Guide to

Evaluation
of Instruction

Guide to

Evaluation
of Instruction

Evaluation of Instruction

Introduction

Student evaluation of instruction was a part of academic life at UCLA even before University of California President Charles Hitch first implemented the requirement in a 1969 memo addressed to University of California Chancellors and Members of the Academic Council. President Hitch felt strongly that the considered opinion of the students themselves must always be taken into account, while recognizing that student evaluations provide only one of many useful tools for assessing teaching effectiveness. As a response to President Hitch, UCLA formed the "Task Force on the Evaluation of Teaching" in 1970, whose mission was to assess departmental and student evaluation practices at the University, as well as throughout the United States. Upon review, the Task Force recommended that "an ongoing process of regular evaluations be conducted emphasizing the facilitation of faculty growth [and] the need for feedback for the development of UCLA's instructional programs." Particular significance was placed on the need for student surveys of teaching effectiveness for every course for every instructor. Following the Task Force guidelines, UCLA's Office of Instructional Development has provided this service to the campus community through its Evaluation of Instruction Program (EIP). Presently, EIP distributes more than 300,000 forms annually to approximately 100 departments and programs, and provides forms, analyses, and reports for evaluation of courses, instructors and teaching assistants.

As part of this service, the Evaluation Guide offers insight into the University's policies and procedures regarding the evaluation process, provides interpretive techniques for making sense of student survey data, presents informal evaluation methods and feedback tools, and suggests resources for consultation and assistance. A brief literature review and bibliography of seminal work in the field of instructional evaluation is offered as well. The format of the Evaluation Guide should facilitate quick and efficient browsing by experienced UCLA faculty who wish to find an answer to an immediate question, as well as assist new faculty who may have more extensive questions about the process of evaluation of instruction. The Evaluation Guide is designed to complement other evaluation resources available at the Office of Instructional Development for use by all UCLA faculty members.

Defining and Measuring Effective Teaching

Effective teaching can be defined, very simply, as activities that promote student learning. It encompasses all of those instructor behaviors that foster student learning of the instructor's and/or of the institution's educational objectives. Ideally, the students will also share at least some of these objectives. This definition of effective teaching includes curriculum and course development, advising, and supervision of student research as well as classroom performance. Given this broad definition, no single approach is sufficient for evaluating effective teaching. Rather, student ratings, self-reviews, peer evaluations, and objective criteria such as student performances and improvements are all useful tools for evaluating different aspects of teaching.

Table 1.1 (see next page) indicates some important sources of data that can be used to measure effective teaching. The sources fall into three main types: students, peers, and the instructor him/herself (through self-reflection). Since measuring teaching is clearly not an exact science, the more varied the data sources, the more useful the measurement is likely to be.

Since there is a great deal of focus on student evaluations used to assess teaching at UCLA, much of the remainder of this guide pertains to the implementation of this particular measure. However, the other two sources of data can provide valuable additional insight, and should be considered as part of any comprehensive approach.

Table 1.1	Sources of Information about Teaching

The following is a list of some important sources of information about teaching and their main advantages and disadvantages for evaluation purposes.

Students

Systematic Student Evaluations

These are very important for a global picture of the course. The students are the ones who are doing the learning, so their perception is important. Their response often highlights strengths and weaknesses. However, students are not subject matter experts. Also, students' ratings are sometimes influenced by their own motivations, attitudes and needs.

Interviews with Students

This is a very useful evaluation procedure which can yield much information in a short time. A group of students from a course are interviewed by other faculty about their experience in a course. A structured format is followed and typically, a consensus view of the nature of the course -- its strengths, weaknesses, and problems -- emerges in 15 to 20 minutes. The difficulties with this technique are associated with the training needed to do the interviews and report the results, and the selection and recruitment of the sample of students interviewed.

Long-term Follow-up of Students

Surveys or interviews with seniors and alumni can yield information based on a wider context of university and life experience than given by the usual end-of-course student survey. However, reaching alumni can be difficult so response rates are often low.

Peer Review

Classroom Visits

Visits by other faculty can provide information about the process of teaching. However, correct use of this procedure is time consuming. It is best done when training can be provided, and two or more visits can be arranged by at least two observers. In addition, this technique is most effective when prefaced by a discussion between the instructor and observer regarding the goals of the class.

Colleague Evaluation of Materials

Colleagues have the expertise to evaluate the quality of a course as evidenced by its content and format. Colleagues can also evaluate student achievement as indicated by performance on exams and papers.

| Table 1.1 | Sources of Information about Teaching |

Self Reflection

Teaching Activities, Reports and Self-Reviews

The instructor's statement of his/her goals for the course, teaching methods and philosophy, student outcomes, and plans for improvement are a critical source of information. Oftentimes, there may be external factors, bad classes, difficult teaching problems and experiments with innovative teaching techniques (which may lower ratings initially before ultimately raising them) on which only the instructor can reflect. A systematic self-review has the potential for contributing significantly to the instructor's teaching improvement by focusing on the strengths and weaknesses of the course in light of his/her original course objectives. If the instructor notes broad shifts from the course's original objectives it may lead to a reassessment of methodological approaches when drafting future courses.

Measures of Student Achievement

When appropriate tests are available, measures of student learning are a prime criterion of effective teaching. However, valid direct measures of student learning require considerable developmental effort. Also, interpretation of achievement tests requires some comparable measures of student motivation and interest. There are a number of informal assessment techniques which may be employed to gather this information. (See for example, Angelo and Cross, 1993.)

Suggested readings:

Arreola, Raoul. (2000). *Developing a Comprehensive Faculty Evaluation System: A Handbook for College Faculty and Administrators on Designing and Operating a Comprehensive Faculty Evaluation System*. Bolton, MA: Anker Publishing Co., Inc.

Angelo, T. A., & Cross, K. P. (1993). *Classroom Assessment Techniques: A Handbook for College Teachers, Second Edition*. San Francisco, CA: Jossey-Bass.

Davis, B. G. (1993). *Tools for Teaching*. San Francisco, CA: Jossey-Bass.

Seldin, P. et al. (1999). *Changing Practices in Evaluation Teaching: a Practical Guide to Improved Faculty Performance and Promotion/Tenure Decisions*. Bolton, MA: Anker Publishing Company, Inc.

Data from Students:
Systematic Student Evaluations

The Student Evaluation System at UCLA

The Office of Instructional Development's Evaluation of Instruction Program (EIP) helps to assess and improve teaching at UCLA by providing instructor evaluation services for instructors and TAs across campus. At the end of each quarter, instructors have the opportunity to solicit formal evaluations from students enrolled in their classes. EIP distributes, collects, and processes evaluation forms via a network of departmental coordinators. Interested instructors should first contact their department evaluation coordinator, or may reach the centralized Evaluation of Instructional Program at extension 56939 (www.oid.ucla.edu/eip).

EIP's standard evaluation forms, which cover most teaching situations for faculty and TAs, were designed in consultation with faculty committees, national experts on assessment, and recommendations from UCLA and System-wide surveys of faculty, TAs, and students. While most departments on campus use the standard campus-wide forms, a few departments and schools have devised their own standard forms that are distributed and administered by EIP. The Faculty Consultation Service can work with departments and individuals to develop questionnaires for special needs.

Policy on Data

The Evaluation of Instruction Program seeks to provide as secure an environment for data as it can. The forms are stored for processing in a physically safeguarded location. Data are compiled centrally only for the numerical scores. In addition to individual instructor scores, larger collective comparison data (e.g. for Divisions, or Schools) are also calculated. These compiled individual and comparison data are returned to Department Chairs along with the original student forms. Each Department may devise its own particular policy on how data are presented, distributed, or made openly available. Most often, the forms are returned to the instructor so that the written comments may be read firsthand. Some departments choose to transcribe the comments for small seminars to protect the identity of the students. The numerical data are provided by the Evaluation of Instruction Program in electronic format, and include both departmental scores and appropriate comparison data. The Evaluation of Instruction Program does not release individual instructor data except to the Department or to the individual instructor. External requests for such data are referred to the Departments. External requests for comparison data (without any individuals identified) are generally granted.

