

(Continued from page 18)

CSU Granting Doctoral Degrees

administer the program. The Master Plan, indeed, was visionary and generous at its inception. Higher educational opportunities were extended to those who were striving, by means of a higher education, to become part of the middle class. In this noble process, astounding opportunities were afforded ambitious aspirants to the upper level professions due to a rapidly expanding CSU faculty. This occupational mobility was even more unbelievable for those of us faculty of humble family origins.

No matter how much we delude ourselves, our current efforts only maintain a status quo. The Master Plan no longer benefits CSU and deprives California of the fullest potential from CSU. No stampede to adopt doctoral programs will occur. Only some departments and universities will award doctoral degrees. The UC System will continue with the older professions such as medicine and law. The CSU campuses would emphasize more recent or applied disciplines such as education and business. An increase in total doctoral degrees granted is unlikely. Some doctoral candidates, for example, might attend CSU rather than private universities.

A CSU doctorate should and would have standards as rigorous as UC; we CSU faculty would ensure that. No knowledgeable person would suggest that the academic standards, for example, of Ohio State University are less rigorous than the University of Ohio. The CSU tradition of teaching excellence would not diminish. Faculty still would be required, for promotion and tenure, to run the same grueling gauntlet about their teaching effectiveness. Faculty, we would hope, would no longer feel the need to apologize for being a scholar at CSU.

Cost estimates largely reflect the ideologies of the pronouncing

“experts.” It is doubtful that the total cost for universities will be greater than current costs relative to percentage of tax revenue or state income. Because of the loosening of the UC doctoral degree monopolies, costs could be lower. Competition tends to lower prices for commodities and services. We don’t know about state support until we ask. If we don’t, less worthwhile projects instead will be legislated and funded. With the Chancellor, faculty, and all other university constitutions, our efforts can bring us together as we focus on this common objective. Ideal times and circumstances never will be. The time is now, the place is here and the goal is right.

Leaving succeeding generations with opportunities equal to what we had in the pioneering years of the Master Plan poses great difficulties. The

restless pioneering spirit of California requires our best effort. We guardians of the people’s university can renew our commitment each time we look at our students and see the descendants of the poor, minorities, and the oppressed. We need a vision that reaches beyond the status quo if we are going to leave a legacy. Future generations, then, can remember us as a faculty who helped make a difference that was worth remembering. ■

Editor’s Comment

What are your thoughts? Go to <http://faculty.fullerton.edu/senatenews> and register your concern in the discussion group “PhD or Not for CSU.”

C.E. Tygart, is a Professor of Sociology and a member of the IRB Committee. He was a member of the University Research Committee in 1972, when IRB was established.

A Survey of Faculty Opinions Concerning Student Evaluations of Teaching

Michael H. Birnbaum

Student evaluations of teaching were originally intended to help improve instruction, but they may be doing more harm than good. Because retention, tenure, promotion, and merit salary raises are influenced by student evaluations, faculty members make changes in their courses that *they believe* will improve their evaluations. This article explores beliefs held by members of the faculty concerning how changes in grading standards and content of courses would affect student evaluations and student learning.

A majority of CSUF faculty who were surveyed judged that student learning can be improved by increasing course content and by raising standards for grading. However, they also stated that



these improvements would hurt their evaluations. The majority judged that the current system of tenure and promotion discourages raising standards, encourages lowering of standards, and promotes “watering down” of course content. Most said that ratings are hurt by changes that would improve learning, and that the use of student evaluations of teaching is

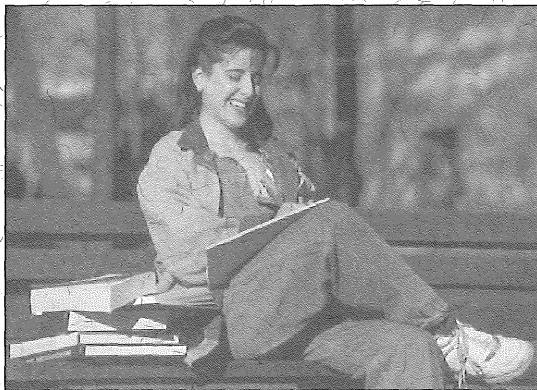
(Continued on page 20)

(Continued from page 19)

A Survey of Faculty Opinions Concerning Student Evaluations of Teaching

harmful to the quality of education. In another survey of 142 CSUF students, the majority gave highest ratings to courses with the least content and the lowest standards; thus, the faculty understands student opinion.

