AC 2009-1742: IMPROVING SCORES ON COURSE EVALUATIONS: EXPERIENCED FACULTY TELL WHAT WORKS

Edward Gehringer, North Carolina State University
Ed Gehringer is an associate professor in the Department of Computer Science and the Department of Electrical and Computer Engineering at North Carolina State University. He has been a frequent presenter at education-based workshops in the areas of computer architecture and object-oriented systems. His research interests include architectural support for memory management, garbage collection, and computer-supported collaborative learning. He received a B.S. from the University of Detroit(-Mercy) in 1972, a B.A. from Wayne State University, also in 1972, and the Ph.D. from Purdue University in 1979.

Improving Scores on Course Evaluations:
Experienced Faculty Tell What Works

Edward F. Gehringer
North Carolina State University
efg@ncsu.edu

Abstract

On many campuses, student course evaluations are the primary means of evaluating teaching, and can have an impact on performance reviews, tenure, and promotion. Thus, faculty desire to improve their scores. But how can they? This paper considers the experience of several faculty who managed to raise their scores significantly. Some of them did so by taking steps that would improve their course or their attentiveness to students. Others simply "dumbed down" their courses. There is an extensive literature on factors affecting course-evaluation scores. It confirms some of what our informants told us, and calls into question some of their observations. In particular, we discuss the issue of leniency at length, and conclude with three recommendations to new engineering instructors on how to improve their own scores. The first is to be more attentive to students and their needs. Next, an instructor should focus on what is being learned rather than what is being taught. And finally, faculty should avail themselves of institutional support for improving teaching.

1. Introduction

In most engineering schools, except for research, teaching is the most important factor on which reappointment, promotion, and tenure are based. And teaching is most often evaluated using student course evaluations. This places faculty in a delicate position, a reciprocal relationship between their students and them, in which each party is assessing the other and influencing their subsequent advancement. For this reason, student course evaluation is one of the most contentious issues [1] in all kinds of academic departments and all kinds of institutions. Instructors rightly point out that other factors should be considered when determining the efficacy of teaching. Thus, in recent years, peer evaluation of teaching [2] has taken its place alongside student evaluations in determining teaching competence. But faculty remain uneasy about their student evaluations, regarding them almost fatalistically as something potentially important over which they have little control.

The goal of this work is to present the cases of a number of engineering and computer-science faculty who did manage to improve their scores, in hopes that they can serve as role models. We identify several aspects of their teaching where change made a difference. Then we compare their observations to what the published literature reveals. We conclude with recommendations for faculty who want to improve their scores.

Our respondents came from two mailing lists, the Engineering Technology listserv, etd-l@listproc.tamu.edu, serving ASEE's Engineering Technology division, and the SIGCSE members list, SIGCSE-members@LISTSERV.ACM.ORG, serving the Special Interest Group on Computer Science Education of the Association for Computing Machinery. The author posted

on these lists in October 2008. The number of responses was not large, but that may just reflect the fact that few instructors have seen much change in their evaluations. A total of 14 usable responses were received. Three of these were from instructors who said their scores have always been high, and offered tips on how they did it. The remaining eleven were from instructors who said their scores rose after they changed something about the way they taught.

2. What our respondents told us

No two of our respondents gave exactly the same answer to how they improved their scores. To bring some order to the discussion, the responses were grouped into categories like interest, organization, and manipulation. Interestingly, no single response included items from multiple categories. A discussion of each category follows.

Show interest in students and their feedback. Four instructors said that showing interest in what students said about the course was the key to getting good evaluations. As one respondent put it,

> Two tips: both effective. The first was one my unnamed chairman at a major institution in Boston showed me. His evaluations were uniformly excellent. He handed them out, then stood at the front of the class, and had students bring the evaluations to his desk. He would pick up each of them and read them - silently, but with great attention. It only took a moment for students to understand how important their evaluations really were. The second tip is to show students verbally, and by your actions, how important the evaluations are. Spend time at the *start* of class telling them how the evaluations are used, hand them out, and then disappear for 10 minutes. Rather than having the students free to leave class after spending two minutes dashing something off, they wait around for your class. You can pick what you are going to discuss: perhaps going over the sample exam you handed out last week.

