

HOW TO IMPROVE YOUR TEACHING EVALUATIONS WITHOUT IMPROVING YOUR TEACHING¹

IAN NEATH

Purdue University

Summary.—The current increased interest in evaluating the teaching of college and university faculty has made course evaluations even more important to the careers of academic faculty. The most important use of teaching evaluations is to guide instructors in revising courses and altering teaching methods to become more effective educators. A second use is becoming more frequent: comparing scores on teaching evaluations across faculty to make major personnel decisions, including hiring, firing, promotion, and tenure. The soundness of these decisions depends upon the validity of the evaluation instrument, and even the most reliable and consistent instrument can easily be compromised if administered inappropriately. Based on a review of the literature, I list 20 ways to compromise teaching evaluations and improve ratings without improving teaching.

More and more college and university administrators are requiring (or 'strongly encouraging') faculty to administer teaching evaluations. With declining revenues, state legislators are demanding proof from institutions which receive some portion of their funding from the taxpayers that students are being effectively taught. With income from tuition and governmental funds often tied directly to bodies in classrooms, administrators want to hire and retain the best instructors to maintain or even increase enrollment and retention. Worried about continuously rising tuition and other costs, parents and students are increasingly concerned about the quality of the education being provided and are demanding more for their money. None of these concerns are inherently bad, but all serve to focus attention on course and teaching evaluations. Instead of publish or perish, faculty can now publish *and* perish—if their ratings on teaching evaluations are not sufficiently high.

Students are in a unique position to evaluate some aspects of teaching effectiveness; collecting this information and interpreting it appropriately can serve all three needs listed above. There are two main uses of teacher and course evaluations collected from students. *Formative* evaluations have, as their goal, revising courses and altering teaching methods to promote more effective teaching. *Summative* evaluations, on the other hand, are used to compare faculty to aid in decisions of hiring, firing, promotion, and tenure. The former use, although not necessarily as easy to implement as one might

¹Address enquiries to Ian Neath, Department of Psychological Sciences, 1364 Psychological Sciences Building, Purdue University, West Lafayette, IN 47907-1364. e-mail: neath@psych.purdue.edu

think (Canelos, 1985), has potential for improving undergraduate education (Marsh & Roche, 1993). There are no valid arguments against conducting formative teacher and course evaluations, and there is no defense for not instituting and encouraging such programs.

The second use—comparing faculty on the basis of their scores—is the focus of this comment. The easiest and least expensive way to collect summative evaluations of teaching effectiveness is to form a committee to write a questionnaire for teaching evaluation. Although the questions will usually have high face validity—they look like they measure teaching effectiveness—there is often no evaluation of construct or convergent validity or of other measures that might indicate a sound instrument. The only analysis typically reported is the over-all score, usually expressed as a percentile. For example, because Professor X ranks in the 55th percentile on a particular teaching evaluation and Professor Y ranks in the 45th, Professor X will get a merit raise this year whereas Professor Y will not. The soundness of this decision rests on the validity and reliability of the measures of teaching effectiveness and on the manner in which the evaluation questionnaire was administered.

Theoretically, a teacher evaluation rating should reflect the extent to which the instructor informed the student about the particular subject matter and fostered further development of educational goals and intellectual skills. In practice, there is no agreed-upon definition or criterion measure of effective teaching. Typically, some composite measure is used that includes items such as student-reported gains in factual knowledge, gains in thinking or decision-making skills, and discovering implications of the course material (e.g., Cashin & Downey, 1992). Thus, evaluations of teachers usually rely on self-reported improvements by individual students. Because evaluations of teaching are based on subjective feelings of improvement, they are susceptible to many outside influences that may not be related to being an effective teacher. The validity of even the most reliable instrument can be compromised quite easily if it is administered or interpreted in an inappropriate way.

What follows are 20 tips that Professor Y could use to increase teaching effectiveness ratings.

