On Being "Rigorous"

Kathleen M. Galotti

Since my high school graduation, I have studied and worked in "elite" institutions. I attended college at a top-10 rated liberal arts college and teach now at a competitor school; the years in between were spent in graduate study at an Ivy League university. Throughout these years, one administrator after another has spoken about the general level of talent surrounding us. We were (in my student days) frequently reminded that we were "the cream of the crop," the next generation of the world's leaders, the finest minds being trained for tomorrow. That fact obliged us to exert ourselves in mastering the knowledge and skills we'd need in the future.

Today, at the job I love very much, I often remind myself how privileged I am to be able to teach and guide tomorrow's future movers and shakers. I also am reminded frequently by administrators here of my enormous obligations to my students. In this essay, I want to explore the responsibilities I think I have as a teacher of such talented students and how those responsibilities need to be and should be balanced against a respect for and welcoming of student diversity. In particular, I want to explore what it means to be rigorous and to have high standards, how grading practices relate to these qualities, and whether broadening teaching practices to honor students with different cognitive or learning styles will necessarily un-

What does it mean to have high standards? Does broadening teaching practices to honor students with different learning styles necessarily undermine a commitment to rigor?

KATHLEEN M. GALOTTI, professor of psychology and cognitive studies, joined the Carleton faculty in 1983. She earned her B.A. degree at Wellesley College and her M.A., M.S.E., and Ph.D. degrees at the University of Pennsylvania.

dermine a commitment to rigor. I should say at the outset that I don't think the points I make below apply only to elite places—any institution of higher education has its share of bright and motivated students, and faculty everywhere have (or should have) the desire and obligation to serve them. In fact, I believe my arguments apply, *mutatis mutandis*, to any teacher of any student.

RIGOR AND HIGH STANDARDS

One of the biggest responsibilities I think any teacher has is to set clear and high standards for her students, then to develop curriculum, assignments, and experiences that allow a student to meet those standards. Often, "good teaching" is associated only with the curriculum development aspect of pedagogy. In my view, this misses the mark.

Students, like other people, work for incentives. Rational students prioritize their time and energy in accordance with payoffs. When they are in a fairly non-demanding situation, where mediocre or "okay" work is acceptable, they will eventually find themselves meeting that standard. If they are bright, that standard will be all too easy to meet and little intellectual "stretching" or development will take place. Conversely, if only their best efforts are enough, students will find themselves developing new skills, new abilities, new ways of knowing and thinking. Of course, the level of rigor has to be calibrated to the general ability of the relevant student body: Asking people to do more than they are capable of leads inevitably to a great deal of frustration. The trick, the art of teaching, is to learn how to achieve this calibration and how to individualize it for different students.

Now, students don't always immediately appreciate being pushed. And, as I remember all my experiences with learning, it isn't necessarily "fun" in the middle of the process, especially when facing a new and daunting challenge. It's hard, it's scary, it's tempting to give up, it's frustrating more often than it is exhilarating and invigorating. The "fun" part comes after the challenge has been mastered—it's often a retrospective sort of thing.

Students at Carleton complain a lot about the stress of classes and assignments and how much is expected of them. Until, that is, someone makes a real move to reduce that stress. Then, students resist more than any other constituency (faculty, staff, administrators) stress-reduction initiatives such as reducing the number of courses required for graduation, dropping the idea of

academic distinction in the major, instituting a "common time" during which no classes are scheduled each week. The stresses they face are seen as defining characteristics of their experience, ones worn proudly as badges of achievement. Whether this reaction is one common to most college students or not is hard to say, but I sometimes worry that we (faculty, administrators, parents) implicitly push the "you are top-notch and therefore able to carry superhuman workloads" message far too often at Carleton.

GRADING AND ASSESSMENT

One major mechanism for both enforcing rigor and high standards is, of course, the practice of grading. Not surprisingly, one general source of stress for many students is grades. Concern over the assessment of one's performance, and one's relative standing, are as common today as ever. Most of our students are used to being in the top percentiles of students (e.g., in their high school class), and it comes as a great shock to many that half will be "below average" in this new environment.