Retaining Data

Most departments develop a reliable system for storing numerical data from teaching evaluations. It is less common for departments to retain written comments. Because such data are often used when compiling dossiers for personnel decisions, faculty should be careful to keep copies of their own evaluations. Even normally reliable systems sometimes have unexplainable lapses, and it can be extremely difficult (if not impossible) to re-establish such data after the fact. In addition, it may be useful to annotate records with information that might provide insight into any anomalous results (e.g. "the class was scheduled into a room that was frequently impacted by construction noise," or "it was my first attempt to develop group projects and they did not work the way I had hoped.")

Some departments retain only overall ratings, and again, instructors would be better advised to keep data which encompass all the individual items on the form. Such information can often expand on understanding why certain classes may have been rated higher or lower.

Most evidence of teaching data is used within a period of six years after the time it was collected. Based on actual requests for re-establishing older records, eight years would provide a more certain time frame. Comparison data, such as to other instructors in the department or to the overall University mean, should likewise be kept in order to provide bases for

comparison, should later disputes arise. Faculty who are nominated by their departments for teaching awards also find some of the written comments useful in documenting what students find particularly compelling about their teaching.

Disputed Data

In the infrequent situation that the integrity of the data are disputed (e.g. if forms are intercepted by the instructor or the chair before processing, or considerably more forms are returned than the number of students enrolled in the course) what facts are known are forwarded to the Committee on Teaching for any further consideration of action. Other Senate Committees may become involved as appropriate.

Student Questionnaire Administration Procedures

Since student ratings are sensitive to a wide variety of situational factors, it is very important that questionnaire administration procedures are consistent. The goal is to get candid feedback focused on instructional issues from as many students in the course as possible. The following general guidelines for collecting student ratings may help instructors achieve this goal:

Protect Student Anonymity

It is important that students not be identified in any way with their evaluations, otherwise less than candid responses are likely. In particular, students fear that an adverse rating might negatively impact their course grade. It is, of course, difficult to maintain confidentiality of student raters in small classes and individual study courses. There is no simple solution to this problem. One option is to have the departmental coordinator type up the responses in such cases.

Have a Third Party Administer Evaluations

In order to protect anonymity, evaluation questionnaires should not be administered by instructors or TAs. Rather, a responsible student or the department evaluation coordinator should be appointed to collect the completed questionnaires and deliver them to the department office.

Time questionnaire administration appropriately

Ratings should be collected during the last two weeks of the quarter. Students should be forewarned that evaluations will be done on a certain date so that they will be in class and will be prepared. Administration of evaluations at the final exam or other test is not recommended.

Emphasize the Importance of Evaluation

It is advisable to give students some context for the evaluation, especially for first year students. It is useful for them to know that the department and instructor value their comments, and the use to which they will be put. Distributing questionnaires at the beginning of the class period and allowing sufficient time for students to complete them, all contribute to the sense of importance placed upon a student's opinion, and are hence likely to produce more constructive results.

Suggested Readings:

Arreola, Raoul. (2000). *Developing a Comprehensive Faculty Evaluation System: A Handbook for College Faculty and Administrators on Designing and Operating a Comprehensive Faculty Evaluation System*. Bolton, MA: Anker Publishing Co., Inc.

Davis, B. G. (1993). *Tools for Teaching*. San Francisco, CA: Jossey-Bass.

Interpreting Quantitative Student Evaluations for a Course

Several weeks after the end of each quarter, instructors will receive their students' responses to the formal questionnaires, along with a summary sheet of statistical information, including mean, median, and standard deviation of the questionnaire responses. Instructors should keep the following points in mind when interpreting the results:

Procedures

Were reasonable procedures (such as suggested in the preceding section) used to collect the ratings?

Sample Size/Response Rate

How many students in the class completed the questionnaires? Responses from less than two-thirds of the class should be viewed with caution. The minimum number of student raters in a class should be 8 to 10; a sample of 15 or more students increases reliability and generalizability of the results. It should be noted, however, that even in such "small sample" situations the qualitative comments may still be extremely valuable.

Making Comparisons

The average ratings can be interpreted according to an absolute scale or relative to the ratings of other courses and instructors. For example, a mean rating of 6.5 on a 9-point scale for overall course evaluation may seem "above average." Taking them at their word, students rated this course as adequate. It may be, however, that 70 percent of similar courses are rated above 6.5. Thus, relative to other courses, the 6.5 rating was in the lower third. Which interpretation is correct? The answer is that both interpretations are useful. The course was judged positively, but students were not particularly excited by it.

The Campus Context

In making comparisons, it may be helpful to consider the campus context. Means do vary considerably between departments and divisions, or between different kinds of courses. Departments are, therefore, encouraged to keep records regarding their own norms.

Variability of Responses

The variability of student responses is important diagnostic information. For example, consider an average course rating of 7 on a 9-point scale with a small standard deviation, say of 1.0 or so. This means that most students rated the course as 6, 7, or 8. On the other hand, the same average rating with a larger standard deviation, say 3.4, indicates a greater range of ratings that may suggest problems with the course. It is also important to look at the total distribution of ratings. For example, there are sometimes bimodal distributions of ratings in which a large group of students really liked the course and another large group of students disliked it. This could indicate two different ability or interest groups in the class, which would be worth exploring further for future iterations of the course.

The Importance of Other Variables

Although student response to a course is important for evaluation and teaching improvement, it should not be used as the only measure of teaching. Student ratings are affected by many variables including course content, amount of work, expected grades, class size, and students' own needs and attitudes.

Suggested Readings:

Arreola, Raoul. (2000). *Developing a Comprehensive Faculty Evaluation System: A Handbook for College Faculty and Administrators on Designing and Operating a Comprehensive Faculty Evaluation System*. Bolton, MA: Anker Publishing Co., Inc.

Cashin, William E. (1990). *Student Ratings of Teaching: Recommendations for Use*. (IDEA Paper No. 22). Manhattan: Kansas State University, Center for Faculty Evaluation and Development.

Cashin, William E. (1995). *Student Ratings of Teaching: The Research Revisited*. (IDEA Paper No. 32). Manhattan: Kansas State University, Center for Faculty Evaluation and Development.

Davis, B. G. (1993). *Tools for Teaching*. San Francisco, CA: Jossey-Bass.

Interpreting Quantitative Student Evaluations for Personnel Decisions

Faculty members are often concerned about how evaluation results will be used by their departments in administrative personnel decision-making. Although details of the process differ among departments, the substantial body of research literature on the use of student course and instructor evaluations suggests that the following practices be observed:

➢ Guidelines for administration and explanations of how ratings will be used should be consistent and well-publicized.

➢ Small differences between scores should be kept in perspective.

➢ Multiple sets of student ratings should be used in administrative decision-making.

➢ Global ratings (i.e., overall ratings of the instructor) are more often used than specific items (such as the instructor's organization or communication skills) for making personnel decisions. While this may be the most convenient measure, decision-makers might note that global ratings are also those most likely to reflect personal bias on the part of students.

➢ Ratings for any single instructor or course should be considered in conjunction with university, college, division, department, and even specific course norms.

➢ Multiple sources of information should be used in administrative decision-making. In other words, numerical ratings should be only one piece of the larger picture.

Suggested Readings:

Arreola, Raoul. (2000). *Developing a Comprehensive Faculty Evaluation System: A Handbook for College Faculty and Administrators on Designing and Operating a Comprehensive Faculty Evaluation System*. Bolton, MA: Anker Publishing Co., Inc.

Cashin, William E. (1990). *Student Ratings of Teaching: Recommendations for Use*. (IDEA Paper No. 22). Manhattan: Kansas State University, Center for Faculty Evaluation and Development.