Tip 1: Be Male

The effects of instructors' gender on evaluation ratings are complex and interact with a variety of other variables (e.g., Beck & Drew, 1989; Feldman, 1993; Lueck, 1993). Nonetheless, a preliminary conclusion appears to be that it is better to be male than female. Changing your gender, if female, can boost your effectiveness ratings. Students often expect more support from female faculty than from male faculty, and, when this extra effort is not forthcoming, students often downgrade their ratings of teaching effectiveness

(Bennett, 1982; Unger, 1979). If you are female, do not be very demanding of your students; students tend to be more critical, particularly on items measuring faculty's availability and course stimulation when the instructor is female (Hearn, 1985). This effect is compounded if you teach a technical course, particularly to liberal arts majors. If you do happen to be female and, for whatever reason, do not wish to become male, all is not lost. The Bem Sex-role Inventory is a scale that measures the preponderance of masculine and feminine traits in an individual. Study the scale and learn how to be less feminine and more androgynous; your ratings will improve (Bray & Howard, 1980).

Tip 2: Be Organized

Be as organized as possible. One prominent question on teaching evaluation forms asks students to rate how organized the instructor was. Although one cynic has suggested that "well-organized garbage still smells" (cited by Cashin & Downey, 1992, p. 564), your score on this dimension still factors into your over-all effectiveness rating. Start class on time, end class on time, bring extra chalk or overhead pens to class, and keep all of your lecture notes in a three-ring binder with neatly punched holes.

Tip 3: Grade Leniently

Higher grades correlate very highly with individual ratings of teaching effectiveness (DuCette & Kenney, 1982; Edwards & Waters, 1984; Hoffman, 1983; Meredith, 1982), and grading harshly is a sure way to lower evaluations (Chacko, 1983; Vasta & Sarmiento, 1979). The central issue in evaluating teachers is not whether there is a relationship between grades and ratings of effectiveness but rather interpreting that relationship. While it is clear that effective instruction, which should increase grades, can give rise to positive evaluations, lenient grading standards can also give rise to more positive evaluations. Although correlation does not imply causation, in this case the result—higher ratings—is the same in both cases. Opt for lenient grading and then defend grades on the basis of effective instruction; your ratings will back you up.

Tip 4: Be Present at Your Evaluation

Make sure that you administer your own teaching evaluations. Many of the recommendations below require your presence rather than that of a neutral proctor. In addition, you will have control over when they are administered, e.g., after a particularly entertaining lecture rather than after a poor one, how the instructions are delivered, and whether the mood of your students is favorable for obtaining inflated ratings.

Tip 5: Administer Ratings Before Tests

Have yourself evaluated immediately prior to a midterm or final exami-

nation, not afterwards (Dizney & Brickell, 1984). Many students do not perform as well as they think they should on tests, and this disappointment can reduce their ratings of your effectiveness. Moreover, administering the evaluation immediately after the students have reviewed course material for a test can inflate their subjective feeling of how much they have learned and, by extension, of your effectiveness in educating them. If possible, inform students of their potential grades in the course prior to administering the evaluation, particularly if the grade distribution for your course is lenient (see 'Grade Leniently' above).

Tip 6: Provide the "Correct" Instructions

Make sure that you inform your students that the purpose of the course evaluations is for administrative purposes rather than for actual course improvement (Aleamoni & Hexner, 1980; Tetenbaum, 1977). One reason for this effect of instruction may be that students often feel a lack of confidence in the use of evaluations for summative purposes and inflate their ratings to compensate (Dwinell & Higbee, 1993). Deliver the instructions in a relaxed, even humorous way so as to induce a positive set (Dizney & Brickell, 1984). One instructor I know has honed his instructions to a fine point: "These evaluations will be used by the administration to determine whether my children, ages 4 and 6, will eat for the next 6 months."

Tip 7: Use Appropriate Answers

Unfortunately, the basic format of the evaluation questionnaire does not seem to affect ratings, whether the comparison is open-ended or fixed-alternative (Braskamp, Ory, & Pieper, 1981) or whether it is paper-and-pencil or computerized (Cates, 1993). All is not lost, however. Rating scales have answers ordered along a dimension, such as strongly agree, agree, neither agree nor disagree, disagree, strongly disagree. Fortunately, these are the types of scales most often used, but more importantly, they afford more robust halo effects than other scales (Edwards & Waters, 1984). A halo effect is the tendency to overrate other qualities if some abilities are also rated highly. If possible, then, begin the questionnaire with a statement that all students strongly agree with; the halo effect may carry over to questions about your actual teaching ability.