Perhaps in response to signs of student stress, there has been an issue of grade inflation throughout most institutions of higher education, Carleton included. When top-rated universities are graduating over half their students with Latin honors, it is a natural worry that the "honor" starts to lose meaning. And if every student who puts forth minimal effort can come to count on receiving at least a "B" (as is the case in many classes that I've taken or heard about), the very notion of grades as indicators of achievement or mastery becomes hollow (Rosovsky and Hartley, 2002, explore the problem and consequences of grade inflation in a thought-provoking essay).

If grades cause stress, or else give rise to inflation that undermines meaning, maybe it is time to question why we give grades at all. The standard answer offers two reasons: 1) incentive to meet standards, and 2) information about how well standards are being met.

As a teacher, I've found that grading work typically elicits better performance from students than not grading it. Papers students know will be graded typically are better written and organized; likewise with oral presentations, quizzes, etc. This phenomenon is understandable, and goes back to the idea that a rational student will align his behavior with the payoffs for that behavior—nongraded work has only the payoff of personal satisfaction, while graded

work offers, in addition, more visible and perhaps public affirmation of a job well done. Grades seem to lend more importance to assignments.

Some colleagues make additional arguments that grades provide crucial information to students, allowing them to more accurately diagnose their strengths and weaknesses. A professor who gives out "charity" Bs, then, is seen in this view as doing a grave disservice to her students. By not accurately assessing their work, she gives false encouragement and information to them.

Of course, there are counterarguments to the utility of grades. One is that providing external incentives, even symbolic ones, can undermine an activity's inherent appeal. Psychologists Lepper and Greene showed this result in a study of preschool children who liked to draw and spent a lot of their free time doing so. When the investigators started giving "good player awards" for drawing, the activity suddenly became less fun for the children. By analogy, it can be argued that providing grades as an incentive for academic work makes that work inherently less fun.

A second argument has to do with the meaningfulness of the discriminations made by grades. This argument goes as follows: Separating the A students from the B, C, D and F students may make sense when you are dealing with a large and diverse population. However, imagine taking all the A and B students only, enrolling them in a program, and again trying to sort them into five groups: the As, Bs, Cs, Ds and Fs. Now cull, and sort again. At some point, the distinctions start to lose their meaning. The difference between an A and a B (or an A- and a B+) start to become miniscule and unreliable—the evaluator needs to start finding fairly small nuances to accomplish the sort. It has been argued that, at the "elite" schools where I've been, we are in this situation and the differences between students receiving different grades are negligible.

A professor who gives out "charity" Bs, then, is ... doing a grave disservice to her students. By not accurately assessing their work, she gives false encouragement and information...

"Grade eggs, not people!" was a banner cry I remember from my own college days in the 1970s. The protest was against providing any kind of labels to people at all, against reifying a number (such as a grade point average, an SAT score, an IQ) and imbuing it with importance. Grades in high school

were seen (cynically) as providing validation for entry into college, grades in college as validation for entry into graduate, law, business, or medical school. The sense was that all these institutions of higher learning ought to find a better, fairer way to select their applicants than by making feeder institutions use grades to sort.

MULTIPLE INTELLIGENCES AND LEARNING STYLES

Another argument against grades can be derived from the arguments advanced against the use of IQ scores—that they reduce performance unreasonably to a single number. This means that important aspects of performance are ignored or given short shrift.

Psychologist Howard Gardner broadened the concept of "intelligence" a few decades ago when he introduced the concept of multiple intelligences—noting that there are very distinct realms of achievement in which people can and do excel. For example, in addition to the traditionally prized academic intelligences (linguistic and logical-mathematical), there are the following: musical intelligence, bodily-kinesthetic intelligence, spatial intelligence, interpersonal intelligence, intrapersonal intelligence, as well as the recently added naturalist intelligence. Gardner's recent book (1999) presents detailed descriptions of each of these and makes the arguments for their existence.

Though certainly not without controversy, Gardner's claims have a lot of intuitive appeal. Who hasn't encountered the student who excels in one domain (say, using computers or statistical analysis) but struggles mightily with another (say, writing papers or delivering oral presentations)? Or the student with mediocre performance on tests and papers, who gives a brilliant piano recital or delivers a stellar athletic achievement?

Related claims about the existence of important individual differences come from researchers studying learning styles (see Sternberg, 1997; Rayner, & Riding, 1997 for overviews). The idea here is that, in addition to differing in terms of physical characteristics (e.g., hair, eye, or skin color, height, weight) and in terms of intellectual ability (e.g., intelligence or intelligences), people also differ in reliable and predictable ways in how they use their intellectual abilities, the way they approach situations or problems. These approaches may relate more to personality differences than to ability differences and have come to be called cognitive, thinking, or learning *styles*.