The first tip would be against policy at many institutions, because it compromises the anonymity of responses. According to the respondent, it was against policy at his institution too, but was nonetheless very effective. Another variant on this preserves anonymity:

> I'm not sure if anything I do has improved my evaluations, but there are certain recommendations that I give to new faculty, and I follow these myself. They assume that the instrument is an in-class paper-based evaluation. 1. Give evaluations at the start of class. That way there is incentive to write a lot of comments since the longer the evaluations take, the shorter class will be. 2. Give the class advance warning, and try to focus their attention.   My standard rap goes something like this. "The College uses the information from evaluations in the reappointment and tenure process. It is one way that students have a voice in personnel decisions. I also find your comments to be very helpful as I plan changes to the next offering of this course. So I have two requests. First, when you are doing the evaluation, really read the questions and answer them—don't just go into autopilot and give the same answer for everything. Second, during the next week I'd like you to think about a few things on your back burner. ‹fill in here a few questions you have, perhaps about the textbook, or topic coverage. I think the actual things you suggest are less important than the fact that you are suggesting some things and asking them to think about the course› Think about how you would complete these sentences: this course is really great except for …"

Another aspect of showing interest in students is talking to them:

> I really don't think my course evaluations have changed. I have generally always received excellent responses from my students. I do try to know my students and talk with them as much as possible (both in and out of the classroom).

John Edgerton, Professor Emeritus of Engineering and Science Technology at the University of Akron came at it from a slightly different angle, focusing on respect for students.

> I have rarely made much effort to improve my student course evaluations. They have always been relatively high. Recent scores mathematically could not get much higher. What I have always done, even before student course evaluations, is: 1. Treat the students with complete respect. 2. Work hard and make sure I know what I am talking about. 3. Encourage students to ask questions, and treat all questions and questioners with complete respect. 4. Have a grading system that is fair, transparent, and objective; admittedly not difficult for my subject area, Electronic Engineering Technology, but I am sure do-able for any subject. By the way, I do grade spelling, grammar, content, and clarity of expression on written work I do assign. And I clearly explain why I do that. 5. Set high standards and encourage all to achieve them. 6. Be approachable with posted office hours that I meet and exceed. 7. Treat students as the customers they are; ultimately they pay my salary.

Be organized. Five of the other respondents wrote about various aspects of course organization. The first emphasized creating a set of FAQs and checklists:

> Each course has a web site that contains all the materials for students to download, FAQs for each assignment and exam, and other helpful material. The web site is updated each time I'm going to teach the course again. Many of my students are distance students and when you combine that with my being an adjunct and, therefore, not having regular on-campus office hours, virtually all student communication occurs by email. I use the following technique for each course I teach.
>
> *Technique:* I've learned to turn email communication into an advantage. Whenever a student asks a question I provide a carefully thought out reply to the student and then put a copy of his or her email and my reply in a "questions asked" folder for the course in question. (If I use a grader or assistant I ask them to forward non-routine questions to me for this purpose, including their answers.) Sometimes I find these questions very thought provoking - helping me see that my lectures or assignments were not clear, opening up new avenues of thought or new perspectives on the subject at hand, giving me new concepts to use in my courses, or occasionally pointing out to me that my own understanding is muddled or incorrect or incomplete.
>
> When preparing for the next offering of any given course, I open up the "questions asked" folder from the previous offering and do several things with these questions and answers:
>
> - In some cases I research the issues at hand further, because the student question has made it clear that this is appropriate.
> - I update the FAQs that I have provided them for each assignment and each exam.
> - I update the course assignments to use clearer wording for any areas where there was student confusion.
> - I update the course slides/notes/reference lists in a similar manner, introducing new ideas or discussion points, refining things, etc.
> - I maintain a bank of exam questions and I update these, if needed.
> - As part of my "assistant/grader training program" I have each assistant or grader study these FAQs to better understand the subtler aspects of each assignment.
> - I also have a set of guides and checklists for grading each assignment. These are updated each time and serve as guides for myself and my graders/assistants.

The other comments in this category were much shorter. One said his "course organization" score was much higher in courses where he used a textbook. Another said he "started out with firm rules on the first day. Made a good first impression …" The fourth referred to using WebCT Vista to give the students access to notes, PowerPoint slides, and grading rubrics. The fifth one said that, while he didn't have course evaluation scores yet, he was writing his own

textbooks and the students seemed to respond much better to it.

Special training. Two respondents cited special programs as the reason for their improvement. The first was a teaching-effectiveness seminar:

> I attended an "Essential Teaching Seminar" sponsored by ASME, hosted at MIT and conducted by former West Point faculty in June 2006. This three-day interactive training workshop demonstrated proven techniques for effective learning in engineering courses. An instructional improvement was reflected in higher student ratings for the course in Fall 2006 than the previous semester.