Cashin, William E. (1995). *Student Ratings of Teaching: The Research Revisited*. (IDEA Paper No. 32). Manhattan: Kansas State University, Center for Faculty Evaluation and Development.

Davis, B. G. (1993). *Tools for Teaching*. San Francisco, CA: Jossey-Bass.

Further Help for Faculty Concerned about Student Evaluations

Perhaps the most common concern that faculty members express about the evaluation process is that students do not take the evaluations seriously and that students are not aware of the gravity of their input into the tenure and merit review process. Research and experience show that instructors who openly announce to students that they themselves take student input seriously are usually the recipients of the most constructive comments.

The vast body of research on student ratings that has accumulated over the years also suggests that student ratings correlate highly with the ratings conducted by one's own colleagues and even, in some instances, with self-ratings. The high correlation holds up, however, only when students are asked to judge selected aspects of teaching. Students as consumers are well-equipped to assess the clarity of presentation, to comment on organizational matters, to rate the quality of student/instructor interaction, and even to assess their own learning to some extent. Students are not, however, the best judges of whether what they are learning is current, whether the instructor's perspective is biased, or even whether the selection of course material was appropriate for the achievement of course goals.

Instructors occasionally deride quarterly evaluations because they believe that students cannot make accurate judgments regarding a course or instructor until the students have been away from the course, or even the university, for several years. While it has proven to be very difficult for researchers to obtain a representative sample in longitudinal follow-up

studies, the research shows that, in general, alumni who had been out of school for five to ten years rated instructors much the same as did the students currently enrolled. Current research also provides little substantiation for the widely held belief that only warm, friendly, humorous instructors win the "ratings game." Most students are quite able to discriminate between glossy presentations with little substance and less flashy lectures with relevant content.

If a faculty member's number one concern about evaluation of instruction results is how they will be used in personnel decision-making, the number one concern among students is that their feedback will not be acted upon. It is, therefore, crucial that having conducted any type of feedback activity with students, instructors be seen to respond to the results. This may not be as easy as it sounds, since bimodal distributions, for example, can make obvious courses of action elusive.

It is important for faculty—particularly first-time faculty—to remember that some students can be insensitive or may lack the maturity necessary to deliver constructive criticism in a non-threatening manner. At times, their comments are overstated and short on diplomacy. While such comments can be very discouraging, if they come from only a few students, they represent only an extreme point of view. However, if such comments come from a majority of the students, advice from a trusted peer or from an objective consultant might be useful. Even if painful, they may contain insight into teaching issues that can be addressed – but again, only if they present a cogent argument, not just a personal attack.

Suggested Readings:

Arreola, Raoul. (2000). *Developing a Comprehensive Faculty Evaluation System: A Handbook for College Faculty and Administrators on Designing and Operating a Comprehensive Faculty Evaluation System*. Bolton, MA: Anker Publishing Co., Inc.

Davis, B. G. (1993). *Tools for Teaching*. San Francisco, CA: Jossey-Bass.

Chen, Y. & Hoshower, L. B. (2003). Student evaluation of teaching effectiveness: an assessment of student perception and motivation. *Assessment and Evaluation in Higher Education* 28(1): 71-88.

McKeachie, W. and Svinicki, M. (2006). *McKeachie's Teaching Tips*. Boston, MA: Houghton Mifflin Company.

Consultation

Faculty are well advised to seek consultation when deciphering their teaching evaluations. Quite often an outside perspective offers insight to evaluation data that may not be apparent to the recipient of the evaluation. Speaking with peers or mentors within one's department or at other universities may help discern new approaches to instructional improvement.

Instructors who wish for advice outside of a departmental context are welcome to contact the Office of Instructional Development's Faculty Consultation service. The office offers consultation to faculty on all aspects of teaching and learning. The objective is to help them reflect on their classroom practice and suggest strategies for improvement. Faculty Consultation services take the form of individual consultation by appointment, workshops for groups of interested faculty, and seminars in college teaching tailored for individual schools/departments. Professional consultation is given based on an analysis of some initial inquiry, or on the provision of such data as a classroom video taped by Audio Visual Services or a set of evaluation results administered through the Evaluation of Instruction Program (EIP). All consultations are strictly confidential and participation by faculty is entirely voluntary.

Suggested Readings:

Brinko, K. T. (1991). The interactions of teaching improvement. In M. Theall & Franklin (Eds.), Effective practices for improving teaching. *New Directions for Teaching and Learning*, No.48. San Francisco, CA: Jossey-Bass.

Cohen, P. A. (1980). Effectiveness of student rating feedback for improving college instruction: a meta-analysis of findings. *Research of Higher Education*, 13 (4), 321-341.

Seldin, P. et al. (1999). *Changing Practices in Evaluation Teaching: a Practical Guide to Improved Faculty Performance and Promotion/Tenure Decisions*. Bolton, MA: Anker Publishing Company, Inc.

Supplemental Types of Evaluation

While end of term student questionnaires continue to be the predominant type of instructor evaluation in most departments, various instructional needs often warrant the inclusion of additional evaluation strategies. The usefulness of an evaluation strategy depends on the desired level of detail necessary and the timeframe within which the data are required. The following section offers a range of techniques that vary from informal feedback strategies to help address current term teaching improvement, to more elaborate processes that garner in depth evaluative data for ongoing improvement.

Informal Feedback Strategies

Questioning—A very simple tool for checking effective teaching is to incorporate specific questions within a lesson to gauge student understanding of the material. For example, an instructor may ask students to verbally answer a question similar to one that will be asked on an exam. This tool is more useful than simply asking if students have any questions because students who are confused may not be able to articulate their questions. Moreover, some students may falsely believe they understand the lesson and not ask questions. Checking for understanding within a lesson helps the instructor discover students' level of learning and to make adjustments during the lesson itself.

Classroom Response Systems—A problem with simple questioning is that an instructor generally will get a response from only one or two students rather than the entire class. This problem can be resolved with a few strategies that fall under the Classroom Response umbrella.

The first strategy is the easiest to implement. An instructor asks a multiple choice question or makes an agree/disagree statement about the material. Students indicate by the position of their thumb whether they believe the answer is A (upright), B (sideways), or C (downward) or Agree (upright) or Disagree (downward). The instructor can then quickly look around the room to determine how many students have the correct answer.

The second strategy involves the use of colored index cards. Its method is identical to the first strategy except that the instructor is using color coded cards for the responses. The advantage of using colored index cards is that they are easier to see than thumbs.

The third strategy involves the use of hand-held remote controls ("clickers") to measure student responses. The technology is linked to software in a computer—either a laptop or a classroom computer—and can keep a record of student responses. Many instructors use this technology by imbedding the question into their presentation software. Both the instructor and students receive immediate feedback to the responses. In addition to the recordkeeping aspect of this system, a primary advantage of clickers is student anonymity in their responses in the classroom. A major disadvantage is the cost and performance reliability of the clickers themselves.

Open Class Discussion—This technique can be used either during the class session or by monitoring student online discussion. By asking discussion questions that require critical thought, instructors are able to gauge students' understanding of the lesson material and whether they are making necessary connections to other course material. Many times students believe they know the material but their misunderstandings are revealed during discussion.

Minute Paper—This evaluation tool is done at the end of class several times during the quarter. It derives its name from the fact that students spend no more than one minute answering any number of questions. The instructor reads the responses before the next class meeting and responds appropriately. Examples of questions asked are

➢ What was the most important thing you learned during class?

➢ What unanswered questions do you have?

➢ What was the muddiest point for you?

➢ At what point this week were you most engaged as a learner?

➢ Can you summarize today's lesson in one sentence? If so, please summarize it.

➢ What has been most helpful to you this week in learning the course material?