For example, we have heard proposals about visual versus verbal learners, the idea being that different learners find different approaches to acquiring information—either through viewing spatial depictions versus hearing linguistic explanations—more comfortable, intuitive, or natural (e.g., Edwards & Wilkins, 1981). There are proposals for whether people are extraverted or introverted, approach information judgmentally or not, in some of the work on the Myers Briggs Type indicator (Myers & McCauley, 1985). Other investigators adopt the concept of "need for cognition" to discuss individual differences in motivation to engage in cognitive tasks such as problem solving (Klaczynski & Fauth, 1996; Stanovitch & West, 1998, 2000).

The point here is that there exists documented evidence of apparently stable, personality-related traits, that seem to impact the way people learn, acquire or interpret information, make decisions, and perform other higher-order cognitive tasks. Although the evidence is not without its critics, once again there is some good reason to believe that people approach learning differently and in ways having to do with qualities other than their intellectual ability. Indeed, it's not a far-fetched idea to believe that some learning styles may be gender or perhaps socioeconomically or ethnically related (e.g., Belenky, Clinchy, Goldberger & Tarule, (1986/1997); Galotti, Clinchy, Ainsworth, Lavin, & Mansfield, 1999).

So the dilemma I see boils down to this: On the one hand, we want to adopt and maintain high expectations and high standards. We want to insist that our talented students "stretch" themselves intellectually. To do less is a disservice to any student, but perhaps particularly to the "elite" students who choose institutions such as ours *because* of the intellectual challenges we purport to offer. On the other hand, many recent and credible proposals suggest that a "one-standard-evaluates-all" model of assessment will not equally benefit (or honor) students with different styles. At the limit, forcing everyone to one standard may be unfair and discriminatory.

What to do? Do we hold fast to traditional standards of rigor, insisting that if the students are bright and motivated enough, they'll be able to adapt (and will, in the long run, benefit from developing the traditionally prized and assessed abilities and skills)? Do we abandon the traditional standards, accepting the argument that they unfairly privilege some cognitive or learning styles

over others? If so, are we willing to risk abandoning the intellectual challenge to stretch?

Is there a way to customize our standards for every student, trying to match our expectations with his or her work? Thus, could we ask a "high verbal" student to continue to produce the traditional term papers and oral reports, but maybe develop new assignments and assessments for students who have other gifts? Of course, then the issue of comparability of standards arises—how would we know if the alternative assignments were as demanding, as comprehensive, and as well assessed as the originals?

Rigor *is* very important to me as a teacher; but I'm becoming more convinced that good teaching *also* requires sensitivity to diversity of various sorts.

Perhaps we should ask all students to assemble a portfolio of work—showing their performance across different types of tasks (making sure that the portfolio contains a product relevant to each major type of learning we hope to see). But again, how are rigorous standards to be applied to the assessment of the portfolio? Is customization, at the limit, equivalent to an abandonment of standards? If the "assessment" of the portfolios becomes so individualized, at the limit, then no comparisons can be made between one portfolio and another. This process would lead to a situation where every student is deemed to be "special" in "his or her own way"—in other words, the Barney (the dinosaur) view of assessment, where nothing is better or worse.

PERRY'S EPISTEMOLOGICAL MODEL AS A METAPHOR

I don't, unfortunately, have definitive answers to any of the above questions. I have mainly an uncomfortable feeling that some middle ground needs to be found. Rigor *is* very important to me as a teacher; but I'm becoming more convinced that good teaching *also* requires sensitivity to diversity of various sorts. Yet the "we need to find a middle ground" answer seems dangerously empty and hackneyed—I feel, as my students say, like a "wuss" simply to end with that idea.

So here's an analogy I want to present, one that I think is meaty and non-wussy. It comes from the work of William Perry, a researcher who ran the Harvard University Bureau of Study Counsel for several years in the 1940s and

1950s. Perry worked with undergraduates (most of them male) at Harvard, and published a report of the qualitatively differing patterns of thought he noticed in them as they proceeded through their undergraduate years (Perry, 1970, 1981). Only a brief summary of the results is possible here.