A recent issue of *American Psychologist* featured the controversy on validity and biases of student evaluations of teaching. Meta-analysis of studies concluded that less than one-sixth of the variance of evaluations is



associated with educational performance. Some authors warned that ratings are so complicated that anyone using them for practical purposes must understand nonlinear, nonadditive, multidimensional modeling of confounded judgment data.

My field of research is human judgment. With the same methods used in student evaluations, I found that the number 9 is judged to be significantly "bigger" than 221. Since $9 < 221$, we should be careful not to evaluate faculty by the same methods that lead to wrong conclusions.

Apart from the actual validity of student evaluations is a potentially more important question, namely, their *perceived* validity. Although some teachers are fired because of student

evaluations, most figure out how to get better evaluations. Do their adjustments promote student learning? No, according to a survey of CSUF faculty.

Two hundred and eight CSUF faculty responded to an email survey. Seventy six faculty members had less than 12 years experience (68 were untenured), 66 had 12 to 24 years, and 64 had more than 24 years. Following are some of the results:

◆ ***If you were to RAISE standards for grades in your class, would it affect your student evaluations?***

Nearly two-thirds of those surveyed (65.4%* or 136) reported that higher standards would result in lower evaluations, and only 3.4% (7) thought the opposite would occur; the others stated no difference. (*Asterisks designate that split are statistically significant throughout this paper.)

◆ ***If you were to INCREASE the amount of CONTENT (material) in your classes, would it affect student evaluations?***

About two-thirds (65.9%*) responded that increasing content would decrease student evaluations, against only 4.8% who stated the opposite. The theory proposed is that with less content, the student believes that the instructor was very successful in teaching the subject. Because students do not know what content should have been included in the course, they will not know that important material has been omitted until later, long after the evaluations are done.

◆ ***Are student evaluations influenced by such variables as the teacher's personality, attractiveness, gender, race, dress, religion, ethnicity, sexual orientation, or disability status?*** In response to this question, only 16.8% responded that student ratings are "unbiased"; 52.4%* responded that students are biased in favor of certain groups; 26%*

responded that students are biased *against* certain groups.

(The questionnaire defined student learning as "knowledge of the subject matter, as might be measured by objective, standardized exams...the sum of knowledge and skills that the student retains from the class and will be able to use in the future.")

◆ ***How would increasing the content covered in class and in assigned readings affect student learning?***

45.2%* said that increasing content would increase student learning compared to 27.9% who thought the opposite.

◆ ***How would raising standards for grading affect student learning?***

57.2%* responded that raising standards would increase student learning against only 7.7% who indicated the opposite. The theory most often expressed was that students will work to achieve a certain grade. If less is required to pass, students ease off in their studies, so they learn and retain less.

◆ ***Does the current system of promotion and tenure give incentives to RAISE standards for grading?*** A surprisingly high 92.3%* stated "no" compared to only 5.8% who said "yes."

◆ ***Does the current system of promotion and tenure encourage faculty to LOWER their standards?*** 70.2%* said "yes" against 28.8% who said "no."

◆ ***Does the use of student evaluations encourage faculty to "WATER DOWN" content in their courses?***

72.1%* said "yes" against 26.9% who said "no."

Thus, the majority opinion of the faculty is that the incentive system for tenure and promotion causes faculty to lower standards and water down

(Continued on page 21)

(Continued from page 20)

A Survey of Faculty Opinions Concerning Student Evaluations of Teaching

courses, which most faculty members believe will decrease student learning. Apparently, the majority of faculty believe that the incentive system has the opposite effect of what a citizen in favor of quality education would support.

◆ **Over the years, have you changed the amount of material presented in your classes?** 48.6%* said that they now present less material against 14.9% who said that they present more material, and the rest indicated no change.