Tip 8: The Smaller the Better

Teach small classes; do not teach large classes. There is often an inverse relationship between class-size and ratings on teaching evaluations, and this effect is larger for ratings that measure the instructor's interactions and relationships with the students (Feldman, 1984). There are a variety of reasons for this relationship, including the fact that smaller classes tend to be higher-level courses, tend to afford more faculty-student interaction, and tend to be

accessible only to majors. In addition, the smaller the class (particularly if there are 25 or fewer students), the more likely any bias factor(s) will influence the ratings and mask your true teaching ability (Feldman, 1977).

Tip 9: The Higher the Better

Teach only high-level courses and avoid all low-level courses (Neumann & Neumann, 1985). Higher-level courses tend to have smaller enrollments, more majors, and typically are electives. Moreover, student interest is usually higher, or they would not have pursued the subject as far.

Tip 10: Cross-Listings are Bad

Do not allow nonmajors to take your course. For example, for a psychology course, require all students to be psychology majors and prohibit students from other departments and particularly from different schools. Students generally rate classes in their major higher than classes outside their major (Gilmore, Swerdlik, & Beehr, 1980; Romeo & Weber, 1985).

Tip 11: Required Courses are Bad

Do not teach required courses, especially a statistics course (Brandenburg, Slinde, & Batista, 1977; DuCette & Kenney, 1982). Required courses, although educationally important, tend not to be well-liked even by students within the major. As Reynolds (1977) stated in his title, "Students who haven't seen a film on sexuality and communication prefer it to a lecture on the history of psychology they haven't heard."

Tip 12: Gimmicks are Good

Show lots of films, perform lots of demonstrations, and use the latest technology. For example, although using computers on a particular task may not lead to better learning than using the old-fashioned paper and pencil, students prefer it and this shows up in teaching evaluations (Rohde, 1993). Also, do not be concerned about whether the film is actually relevant: you can get a substantial increase in ratings following the showing of a basically unrelated videotape (Silvestro, 1979).

Tip 13: Entertain

If you have a flair for the comedic, go for it. Results show that ratings of over-all effectiveness of a teacher are positively related to the use of humor (Van Giffen, 1990). Do not worry if the entertainment value reduces or eliminates the educational value. For example, Naftulin, Ware, and Donnelly (1973) report a study in which an actor was hired to play the part of a visiting instructor, Dr. Fox, M.D. He delivered a lecture to medical doctors which was very enthusiastic, expressive, and entertaining. His ratings of teaching effectiveness were very impressive, sufficiently high to win a teaching award at most institutions. It did not seem to matter that the lecture was almost entirely content free.

Tip 14: Fulfill Students' Expectations

Although the prior expectations of students can influence ratings of a teacher's effectiveness (Barke, Tollefson, & Tracy, 1983; Bock, 1979; Terry & McIntosh, 1988), it can be difficult to take advantage of this because evaluating preexisting opinions may have an unwanted side-effect. Parish and Campbell (1977) had all students fill out evaluation forms after taking a course, but half of the students also filled out precourse evaluations. Although pre- and postcourse evaluations were correlated, the students who had filled out the precourse evaluations were much more critical of their instructors on their postcourse evaluations than students who did not fill out the precourse evaluation form. Perhaps an easier way to take advantage of students' expectations is to find out informally what the reputation of a course is and teach the course to be consistent with or even easier than the reputation.

Tip 15: Teach Only Male Students

There is some suggestion that female students may be more fair and honest than male students and that they do not believe as strongly that their grades will be jeopardized by their evaluations (Taylor & Ricketts, 1981). Males are more likely to inflate ratings than females because they are less prone to honesty and more likely to believe that retaliation by the instructor is possible. This undoubtedly interacts with gender of instructor, however, where being male is better (see above). Just to be safe, though, do not let women take your course.

