

## Why the University Should Abolish Faculty Course Evaluations

Clark Glymour  
Draft, November, 2003

In company with several other University bodies, the Faculty Senate has recently approved a new Faculty Course Evaluation instrument, to be filled out on-line. Both for students and for faculty the decision is unfortunate, and quite possibly for some faculty, sometime, it will be very unfortunate. What the Faculty Senate ought to have done was to recommend investigation of a more serious process for estimating the quality of instruction, leading towards the end of University sponsorship of Faculty Course **Evaluation** forms. Students should be entirely free to organize and publicize their own on-line evaluations of courses and faculty, but the results should not have the *imprimatur* of the University itself. I will start my argument with some anecdotes.

In 1969 Princeton University for the first time admitted about 20 African American students, and nearly half of them enrolled, with sixty other students, in my Introduction to Mathematical Logic. Every one of these smart, brave, ambitious black students failed my course. I thought hard over the summer about why, and formed this hypothesis: the lecture course had based grades on a mid-term and a final and homework. If the African American students were missing some background, or not good at test taking or at judging how well they understood the material, the course structure offered no way for them to make up for those disadvantages by extra effort. The next year I changed the structure of the course. Using a text that divided the material into a great many short chapters with many problems, a student could take a test on a chapter at any time during regular work hours and have it graded immediately; if the test was passed, the student could go on to the next chapter; if the student failed, another test on the same chapter could be taken after a two-day wait; tests on a chapter could be taken repeatedly until one was passed. Lectures were replaced by scheduled problem solving sessions in which the students asked me how to do problems in the text, and I worked the problems out for

that I had contrived the arrangement to save the trouble of preparing lectures. Moral ; *Student evaluations are more influenced by formats meeting their expectations than by how much they and their classmates learned.*

Several years ago I served on a committee established by the Dean of Arts and Sciences at the University of Pittsburgh to review the case for tenuring a young mathematics professor there. The man in question had been promoted to Associate Professor without tenure, an entirely anomalous position at Pitt, as at many other universities (but not, of course, at Carnegie Mellon.) The Chair of the Mathematics department made the case for promotion to the committee: the fellow had not been given tenure previously because, although his scholarship was excellent, his faculty course evaluations were unacceptably low, but they had since improved, and so he should now be tenured. Committee heads nodded as the Chair went on about how the Mathematics department valued teaching. I asked the Chair these questions, with the following answers: Was there any evidence other than FCEs that the fellow was a poor teacher? There was not. Prior to the previous decision to promote him without tenure, what had the professor been assigned to teach? Sections of Calculus and of Differential Equations. Were there many such sections? There were. Did they use the same texts and give the same examinations? They did. On average, how did students in his sections do on the final examination compared with students in other sections of the same courses? Here was the give-away: On average, they

departments by the amount of mathematical content they believed to be typical of courses in each department. In every case, the rank ordering by mathematical content was the same as Dean Modell's ranking by FCEs. Moral: *Unless students are committed to a mathematical curriculum (as are, for example, majors in Mathematics, Computer Science, Statistics and Physics) the more mathematical content a course has, the less students tend to like it.*

For my first five years at Carnegie Mellon, each semester I taught a required Introductory Philosophy course to about 250 students in one big lecture class. I was interested in whether the course improved students' reasoning abilities, which had not been directly

kinds. I know of no good studies that show that courses in which more learning actually goes on—or more that is worth learning is taught—measured, for example, by pre-test and post test performances are more highly valued by students *for that reason* than are less instructive courses. But even if that were so, the FCEs are heavily biased instruments: biased against faculty who have formal approaches, who let students know their grading will be rigorous, who aren't comely, who adopt original methods of instruction. *No student would agree to be evaluated by such criteria. No promotion committee would explicitly count such considerations against promotion of a faculty member, but implicitly it is done all the time.*