



Student Rating Forms

[From the hard copy book *Tools for Teaching* by Barbara Gross Davis; [Jossey-Bass](#) Publishers: San Francisco, 1993. Linking to this book chapter from other websites is permissible. However, the contents of this chapter may not be copied, printed, or distributed in hard copy form without permission.]

Student rating forms, also called student end-of-course questionnaires or student evaluation forms, are traditionally administered at the end of the term to solicit student evaluations of a course. (For suggestions on using student rating forms midsemester, see "[Fast Feedback](#).") Typically such end-of-course information is used by faculty committees and administrators to make personnel decisions regarding an instructor: merit increases, promotion, tenure.

At one time, it was considered controversial to administer student rating forms. Now such forms have become commonplace because it makes sense to survey students to find out what they think about their experiences in the class over the term and also because a substantial body of research has concluded that administering questionnaires to students is both valid and reliable. Here is some of what we know about student rating forms from the research (adapted from Davis, 1988):

- Ratings of overall teaching effectiveness are moderately correlated with independent measures of student learning and achievement. Students of highly rated teachers achieve higher final exam scores, can better apply course material, and are more inclined to pursue the subject subsequently. (Sources: Abrami, Apollonia, and Cohen, 1990; Braskamp, Brandenburg, and Ory, 1984; Cohen, 1981; Kulik and McKeachie, 1975; McMillan, Wergin, Forsyth, and Brown, 1986; Marsh and Dunkin, 1992)
- An instructor's ratings for a given course tend to be relatively consistent over successive years; there is not much variation in student ratings for an individual instructor regardless of whether the form is administered to current students or to alumni. (Sources: Braskamp, Brandenburg, and Ory, 1984; Centra, 1979; McMillan, Wergin, Forsyth, and Brown, 1986; Marsh and Dunkin, 1992)
- There is little or no relationship between the following characteristics of students and their ratings of instruction: age, grade point average, year in college, and academic ability. No consistent relationships have been found between student ratings and such variables as the amount of homework assigned or grading standards. (Sources: Braskamp, Brandenburg, and Ory, 1984; Centra, 1979; Kulik and McKeachie, 1975; McKeachie, 1979; McMillan, Wergin, Forsyth, and Brown, 1986; Marsh and Dunkin, 1992)
- Researchers do report the following relationships:
 - Students tend to rate courses in their major fields and elective courses higher than required courses outside their majors. (Sources: Kulik and McKeachie, 1975; McKeachie, 1979; Marsh and Dunkin, 1992)
 - Faculty tend to receive more positive ratings than graduate student instructors. (Source: Marsh and Dunkin, 1992)
 - The gender of a student has little effect on ratings. The gender of an instructor, however, may have an impact. Though some studies report no relationship between a professor's gender and student ratings, others show that adhering to a gender-appropriate teaching style may be rewarded by higher evaluations. (Sources: Basow and Silberg, 1987; Bennett, 1982; Kierstead, D'Agostin, and Dill, 1988; Marsh and Dunkin, 1992; Statham, Richardson, and Cook, 1991)
 - Ratings can be influenced by class size (very small classes tend to receive higher ratings), by discipline (humanities instructors tend to receive higher ratings than instructors in the physical sciences), and by type (discussion courses tend to receive higher ratings than lecture courses). (Sources: Cashin, 1992; Feldman, 1984; Marsh and Dunkin, 1992)
 - Students' expectations affect their ratings: students who expect a course or teacher to be good generally find their expectations confirmed. (Sources: Marsh and Dunkin, 1992; Perry, Abrami, Leventhal, and Check, 1979)

Many campuses have standard rating forms that all faculty administer and centralized procedures for analyzing the data. Check with your department to find out what policies exist on your campus. If you are free to design and administer your own questionnaire, the suggestions below (adapted from Davis, 1988) will help you make the most of student rating forms. If you use the standard campus questionnaire, these suggestions can improve the way you administer the forms and interpret the results.