Tip 16: Be Like Your Students

Adopt and promote values, attitudes, and beliefs consistent with those of your students, regardless of your actual beliefs. Students generally give higher ratings to an instructor when the students' attitudes are congruent with the instructor's perceived attitudes (Follman, 1975; Hofman & Kremer, 1980; Kovacs & Kapel, 1976; Morstain, 1977). Thus, you might be more liberal if you teach at Berkeley but more conservative at Baylor. If you are successful, students are more likely to identify with you, and the more they do this, the more they will rate the course more favorably, perceive greater progress on their part toward course-related objectives, and report higher motivation to learn (Thomas, Ribich, & Freie, 1982).

Tip 17: Teach What They Want How They Want It

Change your teaching style to that preferred by your students (Cooper & Miller, 1991; Drummond & McIntire, 1977). You might not be as successful when adopting a new or foreign teaching style every semester, but your ratings may very well improve.

Tip 18: Pick Successful Students

Set high minimum grade standards on prerequisite courses as a criterion for enrollment in your course. This will prevent poor students from enrolling and permit only those with high GPAs to take your course. Successful students often perceive internal factors as more important causes of their performance and unsuccessful students often perceive external factors as more important causes of their own performance. Internal attributions for performance are often significantly related to positive course evaluations, and external attributions for performance have been shown to be related to negative course evaluations (Ames & Lau, 1979; Arkin & Maruyama, 1979). Thus, a good student will not attribute failure to you.

Tip 19: Evaluate Everyone

Make sure that *all* faculty at your institution are evaluated, especially those faculty who do not want to be. After implementing this change, the distribution of scores will become much wider, particularly at the lower end, and your percentile score will improve. The reason is straightforward: those faculty who choose not to be evaluated typically receive worse evaluation ratings than those faculty who choose to be rated (Moore, 1978). In arguing for this change, you will also have the advantage of the moral high ground.

Tip 20: Examine Your Evaluation

A final tip—the only one I have personally used and verified—is to examine your evaluation data for accuracy. One semester, as I was looking at the teacher evaluation scores for my course on ‘Introduction to Cognitive Psychology,’ I noticed several forms were from a botany course. These students apparently did not think much of that particular course, but because they were included in my summary data, my ratings were lowered. When I removed the botany ratings, my ratings improved.

CAVEATS AND CONCLUSIONS

It is clear that the factors listed above, although not necessarily related to the ability of a person to educate students, can elevate ratings on evaluations of teaching. Despite the fact that they are not all equally effective, there is sufficient evidence to suggest that most of them can influence individual ratings independently of the questionnaire. If you were Professor Y, the teacher judged less effective than Professor X, and you employed a combination of these techniques, the literature on course evaluation suggests that you should now be rated as a more effective teacher than Professor X.

Obviously, you should not use these techniques when conducting evaluations for formative purposes because they greatly distort students’ ratings of your effectiveness and you might start thinking that you were an outstanding educator. Ideally, then, you should have yourself evaluated twice: once, us-

ing the above techniques for summative purposes and a second time without using the techniques for formative purposes.

Occasionally, an institution will establish a minimal set of standards for implementing evaluations of teachers for summative purposes. If, at your institution, teaching evaluations for summative purposes (1) are administered and implemented uniformly and universally and not by the presiding faculty member; (2) are not based solely on students' subjective evaluations; (3) include objective (not self-reported) data about students' GPA, major, and class level; (4) take into account whether the course is required or an elective, a seminar or a lecture course, a small or large enrollment, and an upper or lower level course; (5) take into account the grade distribution relative to other comparable classes; (6) take into account the extent to which gimmicks and Dr. Fox-like factors may be present in the classroom; (7) take into account the gender of the instructor and of the students; (8) ensure interpretation only by personnel trained in advanced statistical measurement, experimental design, and education; (9) are based on responses from more than 25 students; and (10) are thoroughly and routinely evaluated for both reliability and validity under the conditions in which they are normally administered at each particular location, then you are out of luck. You might have to become a better teacher to improve your ratings.