Index Card—A variation on the Minute Paper is for the instructor to write the responses to the following questions on a 3 x 5" index card following a lesson: "What worked? What didn't work? What are some ideas for changing the lesson?" The 3 x 5 card limits the amount of information than can be written down and serves as a reminder to write down ideas but to only spend a few minutes writing them down. Attach the card to the lesson notes to serve as a reminder the next time the lesson is taught.

Course Exams and Assignments—Student success on course exams and assignments are a powerful data source on teaching effectiveness. A short questionnaire at the end of exams can ask students to identify which questions were the most difficult to answer and why they were difficult. A pattern may develop that can be used to make changes. Additionally, an instructor

may ask students to critique assignments. Questions instructors may ask are:

➢ Were instructions clear?

➢ Did the assignment help students learn course material?

➢ Were the expectations reasonable for the time-frame?

➢ How many hours were devoted to completing the assignment?

Mid-Quarter Evaluation—An effective way of gauging student learning and satisfaction is via anonymous mid-quarter evaluations. The evaluations can take a variety of forms. A simple survey asking students to describe what is working, what is not working, and suggestions for change can be conducted via paper-pencil or online. Many of the course management systems have tools that allow anonymous feedback. Instructors need to check with their system's administrator to find out how to do it. Many instructors provide 15-25 minutes of class time to a neutral party for the purpose of getting feedback from students. A more formal method is to use the same forms that are used for course evaluations. One thing to note is that even if course changes cannot be made during the quarter the evaluation takes place, mid-quarter evaluations allow instructors to engage in dialogue with their students regarding the teaching-learning process and students will feel more positive toward the instructor.

Suggested Readings:
Angelo, T. A., & Cross, K. P. (1993). *Classroom Assessment Techniques: A Handbook for College Teachers, Second Edition*. San Francisco, CA: Jossey-Bass.

Davis, B. G. (1993). *Tools for Teaching*. San Francisco, CA: Jossey-Bass.

Duncan, D. (2005). *Clickers in the Classroom*. Upper Saddle River, NJ: Pearson Education.

Data from Students: Interviews with Students

A procedure using interviews with students is an excellent technique for obtaining rich, in-depth information about student reaction to courses and instructors. Two methods have proven useful. The first involves an interview with a group of students. The second procedure uses a series of interviews with single students.

The procedure is simply to have two or more colleagues (either in an instructor's department or from other departments) interview a group of students from a current course. Alternatively, a group of former students may be interviewed. The interview usually takes no longer than 30 minutes, and a brief two or three page report is completed by the interviewers. A group interview requires some planning, but it is not a difficult technique to use and it can yield valuable information. An instructor may request group interviews when he or she wants some candid student feedback at midterm, or after a course is completed. A department may also want to use this technique to get information for personnel actions such as tenure decisions, accelerations, or special teaching awards.

The Interviewers

Usually colleagues from the instructor's department or other departments do the interviews. It is best to use two interviewers, one to ask all the questions, the second to record the responses. Alternatively, the Faculty Consultation service offers consultation regarding the training of interviewers.

Conference with Instructor

The interviewers should meet the instructor to learn the course characteristics, goals, instructor's concerns and make arrangements for the interview. This meeting should decide the issues to be stressed in the interview, such as course objectives, organization, workload, instructor skills, instructor-student relationships, students' attitudes, and the like. The instructor may have special concerns that should be considered in structuring the interview. Decisions should be made about whether the report should be written and/or oral and who should receive copies.

Constructing the Interview Schedule

After meeting with the instructor, the interviewers should formulate some of the questions for the interview. Table 1.2 (see next page) provides suggested questions. Issues frequently arise during an interview which suggests a line of questioning not anticipated in the interview schedule. Probing students' comments may sometimes be more useful and appropriate than asking all the questions on the schedule.

Table 1.2	Sample Interview Schedule for a Focus Group with Students

The instructor of this class has asked for your feedback on his/her teaching of the class. Your comments will be treated with the strictest confidence. Suggestions from this focus group will be summarized and relayed anonymously to the instructor to improve his/her teaching. Please take a few minutes to answer the questions below before we discuss them as a group. Be as specific as possible, giving examples to illustrate your points. The more constructive you are, and the more suggestions you can give, the more you will help your instructor to improve, both in this and in other classes. Thank you in advance for your time.

1. Do you feel that your instructor is well-organized in this class? (Please explain your answer with an example. You might wish to comment on time management, presentation of concepts, or clarity of explanations)

2. Do you find it easy to identify the main points from each class? (Please comment with reference to ways the instructor helps you take notes, or uses other methods to summarize or emphasize main points.)

3. Do the exams/quizzes relate well to the class material? (Please comment on whether questions are fair, and reflect concepts taught in class.)

4. Is the feedback you receive on assignments helpful? How might feedback you receive be improved to help you to learn better?

5. Please identify key strengths your instructor could build on to improve his/her teaching.

6. Please identify any barriers which prevent you from learning in this class. Wherever possible please suggest specific ways in which the instructor could help you to overcome these barriers.

7. How appropriate was the instructor's use of technology in this class? How did the use of technology enhance your learning experience, if at all?

Interviewing the Students

A convenient time for the interview is the last 20-30 minutes of a class period. Unless class time is used for the interview, it is difficult to get a representative group. In conducting the interview, one interviewer should concentrate on asking the questions, the other on recording the answers and comments. Taping interviews is not recommended because of the problem of maintaining confidentiality.

The instructor should briefly discuss the purpose of the interview with the class before the interviewers arrive. After introducing the interviewers, the instructor should leave.

Group Size

The size of the group is an important variable. In a small group (less than 12-15), the interviewers can more easily probe for in-depth information. With a larger group, the number of students who want to comment makes in-depth coverage more difficult. You can divide the class group into smaller groups of about five and select one person from each group to act as a recorder and one as a spokesperson. If you are dividing the group up, then first have the smaller groups meet individually and arrive at a consensus on the predetermined questions. Then after 10 minutes of discussion time, have each spokesperson report one response to each of the questions.

Alternatively, to get a representative sample of students from large classes, the interviewers may want to use the instructor's course list and/or grade book to select a small number of students to participate in the interview. In larger classes, students may feel more comfortable writing down their responses before participating in a group discussion.

Suggested Readings:
Arreola, Raoul. (2000). *Developing a Comprehensive Faculty Evaluation System: A Handbook for College Faculty and Administrators on Designing and Operating a Comprehensive Faculty Evaluation System*. Bolton, MA: Anker Publishing Co., Inc.

Data from Peers

While peer review of teaching may take many forms, at UCLA it most often involves class observation. Classroom visitation is a form of evaluation strongly supported by faculty as a useful source of information. When colleague observation is undertaken for instructional improvement, the most important considerations in establishing systematic and fair procedures are:

Number and Timing of Visits

In courses taught exclusively by the lecture method, at least two visits by each colleague evaluator are advisable. If the instructor employs a variety of teaching strategies (such as lecture, discussion, student presentations, or role playing), it becomes very difficult to choose one or two class sessions that would be typical or would give a balanced picture of the instructor's teaching. In some small classes, the presence of an observer may be more distracting than in a larger class, and frequent observations by several colleagues during a single term might be problematic. The number and timing of visits should probably be worked out between colleague evaluators and the faculty member being evaluated, to assure an adequate evaluation with minimal disruption.

Explicit, Appropriate Criteria and Guidelines

A set of explicit criteria by which colleague observers are asked to judge the quality of teaching will make the evaluations much more reliable and the evaluations made by different colleagues more comparable. For colleagues observing strictly for the purpose of evaluation, the criteria help to guide the observations. For colleagues who have ongoing contact and observation, they help to summarize the impressions developed over numerous observations. The number of criteria should be kept small and appropriate to the type of teaching done in the department. The format may consist of open-ended questions, or rating scales, or a combination of these.