Selecting or Designing the Questionnaire

Use forms that give students the opportunity to provide quantitative ratings and to comment narratively on an instructor's performance. Forms that include both quantitative and narrative data are the most useful for getting the broadest picture of students' reactions. Davis (1988) presents sample student rating forms and items that can be used in end-of-course questionnaires, grouped by such categories as accessibility, organization and preparation, and interaction. Kulik (1991) describes the catalogue of items for instructor-designed questionnaires (IDQs), grouped by categories such as student development, course elements, and student responsibility. Theall and Franklin (1990) offer cautions in developing your own questionnaire: it is a time-consuming process to produce a valid and reliable form. They recommend using specific, unambiguous items. For example: "instructor defines new or unfamiliar terms," "repeats difficult concepts," "provides frequent examples." Especially if you are using these forms for your own improvement, the items you select or develop should capture specific behaviors that are amenable to change.

Select items that reflect your department's and institution's criteria of effective teaching and that are within students' range of judgment. For example, current students can judge how well prepared instructors are, how effectively they make use of class time, how well they explain things and with what level of enthusiasm, and how responsive they are to difficulties the students may be having in the course. Students can also comment on whether the instructor promotes original thinking and critical evaluation of ideas. In contrast, current students are not qualified to judge whether instructors are up-to-date in their field or to rate how adequately a course prepares them for advanced course work in the field.

State each item clearly. For example, "The instructor routinely summarizes major points" is unambiguous, while "The instructor is well prepared and gives fair exams" confounds two different issues.

For at least some of the key items, provide a numerical rating scale. The use of quantifiable items enables you to calculate a class's average response and to note the distribution of responses, both of which are useful in interpreting the results of the evaluation. Use either a 5-point or 7-point scale, with 1 representing "not at all descriptive" and 5 (or 7) "very descriptive." Also provide a "Don't know or doesn't apply" option that students can check.

Include at least one item that asks students about the effects of the course. For example, ask students to describe or rate the knowledge, appreciation, or skills they acquired in the course or their intellectual, personal, or professional growth as a result of the instructor's teaching.

Include at least one quantitative measure on the overall effectiveness of the instructor. For example:

Considering both the limitations and possibilities of the subject matter and course, how would you rate the overall effectiveness of this instructor?

<i>Not at all effective</i>			<i>Moderately effective</i>			<i>Extremely effective</i>
1	2	3	4	5	6	7

Include at least one open-ended item that asks about the overall effectiveness of the instructor. For example, "Please identify what you perceive to be the greatest strengths and weaknesses of this instructor's teaching."

Limit the number of questions about student characteristics. Student characteristics have relatively little influence on ratings of overall effectiveness (Cashin, 1990a). You might want to know, however, whether students are taking the course as an elective or to fulfill a requirement.

Keep the form short. Since students may fill out evaluation forms for all their instructors, questionnaires should be brief.



Administering the Questionnaire

Announce in advance the date on which rating forms will be handed out. Schedule the time sometime during the last two weeks of the term; allow ten to fifteen minutes for this activity. Encourage students to attend that day and complete the form. Do not distribute forms at the final exam, when students are preoccupied with other matters.

Inform the students about the purpose of the questionnaire. It is helpful to stress to students that their ratings and comments are important and will be used by both you and your department. Students may want to know how the completed forms will be handled. Here are some sample 'instructions that can be placed on the rating forms and read aloud:

We hope you will take the time to answer each question carefully. The information you provide will be part of our ongoing efforts to improve the curriculum and the teaching in this department. In addition, your comments will be summarized and used in faculty promotions and reviews [if this is true on your campus]. To maintain confidentiality, these forms will be collected by someone other than the instructor and will not be available to the instructor until after the course grades have been submitted.

Ask students to complete the forms anonymously. Research shows that requiring students to sign the forms inflates the ratings (Cashin, 1990b). In addition, anonymity can eliminate students' concerns about possible retribution (Ory, 1990).