◆ **Over the years, have you changed the standards required to get a passing grade in your classes?** 32.2%* said that they now use lower standards against 7.2% who said that they now use higher standards. Since the majority opinion is that reductions in content and standards are harmful to student learning, it seems sad that so many faculty concede having made changes that they believe reduced the quality of education.

◆ **Please assess the preparation of students who are now enrolled in your college or university, compared to previous years.** The majority (67.3%* or 140) reported that students are not as well prepared now, compared to only 2.4% (5) who said the opposite. When asked what percentage of lower division students possess the study skills one should expect of the top 1/3 of high school graduates, the median response was 40%, with 85 responses below 30% and 134 (64%*) less than or equal to 50%. Apparently, about two-thirds of the faculty think that half or more of our students do not qualify under the State's concept for admission.

One theory is that declining standards for recent new teachers is a cause of

this problem. Based on data published each semester at CSUF, students who plan to be teachers have some of the highest grade point averages (GPAs) on the campus. When asked if students with the highest GPAs are indeed the best students, only 12.7%* thought these "future teachers" are our best students; about twice as many rated these students as below average on the campus, and 55.8% judged them average.

◆ **What percentage of undergraduates who want to be teachers do you think should become teachers?** Nearly two thirds (63%*) of respondents said that less than half should become teachers.

◆ **What percentage of graduates in your department possess the general education, specific skills, and knowledge base that should be required of a graduate?** The median response was 60%. Thus, the average faculty member believes that two out of every five of our graduates are not qualified to receive the degrees we confer upon them.

A sample of 142 lower division students evaluated 89 hypothetical classes, based on combinations of three variables: instructor's individual characteristics (personality), standards for grading in the course, and the amount of content. The students represented 29 different majors; there were also 26 with undeclared majors. I anticipated that this heterogeneous mix of students would hold a variety of different views of what would be the optimal class. However, to my surprise, the students were remarkably homogeneous in their evaluations of courses:

- 94.4% (134* of 142) gave higher evaluations to an "attractive, well-dressed, 36 year old female with a nice personality" than to a "62 year old male with a slight tremor (due to a previous stroke) who doesn't smile in class."

- 92.3% (131*) gave higher ratings to a class with "light" content (less than



100 pages to read in a semester, and nothing else to do outside of class) than to a course with "heavy" content (800 pages to read and homework assignments); only 9 gave highest ratings to courses with the most content. Only 16.9% (24) rated a "medium" level of content as better than the "light" level, although the "medium" course was described as having "300 pages of medium level reading" to do in the semester, and the course might require some study to master the material.

- 97.9% (139* of 142) gave higher ratings to a course with "very easy" standards than to a course with "very hard" standards. Only 14 (9.8%) students gave their highest ratings to a course with "medium-easy" or "medium-hard" standards.

The "very easy" standards course was described as follows: "This instructor gives most students As and Bs, even those who are struggling with the material or who have not been diligent in attendance and study. Only the most clueless student will get a C in this class. If a person has half a brain and attends some of the time, (they get) an A or a B." In the "medium-easy" course most students get As and Bs. "Medium-hard" was a class with 30% As and Bs, 50% Cs, and 20% Ds and Fs. The "very hard" course assigned 7% As, 13% Bs, 40% Cs, 25% Ds, and 15% Fail.

Students gave the highest rating to the course in which the teacher is attractive, where the standards for grading

(Continued on page 22)

(Continued from page 21)

A Survey of Faculty Opinions Concerning Student Evaluations of Teaching

are lowest, and where the content is least. Apparently, the majority of faculty are correct in their understanding of what students like.

What are the conclusions we may draw from this? According to the majority of faculty members, the incentive system (using student evaluations for promotion and tenure decisions) puts teachers in a conflict of interest between making changes that would improve student learning and making changes that would improve student evaluations. An implicit assumption in the use of student evaluations is that the average student is more likely right than the professor. However, it is dubious if a professor should redesign a course to suit anonymous comments by students who have not yet finished one class on the subject. It seems doubtful that students who have not yet taken the next course in a sequence can judge if they were adequately prepared in the first course.

Many students are inaccurate in describing what the teacher said in class when they are *motivated* to be as accurate as possible (when taking

exams). Therefore, is it reasonable to assume that these same students are accurate when they give evaluative descriptions anonymously with no incentive to be accurate and no penalty for libel?

