were not based on our own criteria. The paper, in the student’s opinion, met all the criteria for an A paper. Some gave detailed analyses of their papers, point by point in comparison with the grading criteria. In other words, we had either not applied our own standards, according to these students, or our standards were subject to interpretation.

The criteria emphasized form as much as content. “In the criticisms I didn’t read about how my paper didn’t address certain key issues, or how it wasn’t persuasive. Rather, I read in this that my paper had grammatical mistakes or obvious spelling errors...Comments concerning technical errors vastly outnumbered the comments I received concerning subject matter.” Some students did not like graders’ attempts to revise the wording of their essays. “I wrote ‘one’s ability to rationalize’ and it was changed to ‘one’s ability to reason.’ I see no difference. I should be able to use any choice of wording I want as long as the meaning is clear.” “I feel like my writing style is being attacked. English is something that is a personal attribute such as an artistic talent.... This essay is very logical to me.... Getting off track in a paper is to be expected a little bit.”

Grades did not reflect the instructor’s teaching. “I state that Dewey believes the child should do what they want. I derived this from my notes (9-16-92) where I have written that Dewey believes we should ask the child what he wants and teach that.” Sometimes the student had shown an instructor or TA a rough draft and was given the impression that the draft was an A paper, or could be if certain revisions were made. “When I took my rough draft to Jim, he told me that I have a strong argument with good supporting quotes and evidence. He said that to make my paper even stronger, I should include a brief paragraph about the possibilities that the two views were contradictory. I did include the paragraph he suggested.”

Grading punished originality. “I don’t feel it’s fair that just because my paper had a slightly different make-up from what was decided was an A paper it was automatically wrong. I think my paper and everyone else’s paper should be graded on their own merit.” “I was penalized because I interpreted the question differently from those who got As. This is not uncommon and is extremely unjust!”

Poor grades reflected poor teaching. “The comment at the end was that I didn’t have a clear idea of any of these philosophies. I do not think that is my fault. I attend class faithfully and also did a lot of research. I understood these philosophies as best as I could the way they were taught to me.” “The whole class is complicated and confusing, so I’m glad my paper reflects that.”

The essay format itself was unfair. “In a test situation, I think my key ideas would have been correct, and I would have received a much better grade. Besides, I’m not taking this course for credit in English and I was not aware it was a writing intensive course.” Some students apparently have the misconception that unless a course is listed as “Writing Intensive,” the instructor should not assign written evaluations.

I have to have an A. The student needed a better grade to meet college requirements. “A failing grade was a bit harsh.... I am a fifth year senior graduating December 19th and I need this class to graduate.”

These comments reinforce the findings of Becker *et al.* (1968) that students believe that grades should reflect the amount of effort they expend on a course. They also corroborate Natriello and Dornbusch's (1984) finding that, for students, problems with teacher evaluations are that they are "unsoundly based" or "unreliable." Student comments also indicated that our evaluations were not interpreted reliably. As Natriello and Dornbusch (1984) and Milton *et al.* (1986) found, different students interpreted the same grade differently. Some students were satisfied with a B; others found a B insulting or devastating. One student even appealed an A-minus.

The appeals "backfired" by making our unreliability even more visible. My co-instructor's evaluation was that accepting appeals was a mistake because they undermined our credibility and opened the door to more grade-oriented griping. The TA whose decisions he had overturned felt deeply insulted, even threatening to resign. We will probably not try grade appeals again.

My position was that the appeals had been an important learning experience for us as a teaching team, because we had been forced to confront student interpretations of our grading decisions, their lack of understanding of our expectations, and our own unreliability. We learned more about our students' beliefs about grades and grading, and had opened a forum for them to voice critiques of an aspect of schooling that is (for better or worse) centrally important to them. After all, democracy in education is a theme of the course. How could we justify unfair practices or silence student participation? But admitting fallibility contradicts the culture of academe in which the professor's authority and autonomy are protected. We want students to "think critically" about education, but not about **us**! In a paper describing the process of putting critical pedagogy into action in a secondary school, Bigelow (1990, p. 148) laments that "if teachers' only power were to grade students, that would be sufficient to sabotage classroom democracy."