Criteria should reflect aspects of teaching on which there is broad departmental consensus and for which faculty observers would be in the best position to provide information. For example, faculty observers might be asked how well prepared the instructor was for the class session, but should not be asked to comment on the instructor's accessibility to students outside of class if this has not been observed. Colleagues may reasonably be asked to comment on the instructor's coverage of a topic or on the appropriateness of the teaching strategy, but should not be asked to evaluate student motivation or satisfaction, which can only be inferred at best. Comments about actual student participation, however, would be appropriate.

Special Teaching Situations

Colleague observation and evaluation of clinical teaching present problems analogous to those which arise with small classes. Clinical teaching takes place most often in a one-to-one or small group context. In this case, the presence of several colleague observers would be intrusive and might significantly disrupt the teaching and practice situation. On the other hand, many clinical settings provide a natural situation for colleague observation. Indeed, colleagues often work side by side in clinical settings and frequently observe one another's teaching. Observation for evaluation should take advantage of these opportunities in much the same manner as evaluation based on observation in team-taught courses. There should be some indication as to the content, frequency, and length of the observations on which the evaluation is based. Where such natural observation is not possible and special visits for colleague observation-evaluation are needed, only one observer should have an opportunity to make several observations over a period of time. The development of criteria for the evaluation of clinical teaching should follow the guidelines for regular courses, although the specific items would be different and would be sensitive to the nature and purpose of clinical teaching. For example, in medicine, the observer might be asked to comment on the instructor's integration of biomedical theory and clinical management or on the actual demonstration of a procedure or technique.

Constructive Feedback

Feedback to the faculty member is an important consideration in designing departmental peer review procedures. For evaluations to be useful for the improvement of teaching, feedback and discussion are essential, yet this may present certain problems of confidentiality when colleague evaluation is conducted as part of the personnel process.

Sensitive Implementation

Many instructors are understandably anxious about peer evaluation. Departments implementing new systems of colleague observations should be sensitive to the problems and insecurities among faculty that will inevitably arise. The suggestions summarized above are also useful for instructors to keep in mind when observing their TAs in discussion or laboratory sections and writing up their evaluations.

Suggested Readings:

Arreola, Raoul. (2000). *Developing a Comprehensive Faculty Evaluation System: A Handbook for College Faculty and Administrators on Designing and Operating a Comprehensive Faculty Evaluation System*. Bolton, MA: Anker Publishing Co., Inc.

Seldin, P. et al. (1999). *Changing Practices in Evaluation Teaching: a Practical Guide to Improved Faculty Performance and Promotion/Tenure Decisions*. Bolton, MA: Anker Publishing Company, Inc.

Data from Oneself

In conducting a self-review of a course, faculty members may wish to compare their own pre-course objectives or expectations with perceived post-course outcomes. A model whereby the instructor assesses the abilities and knowledge of students before and after the introduction of an innovation or improvement effort is especially useful for evaluating new courses or courses with significant changes in content or structure. Often, however, instructors may not have planned for a self-evaluation before the course began; the self-evaluation can then look only at course outcomes.

When faculty members are interested in examining their own teaching behavior, rather than course outcomes, they can follow an end-of-course format similar to that used by students. An instructor might use the same form and complete it from the point of view of self-perceived behavior. Alternatively, they might benefit from completing the student form from the perspective of what they expect, on the average, that students will say; instructors might complete a self-evaluation form at the same time that students complete their evaluations and then compare results. Also, instructors can arrange to videotape a class and then observe their performance, focusing on particular teaching skills of interest. While this latter technique can be extremely valuable, it is usually best achieved with the help of a faculty consultant, who can help the instructor to focus on the key elements. Individuals looking at videotapes of themselves often are biased towards seeing only the negative elements.

Faculty should expect that their self-review in most cases will be more favorable than reviews by students. If these self-evaluations are included as part of the dossier (see "Documenting Teaching Using a Teaching Portfolio or Dossier" following), instructors may wish to comment on any deficiencies noted or should not attempt to defend one's self-review as being more accurate than that of the students, or to discourse on every

aspect of the course. Rather it should be considered as additional information that can assist reviewers in interpreting student data or in understanding how the self-review contributed to course changes or modifications in teaching style.

Suggested Readings:

Arreola, Raoul. (2000). *Developing a Comprehensive Faculty Evaluation System: A Handbook for College Faculty and Administrators on Designing and Operating a Comprehensive Faculty Evaluation System*. Bolton, MA: Anker Publishing Co., Inc.

Seldin, P. et al. (1999). *Changing Practices in Evaluation Teaching: a Practical Guide to Improved Faculty Performance and Promotion/Tenure Decisions*. Bolton, MA: Anker Publishing Company, Inc.

Documenting Teaching Using a Teaching Dossier or Portfolio

At a research university such as UCLA, faculty are generally more aware of colleagues' research than their teaching. Further, getting information about colleagues' teaching is more difficult than getting information about their research. There are usually a variety of concrete materials to use in evaluating research, including articles, books, artistic contributions, expert testimony, and various measures of professional recognition. Because teaching is usually a transaction between a single faculty member and one or more students, colleagues' knowledge of this transaction is necessarily limited. The teaching dossier or portfolio, analogous to an art portfolio, has been proposed as one solution to the issue of documenting effective teaching.

Such a portfolio usually contains a thoughtful and thorough narrative supported by appendices of teaching materials. Portfolios are often an essential part of the personnel decision-making process for instructors. Also, a good self-review narrative can be very useful for improving one's teaching, for winning a teaching award and for applying for future positions. As was stated earlier, documenting teaching effectiveness often leads to teaching improvement, and even reward.

Possible Contents of a Teaching Portfolio

Seldin (1991) lists the kinds of materials that one might include in a teaching portfolio. The specific items selected will depend on an instructor's particular teaching assignment and activities. However, the real secret of assembling a successful portfolio is knowing whom to ask for what. Some of his ideas for sources of data are given below. The Office of Instructional Development would like to thank CELT (The Center for Excellence in Learning and Teaching) at The Pennsylvania State University for permission to adapt their "Creating a Teaching Portfolio" guide for the UCLA context in this section.

Portfolio Sources
Data from Oneself

Self-analysis and self-reflection are far too often overlooked in the assessment of teaching and learning, yet they are central not only to the processes of assessing teaching, but also of improving it. Thus they are an essential part of any teaching portfolio.

Faculty members can provide their own perspective on virtually every aspect of instruction. Self-reports should be primarily descriptive as opposed to evaluative—what were you trying to do, why, how, and what was the result? Consequently, these documents will more easily reflect development than those from other sources. Self-reports should also be compared with data from other sources. Because feedback that provides new information is most likely to produce change, it is by virtue of such comparisons that personal growth and improvement occur. Data from oneself might include:

➢ a list of courses taught, with brief descriptions of course content, teaching responsibilities, and student information

➢ a statement of philosophy of teaching and factors that have influenced that philosophy

➢ examples of course material prepared and any subsequent modifications that were made to accommodate unanticipated student needs

➢ a sample syllabus or lesson plan

➢ a record of teaching discoveries and subsequent changes made to courses regularly taught

➢ a description of efforts to improve teaching (e.g., participating in seminars and workshops, reading journals on teaching, reviewing new teaching materials for possible application, pursuing a line of research that contributes directly to teaching, using instructional development services, and contributing to a professional journal of teaching in your discipline)

➢ evidence of reputation as a skilled teacher, such as awards, invitations to speak, and interviews

➢ personal reflections on your growth and change as a teacher (including awards won and indicating future teaching promise)

Data from Others

Obviously, different people can provide different kinds of information about an instructor's teaching. For example, it is probably counterproductive and inappropriate to ask students about the breadth and completeness of an instructor's content knowledge since, from their point of view, such expertise should be a given. The more obvious and appropriate judges of this information would be department colleagues. Likewise, such colleagues are usually not good judges of whether an individual is prepared for class, arrives on time, or is available for office hours. Clearly, getting the right kinds of input from each group of individuals is what will give a portfolio its strength and depth.