REFERENCES

- ALEAMONI, L. M., & HEXNER, P. Z. (1980) A review of the research on student evaluation and a report on the effect of different sets of instructions on student course and instructor evaluation. *Instructional Science*, 9, 67-84.
- AMES, R., & LAU, S. (1979) An attributional approach to the validity of student ratings of instruction. *Contemporary Educational Psychology*, 4, 26-39.
- ARKIN, R. M., & MARUYAMA, G. M. (1979) Attribution, affect, and college exam performance. *Journal of Educational Psychology* 71, 85-93.
- BARKE, C. R., TOLLEFSON, N., & TRACY, D. B. (1983) Relationship between course entry attitudes and end-of-course ratings. *Journal of Educational Psychology*, 75, 75-85.
- BECK, S., & DREW, T. (1989) The impact of physical attractiveness, gender, and teaching philosophy on teacher evaluations. *Journal of Educational Research*, 82, 172-177.
- BENNETT, S. K. (1982) Student perceptions of and expectations for male and female instructors: evidence relating to the question of gender bias in teaching evaluation. *Journal of Educational Psychology*, 72, 170-179.
- BOCK, F. A. (1979) Student expectations of course content affect: faculty evaluations in an abnormal psychology course. *Teaching of Psychology*, 6, 167-169.
- BRANDENBURG, D. C., SLINDE, J. A., & BATISTA, E. E. (1977) Student ratings of instruction: validity and normative interpretations. *Research in Higher Education*, 7, 67-78.
- BRASKAMP, L. A., ORY, J. C., & PIEPER, D. M. (1981) Student written comments: dimensions of instructional quality. *Journal of Educational Psychology*, 73, 65-70.
- BRAY, J. H., & HOWARD, G. S. (1980) Interaction of teacher and student sex and sex-role orientations and student evaluations of college instruction. *Contemporary Educational Psychology*, 5, 241-248.
- CANELOS, J. (1985) Teaching and course evaluation procedures: a literature review of current research. *Journal of Instructional Psychology* 12, 187-195.
- CASHIN, W. E., & DOWNEY, R. G. (1992) Using global student rating items for summative evaluations. *Journal of Educational Psychology*, 84, 563-572.