Designate a student from the class (or a department staff member) to supervise questionnaire administration. You may hand out the forms (always bring several more than the official number of students enrolled), but you should not be present while students complete the questionnaires, and you should not collect the forms. Ask the designated collector to place the forms in a large manilla envelope, noting on the outside your name, the course number, the total number of students present, the total number of forms collected, and the date. The sealed envelope should be delivered to the department office.

Do not look at the forms until after you have submitted final grades for the course. Some campuses provide the faculty with summaries of their rating forms, including computer printouts that show trends and comparative data. If your campus does not provide such a service, you will want to analyze and summarize the data yourself.



Summarizing Responses

Look at the number of students who completed forms and the total class enrollment.

Ideally, you would like a response rate (number of completed forms divided by number of enrolled students) of 80 percent or higher. When less than two-thirds of the enrolled students in classes of one hundred or fewer students or less than one-half of the enrolled students in classes of more than a hundred students have submitted forms, the data should be interpreted cautiously if the questionnaires are being used for personnel decisions, such as merit, tenure, or promotion. (Sources: Cashin, 1990b; Theall and Franklin, 1990).

Keep the data separate for each course offering. Aggregating data for several different courses may obscure differences in teaching effectiveness for various kinds of instruction and may raise questions of proper weighting of the responses in each course. Aggregating data for several offerings of the same course may obscure long-term trends toward increased or decreased student satisfaction.

Do not summarize data if there are fewer than ten questionnaires. Student questionnaires from independent reading courses or seminars with very small enrollments may be accumulated over several terms and summarized when their numbers are sufficiently large. (Source: Cashin, 1990b)

If your department does not already do so, prepare summary statistics for the quantifiable questions. The summary should include the following:

- The frequency distribution of student ratings for each item (the number and percentage of students selecting each response)
- The average response, either the mean (calculated to one decimal point), mode, or median
- The standard deviation (an index of agreement or disagreement among respondents)
- Departmental norms (averages) or comparison norms, if available, on key items for courses of a similar size, level, and type of instruction (for example, laboratory, seminar, studio, lecture)

Summarize the narrative comments. The summary should reflect the entire range of comments as well as their preponderance. To prepare such a summary, read all the students' comments about a single question, develop categories or headings that meaningfully group most of the comments, and record the number of comments made in each category. In deciding what to ignore and what to consider, take into account your goals for the course, your values and emphases, and your teaching style (Lunde, 1988). Also remember that it is human nature to focus on that piercing negative comment to the exclusion of the positive remarks from most of the class. An outside teaching consultant or supportive colleague can help you put students' comments in perspective.



Interpreting Responses

For quantifiable questions, determine the percentage of omitted responses. Some items may be left blank because they do not apply. Items with low response rates should be interpreted cautiously. (Source: Theall and Franklin, 1990)

Look at the average ratings for the quantifiable questions. Average ratings can be interpreted on an absolute scale and in relation to the ratings of other similar courses and instructors. For example, a mean rating of 4 on a 7-point scale for overall course evaluation may be labeled "moderately effective." However, if half of all similar courses receive mean ratings above 4, then this 4 rating falls in the lower half. There is some debate within the field on whether such comparisons are meaningful (Theall and Franklin, 1990), even among courses similar in level (lower division, upper division, graduate), size, format (lecture, laboratory, and so on), and student demographics. Cashin (1992) argues for using comparative data, pointing out that because students tend to rate most items high, a score of 3.5 on a 7-point scale is not really an indicator of "average" effectiveness. In using student ratings to improve your teaching, try to incorporate some comparative information to better understand your own strengths, weaknesses, and accomplishments (Kulik, 1991). Perhaps the best comparative information, as Kulik suggests, comes from noticing changes in the ratings of a course you have taught several times.

Look at the range of student responses for the quantifiable questions. The range provides important information. For example, the average of all ratings for your course may be 5 on a 7-point scale. But notice whether all students rated the course as 4, 5, or 6 or whether some 2 and 3 ratings were balanced out by some 7s. If ratings cluster at the two ends of the scale, then some aspects of your teaching work for one group of students but not for another--an area to explore. The standard deviation also provides useful information. A standard deviation of less than 1.0 (on a 5-point scale) indicates relatively good agreement among the respondents. Deviations above 1.2 indicate a divided class on that item. (Sources: Cashin, 1990b; Theall and Franklin, 1990).