Students

As the immediate beneficiaries of teaching, students are in an ideal position to report and comment on a number of variables, such as which instructional strategies helped them learn the most and whether the instructor came prepared to class, was available during office hours, or provided useful comments on papers. Other data that only students can report involve any change in their level of interest as a result of taking the course, how much the course challenged them, and whether they felt comfortable asking questions. The most common ways of obtaining student feedback about these aspects of teaching include:

➢ interviews with students after they have completed the course

➢ informal (and perhaps unsolicited) feedback, such as letters or notes from students

➢ systematic summaries of student course evaluations—both open-ended and restricted choice ratings

➢ honors received from students, such as winning a Distinguished Teaching Award.

Other materials often referred to in the literature on teaching portfolios are the "products of good teaching." In a sense, these are really a subspecies of the broader category "data from students" and might include:

➢ examples of the instructor's comments on student papers, tests, and assignments

➢ pre- and post-course examples of students' work, such as writing samples, laboratory workbooks or logs, creative work, and projects or fieldwork reports

➢ testimonials of the effect of the course on future studies, career choice, employment, or subsequent enjoyment of the subject

Colleagues

Colleagues within the department/school are best suited to make judgments about course content and objectives, the instructor's collegiality, and student preparedness for subsequent courses. Departmental/school colleagues can provide analyses and testimonials that serve as a measure of:

➢ mastery of course content

➢ ability to convey course content and objectives

➢ suitability of specific teaching methods and assessment procedures for achieving course objectives

➢ commitment to teaching as evidenced by expressed concern for student learning

➢ commitment to and support of departmental/school instructional efforts

- ability to work with others on instructional issues

- reports from classroom observations

- statements from those who teach courses for which the instructor's course is a prerequisite

- evidence of contributions to course development, improvement, and innovation

Administrators

The most significant administrators who may not necessarily be included in the previous category are the department chair and the school or divisional dean. In their supervisory capacity, these administrators are generally well suited to make summary statements about overall performance over time. In doing so, they can help those who will read and interpret the portfolio by organizing and assimilating all the other information from various sources. It is also appropriate that these individuals draw attention to special recognition for teaching such as a university-wide or department/school teaching award, such as the university Distinguished Teaching Award.

Arrangement and Presentation of Portfolio Components

An instructor's teaching portfolio is, and indeed should be, highly personal. There is, therefore, no specifically recognized format. In the most general sense, such a portfolio is likely to contain a short reflective narrative followed by an appendix of supporting documentation. Beyond this, selection and arrangement should be done so as to best reflect the argument you wish to make (for example, that one should be selected for the job, or be given departmental funding to teach a course).

Reflective Narrative

This is a key piece of any portfolio. It includes the major claims an instructor wishes to make about their teaching, and indicates how these claims support their case. It is always advisable to use specific examples which narrate these claims and which give them flavor. For this one can draw on the "Data from Oneself" section.

Supporting Materials/Data/Documents

These elements are used to illustrate the claims and examples in the reflective narrative, and hence to support the overall argument. For this one can draw mostly, though not exclusively, on the "Data from Others" section. One might include, for example, a table of standardized student evaluations, as well as a sample lesson plan or syllabus. Supporting materials are most conveniently located in appendices. They need to be carefully selected so as to not be too lengthy (just pick the clearest example to support the point), and should be arranged and labeled for the convenience of the reader. Points made in the narrative should be referenced to specific pages or parts of the appendices if at all possible.

Getting Feedback as Part of the Process

Once a portfolio outline is complete, and before the final draft is written, it is always a good idea to check it for balance. In particular, it is important to make sure that the "data from others" comes from multiple sources (students as well as colleagues).

In addition to making a strong case for an applicant, a portfolio should reflect the instructor as a person. As is the case for teaching in general, the best portfolios are those that are constantly revised and updated. Input from colleagues and friends can be invaluable in this process.

Suggested Readings:

Davis, B. G. (1993). *Tools for Teaching*. San Francisco, CA: Jossey-Bass.

Edgerton, R., Hutchings, P., & Quinlan, K. (1991). *The Teaching Portfolio: Capturing the Scholarship in Teaching*. AAHE Teaching Initiative. Washington, D.C.: AAHE.

Seldin, Peter. (1993). *Successful Use of Teaching Portfolios*. Bolton, MA: Anker Publishing Co.

Seldin, Peter. (1991). *The Teaching Portfolio: A Practical Guide to Improved Performance and Promotion/Tenure Decisions*. Bolton, MA: Anker Publishing Co.

Shore, Bruce M., et al. (1986). *The Teaching Dossier: A Guide to Its Preparation and Use*. Ottawa: Canadian Association of University Teachers.

Zubizarreta, J. Evaluating teaching through portfolios. In Seldin, P. et al. (1999). *Changing Practices in Evaluation Teaching: a Practical Guide to Improved Faculty Performance and Promotion/Tenure Decisions*. Bolton, MA: Anker.

Developing a Comprehensive Departmental System of Evaluation

The routine use of all of the evaluation methods listed is not recommended. Rather, it is suggested that departments develop comprehensive systems which use these different teaching evaluation methods selectively and appropriately to assess the quality of instruction and to provide a valid picture of an individual's teaching. The basic elements of a comprehensive system are:

Instructor Consent

Teaching evaluations should only be done with knowledge and consent of the instructor.

Teaching Committee

It is recommended that departments, when possible, establish teaching committees to monitor the quality of instruction and conduct teaching evaluations.

Teaching Files

In order to develop a continuous record of evaluation, it is recommended that all departments maintain teaching files for all faculty. Systematic student ratings, course materials, and other information relating to teaching would be placed in the file on an ongoing basis. Such materials can then be used by the faculty member or department to construct a teaching dossier, or portfolio whenever one is needed.

Regular Evaluations

In many departments, end-of-course student ratings and faculty self-reviews are sufficient for regular evaluation.

Special Evaluations

It is not necessary to evaluate in great detail all courses taught by all faculty every quarter. Extensive evaluations using classroom visits or interviews with students should be reserved for instructors who request additional evaluation in order to improve his/her teaching, or when a department has questions about the quality of teaching for critical personnel decisions. Extensive evaluation is also appropriate for new course development or at times of major curriculum reorganization or development.

Different Department Systems

Teaching assignments vary within and between departments, and it is expected that departments will develop their own systems which emphasize different evaluation methods. The following section contains an example of a particular departmental evaluation system.

Information for Students

Certain information from the comprehensive evaluation system can be useful for students in program planning and course selection. Information about the course organization, reading, grading and test procedures, previous student evaluations and faculty comments about their course might be provided to students with the instructor's consent.

A comprehensive system of teaching evaluation will require more time and effort than our current methods, but the additional effort to develop a credible and equitable system is well worth it. Many faculty believe that such decisions warrant a comparable commitment of time. More importantly, given the potential impact on the faculty member's professional career, especially with the increased attention given to university instruction and its outcomes, evaluation of teaching effectiveness merits a serious commitment from everyone involved to provide an accurate and fair assessment. Perhaps most importantly, the process helps everyone on the faculty to become more effective teachers.

Both departmental instructional programs and faculty teaching assignments vary widely. A single fixed set of rules and procedures for evaluation cannot meet the needs of all departments. However, the principles outlined in Table 1.3 (see next page) should provide a useful framework on which departments can build their own evaluation program.

| Table 1.3 | An Example of a Departmental System |

Department X is a moderate sized department (35 FTE) with large undergraduate and graduate programs. Because there are many TA-run discussion sections in the large lecture and lab courses, graduate teaching assistants are an important part of the instructional effort. The department has a Teaching Committee to monitor the evaluation system and to advise faculty on teaching.

Instructor Teaching Files

Files are maintained for all faculty by the staff person who handles academic personnel. Faculty are asked to put copies of all syllabi, other instructional materials, evaluations, and the like, in their teaching file. Faculty are required to get systematic student ratings for all of their courses using a standard form for the evaluation.