Our incentive system has produced a decline in standards that diminishes education. Students are motivated to get good grades, and faculty are motivated to get good evaluations. Unfortunately, both of these interests can be satisfied by reductions in content and grading standards, which diminish education. The finding that the average member of our faculty thinks that only 60% of our graduates have educations to match their degrees is a sign that our institution is in trouble. We should begin to study how our incentive system can be changed to align the interests of students, faculty, and the people of the State. ■

Editor's Comment

A more complete version of this paper may be found at <http://faculty.fullerton.edu/senatenews>.

Michael H. Birnbaum, Professor of Psychology, is Director of the Decision Research Center, member of the Publications Committee, Society for Judgement and Decision-Making, and on the Executive board of the Society for Mathematical Psychology.

Reply to Michael Birnbaum's "Survey of Faculty Opinions Concerning Student Evaluations of Teaching"



Gayle H. Vogt

Michael Birnbaum's article, "A Survey of Faculty Opinions Concerning Student Evaluations of Teaching," again raises the issues of fairness and validity in the personality contests called "evaluations." Fairness is an issue because of student anonymity, a direct violation of procedural due process, a process

required by various education decisions in both the United States Supreme Court and the California Supreme Court.

Recently, a personnel committee member repeated, in writing, a remark he read in a student evaluation. The student comment was false, could have been harmful to a professional reputation, and was seen in a performance evaluation by several other parties. The reader will recognize these markers as the tests of libelous accusations, libel on the part of the student and the professor who repeated the false statement.

Validity is another important issue. Students may feel free to falsely accuse a professor when classroom standards are high and, as a result, grades are lower. Research shows little correlation between learning and high evaluations; indeed, some studies show an inverse relationship. That is to say, the higher the evaluation, the less learning—as measured by test scores—has taken place. What does correlate consistently is story telling: Professors who

entertain with funny stories rank higher.

A colleague, new in her department and unfamiliar with the evaluation game, whose standards were exceptionally high and whose course content was extremely rigorous, was placed on probation because of low student evaluations. When she learned to manipulate her course, her evaluations went up, thus increasing the probability for retention. There are exceptions to these anecdotes,

(Continued on page 23)

(Continued from page 22)

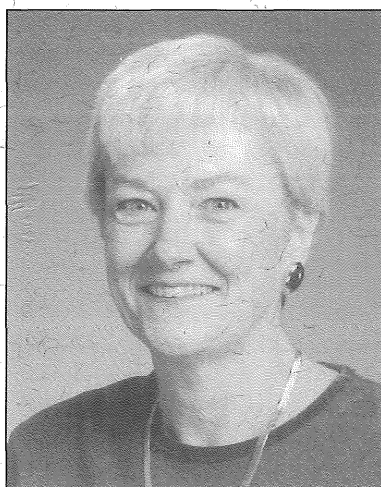
Reply to Michael Birnbaum's "Survey of Faculty Opinions Concerning Student Evaluations of Teaching"

certainly, but the truth of them abounds in and out of academic research literature.

Where a professor with a difficult course and demanding standards also enjoys high evaluations, a little investigation often reveals that students' shopping for an easy class avoid that professor or drop the class after reading the syllabus. This routine, then, eliminates low-achieving students who might write disparaging comments or bubble in low rankings. Some professors deliberately sound tough on the first day or two in order to reduce class size, leaving student scholars who value high standards and an increased level of learning.

Those of us who believe, as I do, that student evaluations of teachers are invalid, unfair, and thus harmful to education also understand that administrators have few alternatives. The disadvantages, though, far outweigh the advantages. Anonymous evaluations create an exercise of power over a precious faculty liberty, that of due process. If the university community retains this performance measure, student anonymity should be eliminated so that faculty members can face their accusers and even sue for libel where warranted. If students were held to the same standards in their evaluations as are faculty members in their grading practices, the entire system would be improved. ■

Gayle Vogt is a member of the Marketing/Business Writing Department. She currently serves on the Faculty Affairs Committee as well as the University Board on Writing Proficiency. Gayle was elected to the Academic Senate in 1990-93 and again from 1995-97.