The second went to a speech pathologist to improve his delivery:

> I would like to remain anonymous, but I have been evaluated for over 17 years teaching in the Associate Degree Electronic Technician Program. For over 15 years my students have commented that I talk too fast. Two years ago I had some extra money, so I paid around $75 out of my own pocket for four 45-minute sessions with a speech pathologist. She taught me to relax, reach for a drink of water, or walk across the room while I allowed time for my students to think.

Manipulation. All of the remaining five responses can be classed as manipulation of some sort, in that they were designed to increase evaluation scores without improving the quality of instruction. The first of these, which at least was not detrimental to learning, said simply,

> Bring food. Cookies, pizza, etc. If you think I'm kidding, I'm not.

"Dumbing down" the course is another kind of manipulation:

> What I did was announce a 10 point curve on the most recent test and dropping of the lowest 2 homework scores immediately before the evaluation.

One instructor said she intentionally stopped making the students do any higher-order thinking:

> I actually increased my student eval scores by about a half a point in a single term. I went from an average score in my various classes of around 3.6 up to 4.1 or 4.2 during 2007-2008 academic year. I followed the advice of a colleague and "make them feel like they're learning a lot without their having to actually do any learning." I implemented this via the following tactics: 1) Completely stopped giving any open-ended design problem type assignments. All assignments were exercises from the end of the chapters in the text. All assignments were at the lowest order "Knowledge" process in the cognitive domain 2A) no longer required the students to access their textbooks at all for any reason except to read the particular assigned exercises from the back of the chapters 2B) Explicitly wrote every word I wanted them to have in their notes onto the board, moved the affective domain down to the lowest "Receiving" level …

> Another point of interest, our student evals also include a question, "How would you rate your level of learning in this course compared to other courses?" My evals were super high, around 4.6 or 4.7 out of 5 on this metric.

The other two instructors said they had heard that these kinds of manipulation had worked, but said they had not personally used them.

3. What the experts have to say

Over the past eighty years, hundreds of papers have been published on factors affecting student course evaluations. Most of them are not very helpful to the present discussion, because they

look at the effect of factors that are not under control of the instructor. An instructor cannot change his class size, his age, or gender. She can't discard a captive audience by making a required course an elective, nor replace cocky sophomores with more appreciative seniors.

But research does reveal that student evaluations are, by and large, reliable [3]. That is to say, they correlate well with other measures of effective teaching. Students are "qualified" to rate their instructors by virtue of the fact that they are with the instructor for an entire semester, and thus are better acquainted with the instructor's teaching than any outside evaluator would be.

So how can instructors improve their scores? Theall [3, p. 47] cites a variety of approaches that can be categorized as *increasing interaction with students*. Instead of focusing solely on presenting material well, they can leave time for questions and discussion of course material.

Beyond that, instructors should *monitor and reflect on* their teaching [4]. They should consider whether their actions are helping to achieve their learning goals for their students. To determine this, they should consider a variety of evidence, notably student reactions to their teaching. This may come from direct personal contact with students, overhearing conversations among students, or from homework or test results. If the experience of students is unacceptable, they need to be prepared to make changes to their methodology and the material they cover.

But, what changes should instructors make? Unless they reflect on their past experience, the changes they make may not lead to better results. A good way of learning to do this is to work with experienced teachers on analyzing the evidence and deciding on a course of action.

Ramsden [5] says that faculty members go through three stages in acquiring skills as a teacher. In the first stage, they concentrate on relating course material to their students. Doing a good job means giving a good lecture. They are focused primarily on their own actions. In the second stage, they concentrate on orchestrating student effort toward producing good work products. This shifts the focus toward student actions. In the third phase, they view education as a cooperative venture, with students and instructor working together to expand students' knowledge and understanding.

A number of studies indicate that a *midterm evaluation, followed by consultation*, can be effective in improving course-evaluation scores. Kulik [6] cites several studies showing that ratings went up after a midterm evaluation (by 0.1 rating point), and after midterm evaluation + consultation (by 0.3 rating point). Emerson et al. [7] report on another study whose authors studied twelve instructors, six of whom had a post-midterm consultation and six of whom didn't. The instructors who went through the consultation received higher ratings from the students, and their students got higher scores on the final exam.

Why is consultation necessary to achieve good results? Lang and Kersting [8] cite research showing that instructors, confronted with feedback that is less positive than desired, tend to work harder. This increases their scores by a small amount. Dissatisfied with just a minor improvement, they instead decide to "work smarter," and alter their teaching methods. But, since they have little expertise in this area, the alteration may hurt instead of help. Groping around, they try one ineffective approach after another, ultimately diminishing the gains they achieved immediately after receiving feedback.