Senior and Alumni Surveys

As part of periodic Academic Senate Reviews, the department has surveyed seniors and alumni about both the program and the quality of faculty instruction. The results for all faculty are summarized and placed in the teaching files.

Faculty Portfolio or Dossier Preparation

Faculty under consideration for merit increases or promotion are asked to prepare a reflective teaching narrative. The narrative includes a self-appraisal in which faculty are asked for self-ratings and comments on student ratings, as well as a teaching activities report in a standard format. Faculty have access to all materials in their teaching file when preparing their dossier and should carefully select appropriate samples for the appendices to their narrative.

Special Review Reports

Faculty at critical career points: tenure decisions, full professor, professor VI, and also for some accelerated advancements, receive a more careful review using group interviews with students, classroom visits, solicited letters from alumni, or other methods depending on what is most appropriate for the teaching responsibilities of the faculty member under consideration. When possible, these special procedures are carried over the year or so preceding the review for advancement.

Teaching/Learning Consultation

Faculty may request special reviews for teaching improvement purposes at any time. Also, the department may suggest to some faculty, based on their departmental review, that they should obtain diagnostic appraisal and assistance using the various programs offered by the department, or the campus Faculty Consultation Program.

Departmental Committees

The department review procedure for advancement uses an ad-hoc subcommittee. All of the material collected for teaching, research, and service is evaluated by this committee and a report and recommendation is made to the full department which votes on the appointment.

Teaching Colloquium

All candidates for departmental hiring are required to give a teaching colloquium as part of the interview process. Students, graduates and faculty are invited to attend and vote on the candidate's performance.

Suggested Readings:

Arreola, Raoul. (2000). *Developing a Comprehensive Faculty Evaluation System: A Handbook for College Faculty and Administrators on Designing and Operating a Comprehensive Faculty Evaluation System*. Bolton, MA: Anker Publishing Co., Inc.

Seldin, P. et al. (1999). *Changing Practices in Evaluation Teaching: a Practical Guide to Improved Faculty Performance and Promotion/Tenure Decisions*. Bolton, MA: Anker Publishing Company, Inc.

Literature Review

There is a vast body of work regarding the evaluation of instruction. The following items are references which have been found most useful, or which have been influential in shaping the discussion of evaluating teaching.

Definitions/Theories of Teaching Effectiveness

Suggested Readings:

Bain, K. (2004). *What the Best College Teachers Do*. Cambridge, MA: Harvard Press.

Cashin, W. E. (1989). *Defining and evaluating college teaching*. (IDEA Paper No. 21). Manhattan: Kansas State University, Center for Faculty Evaluation and Development

Cohen, P. A. (1980). Effectiveness of student rating feedback for improving college instruction: a meta-analysis of findings. *Research in Higher Education*, 13, 321-341.

Hobson, S. M. & Talbot, D. M. (2001). Understanding student evaluations: what all faculty should know. *College Teaching* 49(1): 26-31.

Marsh, H. W. & Roche, L. A. (1997). Making students' evaluation of teaching effectiveness effective: the critical issues of validity, bias, and utility. *American Psychologist* 52(11): 1887-1197.

McKeachie, W. (1997). Student ratings; the validity of use. *American Psychologist* v52, no.11, 1218-1225.

McKeachie, W. and Svinicki, M. (2006). *McKeachie's Teaching Tips*. Boston, MA: Houghton Mifflin Company.

Ory, J. (2000). Teaching evaluation: past, present, and future. *New Directions for Teaching and Learning*, no.83, Fall 2000.

The criterion for teaching effectiveness varies depending on the goals of individual instructors and the mission of their institution. In an effort to address the multi-dimensionality of teaching goals and responsibilities, teaching effectiveness is typically defined in terms of student learning (Hobson & Talbot, 2001). However, focusing on student learning exacerbates the difficulty of defining teaching effectiveness, because learning assessment encompasses equally broad criterion that range from personal development, to increased knowledge, to self-discipline, to career development. Marsh (1983) identified nine factors related to teaching effectiveness: learning/value, instructor enthusiasm, organization/clarity, group interaction, individual rapport, breadth of coverage, examinations /grading, assignments/readings, and workload/difficulty.

While research supports Marsh's multi-dimensional factor analysis, a recent *outcome assessment movement* supports quantifying student learning in terms of job placement outcome (Ory, 2000). This approach attempts to take the onus of evaluation-based decisions off student rating systems, which critics claim are too subjective. Some of the subjectivity in the evaluation process results from students rating areas of teaching effectiveness that they are unable to access, which they identify as, 1) the goals, content, and organization of course design; 2) methods and materials used in delivery; and 3) evaluation of student work, including grading practices (Cohen, 1980; Cashin, 1989; McKeachie, 1997;2006; Bain, 2004).

Student Perspective on Teaching Evaluations

Suggested Readings:

Chen, Y. & Hoshower, L.B. (2003). Student evaluation of teaching effectiveness: an assessment of student perception and motivation. *Assessment and Evaluation in Higher Education* 28(1): 71-88.

Sojka, J., Gupta, A.K. & Deeter-Schmetz, D.R. (2002). Student and faculty perception of student evaluations of teaching: a study of similarities and differences. *College Teaching* 50(2): 44-49.

Research shows that students tend to take teaching evaluations more seriously than faculty and institutional members commonly believe. Students are more willing to participate and offer meaningful feedback when they believe and can see that their input is being considered and incorporated by their instructors and the institution. In general, however, students do not perceive that their feedback is being used. Some studies show that students place most value on evaluations for formative purposes, but research also indicates that students believe their

input should be considered for summative purposes. Students would like to see more specific items related to teaching effectiveness on student evaluation of teaching instruments (Sojka & Deeter-Schmetz, 2002; Chen & Hoshower, 2003).

Faculty Perspective on Teaching Evaluations

Suggested Readings:

Bain, K. (2004). *What the Best College Teachers Do*. Cambridge, MA: Harvard Press.

Gallagher, T. G. (2000). Embracing student evaluations of teaching: a case study. *Teaching Sociology* 28, 140-147.

Hativa, N. (1995). The department-wide approach to improving faculty instruction in higher education: a qualitative evaluation. *Research in higher education* 36(4): 377-413.

Sojka, J., Gupta, A. K. & Deeter-Schmetz, D. R. (2002). Student and faculty perception of student evaluations of teaching: a study of similarities and differences. *College Teaching* 50(2): 44-49.

Traditionally, faculty have been very skeptical of teaching evaluations and have previously opposed such practices. These negative feelings often spring from fear that student ratings will be used or misused for summative decision-making purposes. Moreover, faculty often believe that students do not take evaluations seriously and that ratings encourage grade leniency. Nonetheless, most faculty do pay attention to student feedback. Further, when evaluations are used for formative purposes, instructors show a high degree of motivation to improve their teaching based on student input. Studies have emerged showing how institutions and individual faculty members have begun using evaluations, consultations, and portfolios to qualitatively improve instruction. When faculty are well informed about the purposes of evaluation, much of their anxiety dissipates and willingness to learn from student feedback increases (Sojka, Gupta, & Deeter-Schmetz, 2002; Hativa, 1995; Gallagher, 2000; Bain, 2004).

Summative Uses for Teaching Evaluations

Suggested Readings:

Bain, K. (2004). *What the Best College Teachers Do*. Cambridge, MA: Harvard Press.

Hobson, S. M. & Talbot, D. M. (2001). Understanding student evaluations: what all faculty should know. *College Teaching* 49(1): 26-31.

Hoyt, D. P. & Pallett, W. H. (1999). Appraising teaching effectiveness: beyond student ratings. *IDEA Paper No. 36*. Kansas State University.

Kulik, J. A. (2001). Student ratings: validity, utility, and controversy. *New Directions for Institutional Research* 109, 9-25.