Perry identified three major structures of thought, which seemed to form a developmental progression. Younger college students were often found to make meaning of the world from a dualistic orientation, dividing the world into dichotomies: right vs. wrong, true vs. false, good vs. bad, we vs. others, what They (teachers) want vs. what They don't want. Students at this level attempt to place newly learned theories into these categories and to "learn" through simple "obedience" and adherence to instructor assignments.

The second major structure Perry called "multiplism." It involves a "live-and-let-live" orientation to new or foreign ideas: Anyone has a right to any opinion, and no one else is really in any position to judge or evaluate the validity of those opinions because they haven't had the same life experiences as the person holding a different opinion. Perry notes the liberating feeling this structure provides to many students: No professor or teacher can really comment, fairly, on a student's view of a piece of literature or an experiment—for the student's sincere opinion is just as valid as the professor's.

In the optimal case, students eventually construct a third, more synthetic structure, one that Perry called "relativism." In this structure students remain convinced that very few absolute answers exist. Nonetheless, they come to appreciate that not all answers have equal status, but vary along different dimensions: e.g., internal coherence, explanatory power, plausibility. Thus, standards can be discovered or constructed that allow one to make distinctions among better or worse answers. So, although there is no "right" theory of (say) child development, no "right" interpretation of *Hamlet*, existing theories can be assessed along different dimensions, and overall goodness can be assessed (at least with respect to certain assumptions). I should note that some of Perry's work was later challenged by feminist psychologists who discovered some differences in developmental progression when they studied female undergraduates (Belenky et al., 1986/1997). Nevertheless, for our purposes, the general Perry scheme seems to me to have sufficient generality as a model.

Let me now try to connect Perry's proposal with the idea of what it might mean to be rigorous. An analogy can be made, I think, between Perry's

dualism and a rigid adherence to only traditional methods of assessment. Clinging to *only* one method of assessment, and including *only* traditional topics and assignments in a course because they are the "right" ones, is quite analogous to the first-year student who seeks to sort theories and ideas into only two categories, "right" and "wrong."

Similarly, I am struck by the similarities between Perry's description of the outlook of a multiplistic student ("one opinion is as good as any other") and the notion that all methods of assessment are as good as any other or that no assessment is necessary or possible. Because there is no one true, "right" way to assess student work does not necessarily imply that all methods of assessment are equally good, or equally flawed. Nor does it mean that radical customization (everyone just assemble whatever work is representative of your learning into a portfolio, and we'll all appreciate it) is the only possible response.

Because there is no one true, "right" way to assess student work does not necessarily imply that all methods of assessment are equally good, or equally flawed. Nor does it mean that radical customization (everyone just assemble whatever work is representative of your learning into a portfolio, and we'll all appreciate it) is the only possible response.

Just as good teachers push, nudge, or cajole their students toward Perry's relativism—a position where certain standards are constructed or adopted to assess ideas—so to do we as teachers need to grapple with a never-ending quest to find ways of constructing good assignments *and* good assessments. Certain ones *are* better than others, and we need to be prepared to develop and debate standards to be used to determine the relative worth. This will likely require that our assessments themselves be subject to periodic and rigorous assessment and revision.

Recognizing that there are many paths to a goal is one way of honoring diversity. For example, we want fluent, articulate writers at the end of college. But many different approaches to developing writing skills exist. Having students jointly author (and edit) papers is one way of honoring students with a more extraverted learning style—just as having students independently author a paper honors a more introverted approach. Where appropriate, faculty could

assign more hands-on assignments—being part of a theater production, for example (see Weiner and Hardy, this volume), or creating a piece of furniture (see Clark, this volume). The general point is to make sure that a course's assessments don't draw disproportionately on just one skill, ability, or intelligence. Figuring out the right mix of assignments, and the right mix of assessments, is the key.

So, can or should we be rigorous, especially if we are in elite places? I believe that we must. If we don't adopt and commit to high standards, few of our students will achieve as much as they are capable of. But being rigorous does not mean rigidly clinging to a single way of teaching or assessing. Indeed, the best teachers I know are quite flexible in their work, trying always to diagnose each student's current level of understanding and skill and tailoring a series of assignments and experiences to bring this level up a notch. Our challenge now is to extend that commitment to rigor to the development, design, and assessment of those assignments and experiences. ■