4. The leniency debate

Recall that several of our respondents said that giving higher grades for less work would result in higher evaluation scores. Their sentiments are echoed by Glymour [9], who cites the case of an assistant professor of philosophy who might have been denied tenure over poor teaching, but increased his scores dramatically by promising almost all of the students As. It is widely believed that there is a direct relationship between favorable evaluations of the students by the instructor and favorable evaluations of the instructor by the students. But what does the research say?

There is empirical support for the notion that leniency influences course evaluations. Greenwald and Gillmore [10] observed a positive relationship between grades and course-evaluation results in several hundred courses at the University of Washington in 1993–94. They attempted to explain this correlation by several hypotheses, e.g., that better teachers cause students to learn more and thus get higher grades, that more highly motivated students do better in courses and like their instructors better, or that "students give high ratings in appreciation for lenient ratings." They found that the latter hypothesis to be the only one supported by the data.

However, when studying the relationship between grades and evaluation results, Greenwald and Gillmore did not control for course level or students' interest in taking the course. Marsh and Roche [11] found that these two factors could explain most of the positive correlation between high grades and high evaluations. Moreover, studies done in multisection courses show that high achievement is correlated with high evaluations (that is, that sections which rate the instructor better also get better grades on common exams). If high achievement is positively correlated with high evaluations, and high evaluations are positively correlated with grading leniency, then grading leniency must be positively correlated with high achievement—which is implausible.

A study by Centra [12] included over 50,000 college courses whose teachers used the same evaluation instrument (Student Instructional Report II). It controlled for factors like class size and perceived amount of learning. Amount of learning was measured by such factors as increasing interest in the subject matter and how much the course had helped students think independently. Learning outcomes were strongly correlated with course evaluation, but after controlling for learning outcomes, he found that expected grades were not positively related to evaluations. In fact, in natural sciences, the highest-rated courses were some of the ones with lower grades. Courses that were either "too hard" or "too easy" were rated lower than courses between these extremes.

Eiszler [13] comes at the question from a different perspective—the change in grades and evaluations over time. Using data from more than 37,000 course sections offered between 1980 and 1999, he found that during the 1980s, the percentage of students expecting As or A–s held steady, as did the average scores on course evaluations. But in the 1990s, the percentage of students expecting As and A–s increased by more than 10%, and the average teaching-evaluation rating increased by 0.1 point. Now it is possible that improvements in teaching methods or technology caused students to like courses better by the late 1990s, and thus to invest more work and earn better grades. However, he found that course attractiveness did not increase: in the early 1980s, 70% of students would have taken the course in question if it had not been required, but by the end of the 1990s, only 59% would have.

Another angle is provided by a study by Griffin [14], who tried to measure students' perceptions of grading leniency and correlate them with evaluation results. He asked students whether they believed that the instructor was "a lenient/easy grader." He also asked students what grades they expected in the course, and what grades they thought they deserved. This was used to derive a metric based on expected grade – deserved grade, another measure of leniency. The results showed that leniency was significantly correlated with 11 of 12 ratings items, such as overall instructor rating, overall course rating, course organization, ability of students to seek help, and fairness of instructor. The only ratings item that was *not* significantly correlated with leniency was whether the course content was worthwhile. The magnitude of correlation was such that the difference between the most lenient and least lenient was 0.48 rating point; e.g., if the least lenient instructor expected a rating of 3.50 on some question, the most lenient could expect 3.98.

So, can instructors expect to improve their ratings by becoming more lenient? Most of the research seems to say no, but depending on how one looks at the evidence, one could reasonably say yes. Thus, the answer seems to be a definite "maybe."

5. What's an instructor to do?

Faced with this evidence, what should an instructor do to improve teaching-evaluation scores? First, it seems inadvisable to "dumb down" the course. Not only is the evidence ambiguous on whether this will help; it is likely to lead to students learning less and diminish the value of an education at your school.

The first lesson is to be more attentive to students and their needs. Seek feedback from them, and pay attention to what they are telling you. One well known way of doing that is the "minute paper": At the end of a class, ask students what was the most important thing they learned during the class. The minute paper is the best known of a set of Classroom Assessment Techniques [15], which keep instructors in touch with what their class is learning. Another example is the "one-sentence summary," which asks students to write "Who does what to whom, when, where, how, and why?" ("Who" can also be an non-human agent, such as a force or a virus.) CATs help identify what material the students are having trouble with, and allow the instructor to adapt his/her teaching to the students' learning styles. The Cross and Angelo book [15] describes forty different techniques that can be used in all kinds of classes.