Again, this well-intended attempt to solve a grading problem affected student evaluations of our teaching. Most negative written comments were about grades. There were especially negative comments in the section in which my co-instructor had overturned the TAs grading decisions. There was little appreciation of the work we had expended in trying to make the system more fair. Students also interpreted a flippant remark I had made about the Curators' scrutiny of grades as an indication that we distributed grades according to a Curator-imposed "quota" system with a limited number of As. In fact, our grade distribution was very high, unsatisfactorily high by the Curators' standards.

What Students Believe Grades Should Mean

The following findings are from analysis of student end-of-semester self-evaluations, on which they were asked the question, "What grade do you think you deserve in this class? Why?" The objective of these questions was to probe student

“But I have to Have an A”

beliefs about grades more directly. Of 156 students, 133 responded; the others wrote in a grade without comment. Responses ranged from a few words to two pages. These were coded and sorted according to justifications the student offered. Each entry was identified by the grade the student actually received and the grade the student said he or she deserved.

By far the most common comments came under the general category of **effort**, again confirming the Becker *et al.* (1968) findings. Students reported that they had “worked hard,” and some believed that this effort was not indicated by their grades. The following comments are representative: “I feel my effort was phenomenal.” “I worked extremely hard in this class. I feel that for my efforts I deserve more than a C.” “I gave the class 110 percent effort.” “If grades were based on effort, I would receive an A.” “Because I earned it!” Others defined effort in terms of time invested: “I put a lot of time into this class.” “I put two weeks into my paper.” Students used the self-evaluations as an opportunity to make their effort visible. However, from their studies of high school students, Natriello and Dornbusch (1984) warned that student assessments of effort are unreliable. It is difficult to know what “hard work” means to students without specific knowledge of their effort and the context in which they consider it “hard.”

Some students thought we should consider reading the textbooks as indicating extraordinary effort rather than a minimal expectation: “Because I did all the readings, including **all** of [a particularly difficult text].” Others believed we should consider class attendance: “I attended all but two class meetings.” “I come to class and other students skip most of the time. I think that students who come to class should be rewarded.” In a large class, in which individual absences are not noticeable, attendance does indicate a higher level of commitment. Fourteen said they had sought extra help from an instructor or TA. This was another way, in an impersonal environment, to make effort visible to someone.

A few students noted less ordinary forms of involvement, such as active participation in class discussions (rare in a large class), enjoyment of the class (only seven comments), or discussion of class topics outside of class (only five). The last category was the most appealing to me as an instructor. Remarks such as, “It made me want to read and learn more about the issues. **I often** discussed topics with fellow students and family members” are irresistible. Very few comments, however, referred to such intrinsic motivations.

The second largest category of comments was **learning** or “learned a lot.” “I’ve gained massive knowledge.” “I have gained a great amount of knowledge in this class.” “I have a more than adequate understanding of the material.” The number of such comments may have been skewed by remarks I made in class about grade-oriented versus learning-oriented students. However, the meaning of “learning” in most comments was consistent with a “banking model” of education that entails “absorbing material” rather than critical analysis (Freire, 1990). It was disappointing, personally, that education majors expressed such a reduced concept of learning.

Some complained that essays did not adequately assess their learning of “the material,” and their grades were therefore unfair. This preference for objective examinations has been shown to be an indicator of grade-orientation (Pollio & Humphreys, 1988).

A smaller group, in contrast, commented on the amount of “thinking” the class had stimulated. Again, these were more gratifying comments for an instructor: “This class really made me think.” “I enjoyed the mind-stretching and critical thinking necessary to succeed in this class.” From a nontraditional high LO/low GO student: “It’s all subjective. Give me whatever grade you believe I deserve. I only know that I appreciated not having to take another mindless class.” But A students were more likely to make such comments than those who received Bs and Cs.