Note your highest and lowest rated items. One way to analyze student ratings is to calculate the averages for individual questions and note your highest and lowest rated items. See whether your strengths and weaknesses cluster in patterns on any of the following topics: organization and clarity, enthusiasm and stimulation of student interest, teacher-student rapport, teaching and communication skills, course work load and difficulty, fairness of exams and grading, classroom climate. As a rule of thumb, it is usually cause for concern when a third of the students give low ratings to some aspect of a course (Kulik, 1991). In looking at your highest and lowest rated items, try to identify specific behaviors of yours that might have caused students to give you those ratings. If you do this exercise with a colleague who has administered the same form to his or her students, you can exchange examples of behaviors that lead to high ratings.

From the open-ended comments, identify specific problems. Lunde (1988) recommends reading the comments to pinpoint specific complaints--for example, student anxiety about the degree of structure in the course or your expectations--for students. Then determine whether the complaint is justified (in this case by looking at the syllabus and handouts). If the worry is legitimate, identify specific steps you can take to address the weakness. Keep in mind that students give few detailed suggestions on how to improve a course; they are better at spotting problems (Braskamp, Brandenburg, and Ory, 1984). If you have the time or can prevail upon an experienced

teaching consultant, you could analyze open-ended comments using a grid technique, a method for determining whether students who rate a course highly are saying the same things as those who rate the course lower (Lewis, 1991).

Consider how characteristics of courses can influence ratings. Small classes, electives, and courses in the humanities tend to receive more favorable ratings. For each characteristic the differences are minimal, but together they may be meaningful (Sorcinelli, 1986). It may also be helpful, in interpreting the results, to take into account whether this course is your favorite, one you frequently teach, or a course that you are teaching at the request of the department chair.

Ask a knowledgeable consultant or a colleague for assistance. Some teachers can read their students' ratings and map out strategies to improve teaching. Others find it helpful to review the ratings with a knowledgeable colleague or teaching consultant, available on some campuses. Consultants can help you interpret the results and explore strategies for improvement.

Consider making your ratings available to students. Several departments on campuses around the country--and some entire institutions-- have a tradition of making their course ratings public. Many faculty members in these departments or on these campuses believe that because faculty members work harder at their teaching when they know the results are on view to their peers and to students, teaching is of higher quality when the rating forms are publicly available.



References

- Abrami, R C., Apollonia, S., and Cohen, P. A. "Validity of Student Ratings of Instruction: What We Know and What We Do Not." *Journal of Educational Psychology*, 1990, 82(2), 219-231.
- Basow, S. A., and Silberg, N. I "Student Evaluations of College Professors: Are Male and Female Professors Rated Differently?" *Journal of Educational Psychology*, 1987, 79(3), 308-314.
- Bennett, S. K. "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*, 1982, 74(2), 170-179.
- Braskamp, L. A., Brandenburg, D. C., and Ory, J. C. *Evaluating Teaching Effectiveness: A Practical Guide*. Newbury Park, CA: Sage, 1984.
- Cashin, W, E. "Students Do Rate Different Academic Fields Differently" In M. Theall and J. Franklin (eds.), *Student Ratings of Instruction: Issues for Improving Practice*. New Directions for Teaching and Learning, no. 43. San Francisco: Jossey-Bass, 1990a.
- Cashin, W. E. "Student Ratings of Teaching: Recommendations for Use." *Idea Paper*, no. 22. Manhattan: Center for Faculty Evaluation and Development in Higher Education, Kansas State University, 1990.
- Cashin, W E. "Student Ratings: The Need for Comparative Data." *Instructional Evaluation and Faculty Development*, 1992, 12(2), 1-6. (Available from Office of Instructional Research and Evaluation, Northeastern University)
- Centra, J. A. *Determining Faculty Effectiveness*. San Francisco: Jossey-Bass, 1979.
- Cohen, R A. "Student Ratings of Instruction and Student Achievement." *Review of Educational Research*, 1981, S1(3), 281-301.
- Davis, B. G. *Sourcebook for Evaluating Teaching*. Berkeley: Office of Educational Development, University of California, 1988.
- Feldman, K. A. "Class Size and College Students' Evaluations of Teachers and Courses: A Closer Look." *Research in Higher Education*, 1984,21(1), 45-116.
- Kierstead, D., D'Agostin, P., and Dill, W. "Sex Role Stereotyping of College Professors: Bias in Students' Ratings of Instructors." *Journal of Educational Psychology*, 1988, 80(3), 342-344.
- Kulik, J. A. "Student Ratings of Instruction." *CRLT Occasional Paper*, no. 4. Ann Arbor: Center for Research on Learning and Teaching, University of Michigan, 1991.