The second lesson is to focus on students, not on your own teaching. Do not ask, Did I present this material well? Rather ask, Were students following everything I said? Often student learning is greater when less is covered during a class period. Covering less in class is not the same as "dumbing down" the course, because students can be expected to read the textbook outside of class to learn the extra material. You can even check to make students have read the material by creating a short machine-scorable quiz to be taken online before coming to class. Class material is learned more easily and remembered longer when classes are interactive, not when the presentation is most polished.

Finally, you should avail yourself of the opportunity to get feedback on your teaching and follow through on making changes as a result of it. If your school or department offers

midterm evaluations, take advantage of them. Read the feedback carefully. The advantage of getting feedback during a term is that you have the opportunity to adapt to it. For that reason, it is known as "formative" assessment, rather than the "summative" assessment provided by end-of-term questionnaires. If your department does not have such a policy, or if you want to get feedback more frequently, be sure to check out SALG (Student Assessment of Learning Gains), http://www.salgsite.org.

Most universities now have Teaching and Learning Centers (which may be known by other terms, such as Center for Faculty Excellence). The staff of these centers are usually willing to do one-on-one consultations with you about your teaching. You bring the student feedback to your assessment, and the staff professional will typically discuss it with you for 30 to 60 minutes. Sometimes repeat visits can be scheduled. Evidently, very few faculty take advantage of such opportunities. Directors of such centers have indicated to me that no more than 2%–7% of instructors seek their services during an academic year. Taking that step yourself can be your advantage when it comes to being a successful teacher.

5. Summary

Student course evaluations are generally reliable instruments that can give you valid feedback on your teaching. While there is disagreement over whether they favor more lenient instructors, there is no doubt that instructors can improve by studying the feedback carefully. Some of the most important steps you can take toward improving your own performance are interacting more with students, paying more attention to their needs than to your presentation, and seeking out feedback and then using it as a guide for how to improve.

Bibliography

[1] Harrison, P. D., Douglas, D. K., and Burdsal, C. A. The relative merits of different types of overall evaluations of teaching effectiveness. *Research in Higher Education* 45:3 (May 2004), pp. 311–323.

[2] Chism, Nancy Van Note, *Peer Review of Teaching: A Sourcebook*. Bolton, MA: Anchor Publishing, 1999.

[3] Theall, Michael and Franklin, Jennifer. Looking for bias in all the wrong places: a search for truth or a witch hunt in student ratings of instruction? *New Directions for Institutional Research* 100 (Spring 2001), pp. 45–56.

[4] McAlpine, Lynn and Weston, Cynthia. Reflection: issues related to improving professors' teaching and students' learning. *Instructional Science* 28 (2000), pp. 363–385.

[5] Ramsden, Paul. *Learning to Teach in Higher Education*, Routledge, 1992.

[6] Kulik, James A. Student ratings: validity, utility, and controversy. *New Directions for Institutional Research* 100 (Spring 2001), pp. 9–25.

[7] Emerson, J.D., Mosteller, F., and Youtz, C. Students can help improve college teaching: a review and an agenda for the statistics profession. In Rao and Székely, eds. *Statistics for the 21st century: methodologies for applications of the future*, CRC Press, 2000, pp. 145–172.

[8] Lang, J.W.B. and Kersting, M. Regular feedback from student ratings of instruction: Do college teachers improve their ratings in the long run? *Instructional Science* 35:3 (May 2007), pp. 187–205.

[9] Glymour, Clark. Why the university should abolish faculty course evaluations. Unpublished draft memorandum, November 2003, www.hss.cmu.edu/philosophy/glymour/glymour-universityFCE2003.pdf

[10] Greenwald, A.G. and Gillmore, G. M. Grading leniency is a removable contaminant of student ratings. *American Psychologist* 52:11 (November 1997), pp. 1209–1217.

[11] Marsh, H. W. and Roche, L. A. Effects of grading leniency and low workload on students' evaluations of teaching: popular myth, bias, validity, or innocent bystanders? *Journal of Educational Psychology* 92:1 (March 2000), pp. 202–228.

[12] Centra, John A. Will teachers receive higher student evaluations by giving higher grades and less course work? *Research in Higher Education* 44:5 (October 2003), pp. 495–518.

[13] Eiszler, Charles F. College students' evaluations of teaching and grade inflation. *Research in Higher Education* 43:4 (August 2002), pp. 483–500..

[14] Griffin, B. W. Grading leniency, grade discrepancy, and student ratings of instruction. *Contemporary Educ. Psychology* 29:4 (October 2004), pp. 410–425.

[15] Angelo, Thomas A. and Cross, K. Patricia, *Classroom Assessment Techniques*, 2ed., Jossey-Bass, 1993.