According to 31 students, **improvement** should be taken into consideration: “I believe that I deserve a B because as the semester progressed my grades and my overall understanding of the material improved.” Eiszler (1983) found that college students are more positive about grading systems based on personal growth than about those based on competition with peers, such as the “normal curve.” In fact, 18 of the 31 students who made this argument did receive the grade they thought they “deserved.” In our final grading, we tried to take improvement into account.

A large number took a more technical approach to the question and did not really present an argument. They simply compared their work to the criteria for a particular grade, or they calculated their grade mathematically and said that the result was the grade they “deserved.” This group seemed satisfied with the objectivity and fairness of the grading system and its outcomes; for almost all of them, the grade they said they “deserved” was the grade they actually received.

On the other hand, 28 students critiqued the grading system as unfair; these students were more likely **not** to have received the grade they wanted. Students who received Bs and Cs were understandably much more likely to make such comments than those who received As. The following are aspects they criticized:

Communication of expectations: “It seemed impossible to figure out what the professors wanted. There were too many concepts to grasp them all.”

The method of assessment: “I don’t think the whole class grade should be based on writing papers. It might not be easy for some people.”

The unreliability of evaluation: “I feel the way the essays were graded was not fair. Different instructors graded papers differently. I did nothing different on my D paper than I did on my B paper other than get lucky on who graded it.” Some remarked that grade appeals had made unreliability obvious: How could the first grader have given them a C and the second grader a B?

Only 13 students were self-critical. These students sounded defeated and resigned to receiving a lower grade than they wanted: “I was not able to communicate my understanding perfectly in my writing.” “Writing is not my strength.” This comment was from one of only two D students who had not asked anyone for

“But I have to Have an A”

assistance: “I’ve never been good at writing papers anyway.” Six students cited personal difficulties, such as illness, that had hindered their performance.

Finally, only eight students explicitly cited the GPA requirement as a rationale for assigning them a higher grade. Other students, of course, may have had this motivation, but did not choose to use it as an argument. In the end, 55 students did not receive the grade they said they “deserved.” This level of dissatisfaction could have a great deal to do with negative comments on our teaching evaluations. It suggested that we should dialogue much more with students about our (and their) standards and expectations.

Discussion

In retrospect and as a researcher, I am embarrassed by the *ad hoc*, individualistic qualities of my development as a college teacher documented here. As Rau and Baker (1989, pp. 172-173) argued, “college teaching is most often a self-taught, privately validated activity; it remains uninformed by the growing body of pedagogical theory and knowledge that has evolved over the last half-century.” Had I read the scholarship on college grades, I would have understood the history of grading policies, as well as the college student culture. Instead, I proceeded like a traveller without a map or guidebook, discovering already-charted territory. But this “naturalistic” method may well represent the usual beginning college teaching experience. My next step is to move toward more informed investigations that warrant respectable titles such as “action research” (Kember & Gow, 1992; McNiff, 1993; Schratz, 1992) or “practice-centered inquiry” (Chism, *et al.*, 1987).

I do not believe that a graduate course on “methods of college teaching” could have prepared me adequately for this work, but the question of including such a course should become a topic of policy discussions in graduate schools of education. All academics should enter the college classroom with some understanding of the cultures of academe (administrative, faculty and student) and a critical perspective on the conventions of college teaching and the beliefs that keep them in place. Once in the classroom, we need incentives for innovation and self-study (Cross, 1990; Kember & Gow, 1992; Schratz, 1992). As a teacher educator, I believe that I should model for my students an approach to teaching as simultaneously resolving everyday classroom dilemmas, studying and critiquing the institution that too often creates or exacerbates those dilemmas, and working to transform that institution. This statement, in sum, encompasses the conventional categories of teaching, research, and service.

Teaching, seriously considered, is the most intellectually stimulating and politically significant academic work I do. For example, this examination of a mundane, taken-for-granted aspect of college teaching—grades—revealed all too clearly how my work as a college teacher is subject to the same cultural, political, and philosophical conflicts that are the topics of both my research and my course

curricula. Perhaps I should make these parallels more explicit in my teaching, rather than presenting them as problems in schools in general or “out there in the real world.” And perhaps by resolving the grading dilemma together, students and I will come to a better understanding of teaching, learning, and each other.