Kulik, J. A., and McKeachie, W. J. "The Evaluation of Teachers in Higher Education." In E. N.

Kerlinger (ed.), *Review of Research in Education*. Itasca, Ill.: Peacock, 1975.

Lewis, K. G. "Gathering Data for the Improvement of Teaching: What Do I Need and How Do I Get It?" In M. D. Sorcinelli and A. E. Austin (eds.), *Developing New and Junior Faculty*. New Directions for Teaching and Learning, no. 48. San Francisco: Jossey-Bass, 1991.

Lunde, J. P. "Listening to Students Learn: What Are Their Comments Saying?" *Teaching at the University of Nebraska, Lincoln*, 1988, 10(1), 1-4. (Newsletter available from the Teaching and Learning Center, University of Nebraska, Lincoln)

McKeachie, W. J. "Student Ratings of Faculty: A Reprise." *Academe*, 1979, 6S(6), 384-397.

McMillan, J. H., Wergin, J. E., Forsyth, D. R., and Brown, J. C. "Student Ratings of Instruction: A Summary of Literature." *Instructional Evaluation*, 1986, 9(1), 2-9.

Marsh, H. W., and Dunkin, M. J. "Students' Evaluations of University Teaching: A Multidimensional Perspective." In J. C. Smart (ed.), *Higher Education: A Handbook of Theory and Research*. Vol. 8. New York: Agathon Press, 1992.

Ory, J. C. "Student Ratings of Instruction: Ethics and Practice." In M. Threall and J. Franklin (eds.), *Student Ratings of Instruction: Issues for Improving Practice*. New Directions for Teaching and Learning, no. 43. San Francisco: Jossey-Bass, 1990.

Perry, R. P., Abrarm, P. C., Leventhal, L., and Check, J. "Instructor Reputation: An Expectancy Relationship Involving Student Ratings and Achievement." *Journal of Educational Psychology*, 1979, 71(6), 776-787.

Sorcinelli, M. D. *Evaluation of Teaching Handbook*. Bloomington: Dean of the Faculties Office, Indiana University, 1986.

Statham, A., Richardson, L., and Cook, J. A. *Gender and University Teaching: A Negotiated Difference*. Albany: State University of New York Press, 1991.

Theall, M., and Franklin, J. (eds.). *Student Ratings of Instruction: Issues for Improving Practice*. New Directions for Teaching and Learning, no. 43. San Francisco: Jossey-Bass, 1990.

From the hard copy book *Tools for Teaching* by Barbara Gross Davis; [Jossey-Bass](#) Publishers: San Francisco, 1993. Linking to this book chapter from other websites is permissible. However, the contents of this chapter may not be copied, printed, or distributed in hard copy form without permission.

Available at the UCB campus library (call # LB2331.D37). The entire book is also available [online](#) as part of netLibrary (accessible only through computers connected to the UC Berkeley campus network). It is available for purchase at the Cal Student Store textbook department, the [publisher](#), and [Amazon](#). Note: Barbara Gross Davis is working on the second edition of *Tools for Teaching*.

[Publications and Teaching Tips](#) | [Office of Educational Development](#) | [UC Berkeley](#)

Last Updated 4/11/02

Questions or Comments? [Contact us](#).