Reply to Michael Birnbaum's "Survey of Faculty Opinions Concerning Student Evaluations of Teaching"

Mary Kay Crouch

Michael Birnbaum asked faculty questions about student evaluations; our answers indicate the negative impact we believe those evaluations have on our teaching. Because my responses to the survey generally reflect the majority opinion, in this reply to his findings, I want to bring up a related issue which he doesn't discuss specifically. Yet it seems implicit in the chemistry of faculty evaluation which Dr. Birnbaum does discuss: the design of student evaluations. For the way in which we ask our students to evaluate us is no doubt mirrored in their responses that we then have to defend and/or explain during RTP processes.

In my department, for example, students are given a sheet of nearly blank paper which is titled STUDENT OPINION FORM (note the word *opinion* as opposed to "evaluation").

Instructions for completing the page read: "Please state your opinion of the instruction in the class. Your statement is significant, for it will be placed in the instructor's Professional Review File." A large blank space follows and then, at the bottom of the page, students are asked to put a number to their opinion of the instruction: 5 = excellent; 4 = above average; 3 = average; 2 = below average; 1 = poor.

What these directions tell students more or less is this: "We want your opinion, not your evaluation of instruction. Say what you wish, and whatever you say will be significant to this instructor's evaluation by fellow professionals." Nowhere is there a place for students to comment on their responsibility to the course: "Did you attend class regularly? Did you do all the work expected of you?" No statement says, "Describe your own contribution to and involvement in this class." As Birnbaum comments, students can "give evaluative descriptions anonymously with *no incentive to be accurate and no penalty for libel*" (emphases mine). We put the burden for the course solely on the instructor and encourage our students to do the same. Are we surprised then, at the power students take when they fill out evaluations?

When I read some of my student evaluations—and I am one of those who does read them every



(Continued on page 24)

(Continued from page 23)

Reply to Michael Birnbaum-Mary Kay Crouch

term—I wonder at times if the students and I took part in the same class. Let me give an example. I regularly teach an upper division course which studies the structure of the English language.

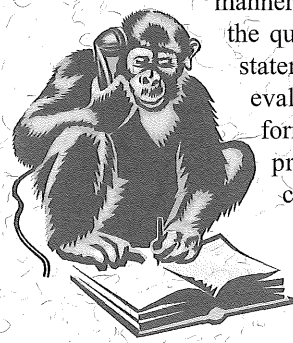


This course is required of students majoring in English and in Liberal Studies who intend to go into teaching. Students often find the course very difficult because they have little background in this area. Last spring, one student opined: “[Dr. Crouch] makes us feel guilty, incompetent, and ignorant for the things that we don’t know rather than praise us or reinforce what we do know.” I won’t give credence to that “opinion of instruction” by defending myself, but I can say that guilt, incompetence, and ignorance are not my goals for the course.

And note what else this comment implies, again something which the survey brings out. The student is not concerned about what s/he doesn’t know. Instead, the student expects praise and warm fuzzies for whatever knowledge he or she has. As a female instructor, I can’t miss the implied mothering/nurturing expectations that lurk behind this comment: “You’re a woman. Be a good mother to us, praise us, and then you’ll be a good teacher.” I teach students; I mother my daughter.

In the language course I’ve discussed here, student evaluations tend to run below the department average for those of us who teach it because most students entering credential programs must receive at least a C, some a B, or they will have to retake it. Since the course represents very high stakes for them, students often vent their frustrations about it and their grades through the end of term evaluation. Therefore, when I sit on our Department Personnel Committee, I pay attention to the classification of the courses for which instructors are being evaluated, especially in high stakes situations for faculty, e.g., RTP decisions or retention decisions for part time lecturers who teach writing. Teaching certain courses can be hazardous to one’s future at the university.

Michael Birnbaum’s survey reveals the divide that exists between what we believe in as teachers and what our students believe about themselves as learners. I hope in the future Dr. Birnbaum will look into the design of student evaluations, because the



manner in which the questions or statements on evaluation forms are presented certainly lends credence (or not) to student responses.