References

- Beard, R.M., & Hartley, J. (1984). *Teaching and learning in higher education*. London, United Kingdom: Harper & Row.
- Becker, H.S., & Hughes, E.C. (1968). *Making the grade: The academic side of college life*. New York: Wiley.
- Bigelow, W. (1990). Inside the classroom: Social vision and critical pedagogy. *Democracy and Education*, 4(3), 13-19.
- Carlson, D. (1992). *Teachers and crisis: Urban school reform and teachers' work culture*. New York: Routledge.
- Chism, N., Sanders, D., & Zitlow, C. (1987). Observations on a faculty development program based on practice-centered inquiry. *Peabody Journal of Education*, 64(3), 1-23.
- Cross, K.P. (1990). Celebrating excellence in the classroom. Paper presented at International Conference on Teaching Excellence, University of Texas, Austin.
- Eble, K.E. (1976). *The craft of teaching*. San Francisco, CA: Jossey-Bass.
- Eiszler, C.F. (1983). The meaning of college grades in three grading systems. *Educational Research Quarterly*, 8 (3), 12-19.
- Elbow, P. (1983). Embracing contraries in the teaching process. *College English*, 45, 327-339.
- Fong, B. (1987). Commonplaces about teaching: Second thoughts. *Change*, 19(4): 28-34.
- Freire, P. (1990). *Pedagogy of the oppressed*. New York: Continuum.
- Frisbie, D.A., & Waltman, K.K. (1992). Developing a personal grading plan. *Educational Measurement*, 11(3), 35-42.
- Frymier, J. (1987). Bureaucracy and the neutering of teachers. *Phi Delta Kappan*, 69, 9-14.
- Gleason, M. (1986). Getting a perspective on student evaluation. *AAHE Bulletin*, 10-13.
- Goldman, L. (1985). The betrayal of the gatekeepers: Grade inflation. *Journal of General Education*, 37, 97-121.
- Jackson, P. (1968). *Life in classrooms*. New York: Holt, Rinehart and Winston.
- Janzow, F., & Eison, J. (1990). Grades: Their influence on students and faculty. In M.D. Svinicki, (Ed.), *New directions for teaching and learning, No. 42, The changing face of college teaching* (pp. 93-102). San Francisco, CA: Jossey-Bass.
- Jedrey, C.M. (1984). Grading and evaluation. In Gullette, M.M. (Ed.), *The art and craft of teaching* (pp. 103-115). Cambridge, MA: Harvard University Press.
- Johnson, B.G., & Beck, H.P. (1988). Strict and lenient grading scales: How do they affect the performance of college students with high and low SAT scores? *Teaching of Psychology*, 15, 127-31.
- Johnson, S.M. (1990). *Teachers at work: Achieving success in our schools*. New York: Basic Books.
- Kember, D., & Gow, L. (1992). Action research as a form of staff development in higher education. *Higher Education*, 23, 297-310.
- Korthagen, F.A.J. (1988). The influence of learning orientations on the development of

“But I have to Have an A”

