Administrators of universities are increasingly emphasizing teaching, especially at the undergraduate level. This is true of both private and public universities. I don’t suppose there was ever a time in American education when administrators ever came out against teaching, but often teaching suffered benign neglect in the sense that good teaching was rarely rewarded and bad (even atrocious) teaching was rarely punished, or even noticed. If a professor were not up to snuff, nothing much was done about it. I’m showing my age, but I can remember a time when numerous faculty members did not prepare syllabi for their courses, did not bother to be too organized in their approach, and often assessed student performance infrequently and haphazardly. Twenty or more years ago, I heard professors speak derisively of having to teach undergraduates.

I would like to think that those days are mostly in the past. I used to ask my colleagues who sneered at undergraduate teaching, “If you don’t want to teach, why did you go into university education?” They would usually reply “to do research.” My response was and is to tell them to resign from the university and go to a research institute. Even the greatest of research universities has a critical teaching mission at the undergraduate level.

The strong emphasis on quality teaching is, in my view, a positive feature of university life today. There is only one problem — no one really knows how to measure “quality teaching.” We try all sorts of things – teacher ratings, portfolios of teaching (with various materials included), observation of teaching by other professors, assessment of professors by teaching assistants, to mention some strategies. All these are useful indicators, but none perfectly assesses quality teaching.

I am not sure it is possible to measure teaching accurately, because of the nature of the beast to be measured. To ask “Is this professor a great (or even competent) teacher?” presumes that all students assess teaching the same way. However, we all know from our own university experiences that assessment of teaching (like most matters of person perception) is a curious and complex interaction between the perceiver and the perceived. Some of my most beloved professors from my undergraduate days were not much appreciated by my friends, and conversely. Whether a teacher is judged to be excellent depends on the student’s personality and how it meshes with the professor’s, the nature of the course, the student’s interest in the course, and numerous other factors. Some professors might be great for serious majors, but off-putting to other students taking the course as an elective.

The easiest way to measure teaching effectiveness is through student ratings, which have become widespread. I think they can be useful if they are collected carefully and systematically, but often they are given on a random day of class and so naturally only attendees fill them out. Sometimes the proportion of the class taking them is small, which leads naturally to the question of what the other students thought: Were they voting with their feet by remaining in the dorm? There is no way to know. (One semester I decided to correct this problem in my own case by handing out the class ratings during the final exam. That experiment was a big mistake, unless you want tired and somewhat irate students to
assess you in that mood state).

Ratings do, I believe, provide useful information, but they should properly be viewed as only one of several criteria. Like the weighting of SAT scores in college admissions, I suspect student ratings have a disproportionate effect on whether professors are perceived as “good teachers” both by their students and by their colleagues. The good aspect of student ratings is that professors can try to change legitimate criticisms of their teaching, such as being disorganized or facing away from the class too much. The bad aspect of student ratings is that professors may also change good educational practice because the students object to it and want an easier course. In my own ratings, students frequently complain that I assign too much reading, that I give cumulative final examinations across the entire course (when “the other section” does not), and that I prefer essay tests to multiple choice tests. They may downgrade my teaching because of these features. (Also, I’m not a lot of fun. I don’t tell jokes). I don’t change these features of my course, to have light reading, multiple choice tests, no final, etc., because I think the educational process would be worse if I took these steps, even though my teacher ratings might go up a bit (and they are reasonably good, anyway).

I can recall taking two courses in college in which I got the highest grade in class (grades were very public back then – posted by students’ names on professors’ doors, usually), and yet I got a B+ in one case and a B in the other. In both cases, the professor told me that he’d been teaching the course for years, knew A work when he saw it, and I hadn’t done it. Both courses were great and I worked hard and learned a lot (in these English and Political Science courses). I wonder if professors of psychology ever do this now; would students complain all the way to the President of the university for getting a B with the highest grade in the class? Do we have courses that thoroughly challenge our best students with demanding written assignments?

In short, I sometimes worry that teacher ratings may serve the unintended purpose of dumbing down higher education. Perhaps that is the natural price we pay for the good effects that teaching assessment can have in helping to improve some features of teaching, but I am sometimes led to wonder about the balance that ratings strike in their effects on teaching.

If you worry about your ratings with students, Ian Neath of Purdue University has written a useful article called “How to improve your teaching evaluations without improving your teaching” (Psychological Reports, 1996, 78, 1363-1372). He surveyed the literature on teaching effectiveness and synthesized the points into 20 helpful tips to improve teaching. For example, Tip 1 is to “Be Male” because males, in general, get higher ratings (although the effects of gender are somewhat complex, Neath reports). Other helpful tips include lists of “do’s” and “don’t’s,” such as: don’t teach required courses (and especially not statistics); do teach higher level classes and do teach only small classes (because both are associated with higher ratings); don’t cross-list your courses (because non-majors rate courses worse than majors), and so on. OK, Neath meant his paper (mostly) as a spoof, and we all know correlation is not to be confused with causation – but the suggestions he makes are based on the research on what determines highly rated teaching. The point embedded in the spoof should lead advocates of teacher ratings as the chief means of evaluation of teaching to pause and reflect.