Perhaps faculty need to do a better job of educating our students about standards by designing evaluations which ask students to comment on the things we believe make up good teaching and learning in our disciplines. ■

Mary Kay Crouch is Associate Professor of English & Comparative Literature



Reply to Michael Birnbaum’s “Survey of Faculty Opinions Concerning Student Evaluations of Teaching”

D. V. Ramsamooj

I want to share the experience of the Department of Civil Engineering, over a period of eight years, of the relationship between student evaluations and student learning in different sections of the same class with common final examinations. A comparison of student evaluations with final examination test scores showed that the best student evaluations were obtained by the professor whose students learned the least.

In September 1988, my department instituted common examinations in Statics and Dynamics. There were as many as seven sections of each class at the peak enrollment period in the School of Engineering, with an average of about three to four sections per semester. Generally, the number of students in each section did not exceed 30, but occasionally a few faculty members taught double sections. The final examinations were prepared by all of the instructors. The student test papers were also graded by

(Continued on page 25)

(Continued from page 24)

Reply to Michael Birnbaum'

all of the instructors; the same instructor graded each test question for all sections. Students also evaluated each instructor and class section by means of student ratings.

There was a great disparity in the average test scores of individual sections — some sections did much better on the average than others. Often there was a difference of two letter grades between the average test scores of the highest and lowest sections of the same class. However, the average final grades assigned to each class, using the same grading scales for all classes, were adjusted because of political reasons and social promotion. This meant that students who would have failed if common standards were used, were allowed to pass the course.

The most interesting point is that there was one professor whose students' performance in both courses was always the lowest. But this professor consistently obtained the highest student evaluation in the Department.

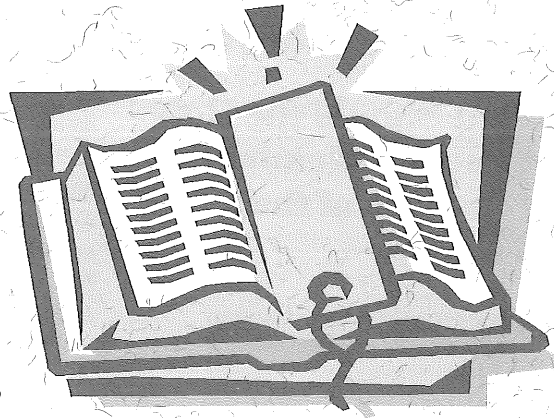
On the other hand, one professor whose students' performance was almost always the best, received only average student evaluations.

It is generally accepted that common examinations are the most objective means of assessing the quality of education. They are used nationally and internationally to judge student learning. Accordingly, it must be concluded that there was an inverse correlation between student evaluations and teaching effectiveness.

Student evaluation is important, as the students may have valuable information about their teachers. However, in a university where there is only a personal or professorial standard, the academic standards vary greatly. In such a system, the penetration into the subject and the academic standards may suffer. In such a system, high evaluations may represent lower standards rather than higher learning.

There are those teachers who students take when they really want to learn, and there are those who students take when they merely wish to pass the course. Low-performing students enjoy lower academic standards and tend to evaluate such easy teachers highly.

Whatever the level of preparation the students have for college, their level of performance at the time of graduation should be adequate in order to protect the public from malpractice or sub-professional work. As in any system, there ought to be a sufficient number of checks and balances. Professional accreditation helps to establish some minimum standard, but



anyone who understands the accreditation process, knows that it cannot guarantee adequate academic standards, as do, for example, the Engineer-in-Training (EIT) and Professional License (PE) examinations. One way of ensuring adequate academic standards is to have the final examination in each course prepared by an external agency. The test papers may be graded by the course instructors for economy and other practical purposes, but they should be open to review by authorized personnel. In this way, low performing classes may be identified and the educational process improved.

Without common examinations we would not have learned that the professor who obtained the highest evaluations also taught the students the least. With common examinations, the self-interests of students, the teachers, and society are aligned. ■

Editor's Comment

If you wish to add your own replies to Birnbaum's thesis and/or to the replies, please go to the discussion group at <http://faculty.fullerton.edu/senatenews>.

D. V. Ramsamoj is a Professor of Civil and Environmental Engineering