- reflective teaching. In J. Calderhead (Ed.), *Teachers' professional learning* (pp. 35-50). London, United Kingdom: Falmer.
- Lowman, J. (1990). Promoting motivation and learning. *College Teaching*, 38(4), 136-139.
- McDermott, R.P. (1987). The explanation of minority school failure, again. *Anthropology and Education Quarterly*, 18, 361-364.
- McNiff, J. (1993). *Teaching as learning*. London, UK: Routledge.
- Milton, O., Pollio, H.R., & Eison, J.A. (1986). *Making sense of college grades*. San Francisco, CA: Jossey-Bass.
- National Commission on Excellence in Education. (1983). *A Nation at Risk*. Washington, DC: U.S. Department of Education.
- Natriello, G., & Dornbusch, S.M. (1984). *Teacher evaluation standards and student effort*. New York: Longman.
- Pollio, H.R., & Humphreys, W.L. (1988). Grading students. In J.H. McMillan (Ed.), *Assessing students' learning* (pp. 85-97). New Directions for Teaching and Learning, No. 34. San Francisco, CA: Jossey-Bass.
- Pulich, M.A. (1983). Student grade appeals can be reduced. *Improving College and University Teaching*, 31(1), 9-12.
- Rabow, J., & Hernandez, A.C.R. (1988). The price of the “GPA perspective”: An empirical study of “making the grade.” *Youth and Society*, 19, 363-377.
- Rau, W., & Baker, P.J. (1989). The organized contradictions of academe: Barriers facing the next academic revolution. *Teaching Sociology*, 17, 161-175.
- Schratz, M. (1992, April). Researching while teaching: Promoting reflective professionalism in higher education. Paper presented at the annual meeting of American Educational Research Association, San Francisco, CA.
- Simon, S.B., & Ballanca, J.A. (Eds.) (1976). *Degrading the grading myths: A primer of alternatives to grades and marks*. Washington, DC: Association for Supervision and Curriculum Development.
- Statham, A., Richardson, L., & Cook, J.A. (1991). *Gender and university teaching: A negotiated difference*. Albany, NY: SUNY Press.
- Summerville, R.M., Ridley, D.R., & Mavis, T.L. (1990). Grade inflation: The case of urban colleges and universities. *College Teaching*, 38 (1), 33-38.
- Weller, L.D. (1986). Attitude toward grade inflation: A random survey of American colleges of arts and science and colleges of education. *College and University*, 61, 118-127.

Comments by Jack Whitehead

I have two of Peggy's papers in front of me: (1) “But I have to have an A”: Probing the cultural meanings and ethical dilemmas of grades in teacher education. *and (2) An action research approach to a contradiction in teaching: Reconciling grades with democratic education. (1994).*

Jack Whitehead
convenes the Action
Research in Educational
Theory Research Group
at the University of Bath,
United Kingdom.

I think you might remember how closely I identified with this paper because of the way in which you had retained your integrity in trying to live your values as fully as you could with your students. We talked about the inchoate nature of some of our work

as we struggled to understand our context. The transformation in An action research approach to a contradiction in teaching: Reconciling grades with democratic education is remarkable. Its methodology is clearly defined in terms of action reflection cycles. You have integrated your dialogues with your students and drawn on the writings of other academics within the action reflection cycles of presentation. I find your writing communicates directly and very powerfully as you move your reader through the living reality of the educative conversations with your students and as you clarify your commitment to value democratic relations in your classroom as its meaning emerges through your practice. I think we all have something to learn from the way in which you have presented the life of your classroom.

I'm worried about your next steps. My worry is focused on your commitment to explore a particular approach to assessment that might be pushing the social and institutional system further than it can accommodate at the present time. I know it might sound presumptuous but I'm going to suggest that you should refocus your enquiry to take into account Karen Guilfoyle's social analysis. You have a range of educational values that you could use to engage with your students. Democracy is one value. Truth, Beauty, Goodness, Social Justice might be others. I think we should all be aware of pushing an institution so far in one direction that it responds by eliminating us.

The example I have in mind is one of our local schools that insisted on trying to develop its curriculum and assessment procedures in a way that did not take into account the changed political realities in Britain after 14 years of right wing, conservative government. By failing to understand the nature of these political realities and the forces against what the school was trying, it took them on headfirst and through a system of inspections, local press, and television coverage had to beat a rapid retreat. I'm supervising some of the staff for their higher degrees and have seen at first hand the damage that they have suffered. Why not build on your success and move sideways into other values that constitute a good social order? Isn't there a danger that by trying to work at ways of eliminating a grading process you are trying to cross a "bridge too far" under the present political realities? I don't think that this is a defeatist analysis. I think it will enable you to broaden the base of your support within and outside your institution from which to move forward. In spearheading a movement it is important not to permit your supply lines to be cut!