
**THE EVALUATION OF TEACHING:
A POSITION STATEMENT FROM THE FACULTY WELFARE
AND PROFESSIONAL DEVELOPMENT COMMITTEE**

(Approved by the Faculty Senate, December 1, 1999)

Teaching is fundamental to this University. This position is clearly stated in the “University-wide Evaluation Guidelines for Promotions and Tenure”:

The primary function of the University is teaching. Research and creative activity, service, and directed professional activity, while important to the life of the University, do not have the central importance of teaching. Therefore, it is essential that excellence in teaching be encouraged and rewarded. (1/21/98, p. 3)

Teaching must also be evaluated, at least according to the “Guidelines for the Evaluation of Teaching” (as approved by the Faculty Council, April 15, 1986). The introduction to that document stated:

The basic purpose of evaluating teaching is to encourage individual faculty development. Systems of evaluation should reflect this aim. It is also acknowledged that summary information of teaching effectiveness is used in the promotion and tenure process, for annual review of all faculty, and for other institutional administrative purposes. . . . Evaluation systems must therefore be designed to provide sufficient information for individual use as well as to assure that results used for administrative and other purposes are derived from a balanced and fair approach that *employs a variety of techniques* (1986, p. 1. Emphasis added).

Teaching is thus to be evaluated. But what does teaching consist of? The “University-wide Evaluation Guidelines for Promotions and Tenure” defines teaching as “activities related to instruction and learning that occur both inside and outside of the classroom” (1/21/98, p. 3). The document lists various aspects of teaching under the following headings:

- Instructing;
- Advising, supervising, guiding, and mentoring students;
- Developing learning; and
- Developing teachers and teaching.

Under the heading “Instructing” appear three components:

- Instructing students in courses, laboratories, clinics, studio classes, libraries, study-abroad programs, and distance education;
- Instructing participants in workshops, retreats, and seminars;
- Facilitating faculty, student, and/or staff learning.

Without going into the same degree of detail, a further twenty aspects of teaching are included under the remaining three headings. They involve aspects of teaching such as advising theses and dissertations, directing collaborative research with students, supervising internships, developing courses, developing computer-assisted instruction and distance learning, evaluating the teaching of colleagues, etc.

Given the myriad of activities that encompass teaching, the means of evaluating teaching are necessarily complex. A single method of evaluation cannot deal with such complexity. Traditionally, this University has relied on one method, however. Student evaluations (often known as “student ratings”) of classroom-based teaching is not only by far the most common means of evaluation, but in most cases it is also the only means of evaluation. A total of 36 Units (Departments within a School or the College or an entire School) responded to the University Teaching and Learning Center’s request for information about teaching evaluation methods, and a great deal of uniformity was found. Students who have been taught

in classroom settings are asked to respond to a series of questions on a four- or five-point scale, using a computer-ready scantron sheet. The majority of Units require that students respond to questions about:

- Accessibility of instructor;
- Course goals and objectives;
- The fairness of examinations and the quality of feedback from the instructor;
- Questions about the instructor (preparation and organization, encouragement given, quality of teaching);
- Questions about the course (difficulty level, challenging, satisfaction gained);
- Questions about instructor/student interaction;
- The instructor's presentation style and knowledge of the subject; and
- Student demographics.

These are important questions to ask. However, there are two major causes for concern, at least when the questions are asked in the traditional form and require the traditional (i.e., scantron) responses. The first is that these types of questions have nothing to do with many of the aspects of teaching outlined above. This is not to argue that questions cannot be asked of teaching involving mentoring, providing internship experience, supervising theses or dissertations, or directing collaborative research with students. It is difficult to know how there can be meaningful scantron-based responses to questions about evaluating the teaching of one's colleagues, developing or redesigning courses, developing computer-aided instruction, and so on. There are, however, much more appropriate ways in which to answer questions such as these, methods that require a more qualitative rather than quantitative judgment.

The second cause for concern is that the use of these formal, traditional methods of evaluating teaching encourages easily quantifiable comparisons to be made across Units and across faculty within Units. Their use in comparisons is so widespread that we suspect that it reveals a not-so-hidden agenda. Methods of evaluation that are simple to use and to quantify may be the best for purposes of comparison, but may be ineffective means of helping faculty improve their teaching, an alternative agenda.

We refer again to the "University-wide Evaluation Guidelines for Promotions and Tenure," which detailed various means of documenting effective teaching. Under the heading "Documented Outcomes" appear fifteen ways to document effective teaching, including student test scores, student publications, placement of graduates, supervision of honors or master's theses or doctoral dissertations, and instructional innovations. A further five ways of making judgments about teaching are included, one of which involves "Statements from students such as information from exit interviews, written comments on examinations, teacher evaluations, and letters from students" (1/21/98, p. 4). The Guidelines clearly give a very limited role to traditional student ratings.

This Committee therefore discourages sole reliance on student ratings, despite the fact that they are easily quantifiable, and wishes to encourage methods of evaluating teaching that give learners a voice while being of assistance to both teachers and administrators. We do not believe that ratings have no role to play. However, we strongly believe that if student ratings are to be used they should be used only in combination with other sources of data. This is not simply the belief of the Committee members; we share this position with those who have extensively reviewed the use of student ratings, both on this campus ("Teaching Effectiveness Subcommittee Report," UNCG School of Education, February 1997) and off campus (IDEA papers #22, 32, 33, published by the Center for Faculty Evaluation and Development, Kansas State University, 1990, 1995, 1996). For example, "student ratings are only one source of data about teaching and must be used in combination with multiple sources of data if one wishes to make a judgment about all of the components of college teaching. Further, student ratings are data that must be interpreted" (Cashin, 1995, IDEA Paper #32). Similarly, after an extensive survey of research on teaching evaluation methods, the first recommendation of the "Teaching Effectiveness Subcommittee Report" was that "a combination of data collection methods be used to create a portfolio of materials to be used in the identification of successful teaching" (1997, p. 9). The second recommendation was that "various

methods of evaluation be assessed and the results of the assessments be reported to the Faculty prior to adoption [of] any combination of methods; that no ONE METHOD take primacy when evaluating teaching or course effectiveness and improvement” (1997, p. 10).

What methods of evaluation should be used? It is our belief that there is no single or simple answer. However, in light of the position laid out in the “University-wide Evaluation Guidelines for Promotions and Tenure,” and followed in the Post-Tenure Review Guidelines, it seems clear that no single rating scale should be used by all faculty across campus. Instead, both guidelines emphasize that evaluations are most appropriately assessed at the Unit level (departments in schools or the College, non-departmentalized schools, or the Library). As stated in the “University Promotions and Tenure Guidelines”: “It is within the Department [Unit as we have defined it] that the criteria for assessing faculty performance are best understood” (Preamble, p. 1).

Our recommendation is thus that the manner of evaluating, the types of questions asked on any rating forms that are used, and the degree of emphasis placed on teaching portfolios should be decided at the Unit level. We further recommend that Units take seriously the position that classroom teaching is but one aspect of teaching. It is important, but so are mentoring, advising, producing teacher-student creative works, designing and re-designing courses and syllabi, and integrating technology into the curriculum. Finally, we recommend that Units consider seriously the impact of using traditional student rating forms that are easily quantifiable. Simply because of their ease of use and analysis they all too easily become the sole evaluative criterion even when combined with other methods.