- CATES, W. M. (1993) A small-scale comparison of the equivalence of paper-and-pencil and computerized versions of student end-of-course evaluations. *Computers in Human Behavior*, 9, 401-409.
- CHACKO, T. I. (1983) Student ratings of instruction: a function of grading standards. *Educational Research Quarterly*, 8, 19-25.
- COOPER, S. E., & MILLER, J. A. (1991) MBTI learning style-teaching style incongruencies. *Educational and Psychological Measurement*, 51, 699-706.
- DIZNEY, H. F., & BRICKELL, J. L. (1984) Effects of administrative scheduling and directions upon student ratings of instruction. *Contemporary Educational Psychology*, 9, 1-7.
- DRUMMOND, R. J., & MCINTIRE, W. G. (1977) The role of cognitive style in student evaluation of instruction. *College Student Journal*, 11, 220-223.
- DUCETTE, J., & KENNEY, J. (1982) Do grading standards affect student evaluations of teaching? Some new evidence on an old question. *Journal of Educational Psychology*, 74, 308-314.
- DWINELL, P. L., & HIGBEE, J. L. (1993) Students' perceptions of the value of teaching evaluations. *Perceptual and Motor Skills*, 76, 995-1000.
- EDWARDS, J. E., & WATERS, L. K. (1984) Halo and leniency control in ratings as influenced by format, training, and rater characteristic differences. *Managerial Psychology*, 5, 1-16.
- FELDMAN, K. A. (1977) Consistency and variability among college students in rating their teachers and courses: a review and analysis. *Research in Higher Education*, 6, 223-274.
- FELDMAN, K. A. (1984) Class size and college students' evaluations of teachers and courses: a closer look. *Research in Higher Education*, 21, 45-116.
- FELDMAN, K. A. (1993) College students' view of male and female college teachers: Part II. Evidence from students' evaluations of their classroom teachers. *Research in Higher Education*, 34, 151-211.
- FOLLMAN, J. (1975) Student ratings of faculty teaching effectiveness: rater or ratee characteristics? *Research in Higher Education*, 3, 155-167.
- GILMORE, D. C., SWERDLIK, M. E., & BEEHR, T. A. (1980) Effects of class size and college major on student ratings of psychology courses. *Teaching of Psychology*, 7, 210-214.
- HEARN, J. C. (1985) Determinants of college students' overall evaluations of their academic programs. *Research in Higher Education*, 23, 413-437.
- HOFFMAN, R. A. (1983) Grade inflation and student evaluations of college courses. *Educational and Psychological Research*, 3, 151-160.
- HOFMAN, J. E., & KREMER, L. (1980) Attitudes toward higher education and course evaluation. *Journal of Educational Psychology*, 72, 610-617.
- KOVACS, R., & KAPEL, D. E. (1976) Personality correlates of faculty and course evaluations. *Research in Higher Education*, 5, 335-344.
- LUECK, T. L. (1993) The interaction effects of gender on teaching evaluations. *Journalism Educator*, 48, 46-54.
- MARSH, H. W., & ROCHE, L. A. (1993) The use of students' evaluations and an individually structured intervention to enhance university teaching effectiveness. *American Educational Research Journal*, 30, 217-251.
- MEREDITH, G. M. (1982) Grade-related attitude correlates of instructor/course satisfaction among college students. *Psychological Reports*, 50, 1142.
- MOORE, M. (1978) Course evaluation by students and self-evaluation by instructors. *Journal of Educational Research*, 72, 22-23.
- MORSTAIN, B. R. (1977) Relationship of student and instructor educational orientations with course ratings. *Journal of Educational Psychology*, 69, 388-398.
- NAFTULIN, D. H., WARE, J. E., & DONNELLY, F. A. (1973) The Dr. Fox lecture: a paradigm of educational seduction. *Journal of Medical Education*, 48, 630-635.
- NEUMANN, L., & NEUMANN, Y. (1985) Determinants of students' instructional evaluation: a comparison of four levels of academic areas. *Journal of Educational Research*, 78, 152-158.
- PARISH, T. S., & CAMPBELL, N. J. (1977) Consistency of students' evaluations of instructors. *Journal of Instructional Psychology*, 4, 30-33.
- REYNOLDS, D. V. (1977) Students who haven't seen a film on sexuality and communication prefer it to a lecture on the history of psychology they haven't heard: some implications for the university. *Teaching of Psychology*, 4, 82-83.

- ROHDE, R. I. (1993) Effect of word processing on students' grades and evaluation of instruction in freshman composition. *Psychological Reports*, 72, 1259-1264.
- ROMEO, F. F., & WEBER, W. A. (1985) An examination of variables which influence student ratings of university faculty. *College Student Journal*, 19, 133-140.
- SILVESTRO, J. R. (1979) Use of video-cassette summaries of childhood in teaching developmental psychology. *Teaching of Psychology*, 6, 171-172.
- TAYLOR, J. S., & RICKETTS, M. S. (1981) Student attitudes about participation in instructor and course evaluations. *Journal of Instructional Psychology*, 9, 18-24.
- TERRY, R. L., & MCINTOSH, D. E. (1988) Do students' expectancies affect their course evaluations? *Educational and Psychological Measurement*, 48, 787-798.
- TETENBAUM, T. (1977) The factor invariance of student ratings of instruction under three sets of directions. *Research in Higher Education*, 6, 11-23.
- THOMAS, D., RIBICH, F., & FREIE, J. (1982) The relationship between psychological identification with instructors and student ratings of college courses. *Instructional Science*, 11, 139-154.
- UNGER, R. K. (1979) Sexism in teacher evaluation: the comparability of real life to laboratory analogs. *Academic Psychology Bulletin*, 1, 163-170.
- VAN GIFFEN, K. (1990) Influence of professor gender and perceived use of humor on course evaluations. *Humor International—Journal of Humor Research*, 3, 65-73.
- VASTA, R., & SARMIENTO, R. F. (1979) Liberal grading improves evaluations but not performance. *Journal of Educational Psychology*, 71, 207-211.

Accepted April 26, 1996.