McKeachie, W. (1997). Student ratings; the validity of use. *American Psychologist* v52, no.11, 1218-1225.

McKeachie, W. and Svinicki, M. (2006). *McKeachie's Teaching Tips*. Boston, MA: Houghton Mifflin Company.

Theall, M. & Franklin, J. (2001). Looking for bias in all the wrong places: a search of truth or a witch hunt in student ratings of instruction? *New Directions for Institutional Research* 109: 45-56.

Teaching evaluations are commonly considered for summative purposes, including tenure, merit increase, retention for non-tenured faculty, promotion, and course assignment decisions. While research generally agrees that teaching evaluations can be used in an effective and meaningful way to inform these decisions, often such data are misused, misinterpreted, or overused. Some institutions use student ratings data as the sole criterion for evaluating teaching effectiveness, and moreover, these institutions often use only global items on student ratings forms to construct their evaluation. Such misuse can breed distrust between faculty and administrators, resentment on the part of instructors for evaluations, and hinder other formative uses of these data.

Instead, researchers recommend that when such data are going to be used for summative purposes, various stakeholders, including administrators, faculty and students, should collaborate

in determining a proper evaluation system. The focus of evaluation should be on desired educational outcomes and whether these outcomes are being met. If student ratings forms are to be used, the instruments must be subjected to rigorous validity tests and analysis. Further, student rating data should be used in combination with other criteria in order to provide a better assessment of teaching, which is inherently a multidimensional construct. Multiple sources of data collected over a span of time and covering a variety of courses taught would be most effective in informing summative decision-making (Hobson & Talbot, 2001; Hoyt & Pallett, 1999; Theall & Franklin, 2001; Kulik, 2001; McKeachie, 1997; 2002; Bain, 2004).

Formative Uses of Teaching Evaluation

Suggested Readings:

Bain, K. (2004). *What the Best College Teachers Do*. Cambridge, MA: Harvard Press.

Gallagher, T. G. (2000). Embracing student evaluations of teaching: a case study. *Teaching Sociology* 28, 140-147.

Hativa, N. (1995). The department-wide approach to improving faculty instruction in higher education: a qualitative evaluation. *Research in higher education* 36(4): 377-413.

Hobson, S. M. & Talbot, D. M. (2001). Understanding student evaluations: what all faculty should know. *College Teaching* 49(1): 26-31.

Hoyt, D. P. & Pallett, W. H. (1999). Appraising teaching effectiveness: beyond student ratings. *IDEA Paper No. 36*. Kansas State University.

Johnson, T. D., & Ryan, K. E. (2000). A comprehensive approach to the evaluation of college teaching. *New Directions for Teaching and Learning*, no. 83. Fall 2000.

Kulik, J. A. (2001). Student ratings: validity, utility, and controversy. *New Directions for Institutional Research* 109, 9-25.

Theall, M. & Franklin, J. (2001). Looking for bias in all the wrong places: a search of truth or a witch hunt in student ratings of instruction? *New Directions for Institutional Research* 109: 45-56.

Using evaluations to inform instructors of their teaching effectiveness and to aid them in improving or enhancing their teaching constitute the formative purposes of teaching evaluations. When used to inform teaching practices, specific dimensions of teaching must be identified and focused upon in order to bring about change. Research indicates that evaluations are most effective in improving teaching when faculty members understand and value the importance of such processes, and an institutional and departmental culture that supports and respects teaching is evident.

In particular, studies also indicate that mid-semester evaluations and feedback accompanied with consultation from a faculty developer or peer are more effective than traditional practices that leave the instructor to interpret end of semester findings by him/herself. Evaluation systems for formative purposes often encompass more than just student ratings of teacher effectiveness. Institutions have begun using portfolios, peer observation, self-review, and more qualitative approaches to improve teaching. Similarly, recent establishment of faculty development centers on many campuses reveals a trend toward investing in the formative uses of evaluations (Hobson & Talbot, 2001; Hoyt & Pallett, 1999; Theall & Franklin, 2001; Kulik, 2001; Gallagher, 2000; Johnson & Ryan, 2000; Hativa, 1995; Bain, 2004).

On-line Evaluations

Suggested Readings:

Johnson, T. (2002). *Online student ratings: will students respond?* AERA, New Orleans, LA, April 1-5, 2002.

Sorenson, D. L. & Johnson, T. D. (2003). Online student ratings of instruction. *New Directions for Teaching and Learning*, no. 96.

Many institutions are moving to using on-line student ratings forms due to the ease of administration and affordability of this format of evaluation. However, institutions are concerned

about the lowered response rates these forms tend to generate. Research shows that response rates can be bolstered through faculty encouragement, reward, and requirement of student completion of forms. On-line forms tend to increase the likelihood that students will offer written comments to the open-ended questions (Johnson, 2002).

Organizational Structure for Evaluation Systems

Suggested Readings:

Arreola, Raoul. (2000). *Developing a Comprehensive Faculty Evaluation System: A Handbook for College Faculty and Administrators on Designing and Operating a Comprehensive Faculty Evaluation System*. Bolton, MA: Anker Publishing Co., Inc.

Hativa, N. (1995). The department-wide approach to improving faculty instruction in higher education: a qualitative evaluation. *Research in higher education* 36(4): 377-413.

Seldin, P. et al. (1999). *Changing Practices in Evaluation Teaching: a Practical Guide to Improved Faculty Performance and Promotion/Tenure Decisions*. Bolton, MA: Anker Publishing Company, Inc.

Institutions structure and locate their evaluation systems differently on their campuses. In general, institutions either organize evaluations through a central office on campus or within individual departments. With regard to the formative purposes of evaluation, the organizational location of evaluation systems can potentially influence the effectiveness of those systems. In particular, certain researchers advocate a departmental approach to evaluation. Advantages with this approach are: instruction and its improvement are often discipline specific and faculty most closely identify with their department and are more likely to seek guidance and information within their department (Hativa, 1995).

Consultation

Suggested Readings:

Cohen, P. A. (1980). Effectiveness of student rating feedback for improving college instruction: a meta-analysis of findings. *Research in Higher Education*, 13, 321-341.

Marsh, H. W. & Roche, L. (1993). The use of students' evaluations and an individually structured intervention to enhance university teaching effectiveness. *American Educational Research Journal*, 30 (1), 217-251.

Murray, H. (1984). The impact of formative and summative evaluation of teaching in North American universities. *Assessment and Evaluation in Higher Education*, 9 (2), 117-132.

Seldin, P. et al. (1999). *Changing Practices in Evaluation Teaching: a Practical Guide to Improved Faculty Performance and Promotion/Tenure Decisions*. Bolton, MA: Anker Publishing Company, Inc.

The preponderance of evidence regarding instructional improvement based on the results of student evaluations of teaching shows that consultation plays a major role in implementing positive change. While many institutions are incorporating professional teaching consultants into their administration, speaking with peers, teaching assistants, and students also benefit instruction (Cohen, 1980; Murray, 1984; Marsh & Roche, 1993; Seldin, 1999).

Other Forms of Evaluation

Suggested Readings:

Angelo, T. A., & Cross, K. P. (1993). *Classroom Assessment Techniques: A Handbook for College Teachers, Second Edition*. San Francisco, CA: Jossey-Bass.

Davis, B. G. (1993). *Tools for Teaching*. San Francisco, CA: Jossey-Bass.

Duncan, D. (2005). *Clickers in the Classroom*. Upper Saddle River, NJ: Pearson Education.

Hativa, N. (1995). The department-wide approach to improving faculty instruction in higher education: a qualitative evaluation. *Research in higher education* 36(4): 377-413.

Hoyt, D. P. & Pallett, W. H. (1999). Appraising teaching effectiveness: beyond student ratings. *IDEA Paper No. 36*. Kansas State University.

Johnson, T. D., & Ryan, K. E. (2000). A comprehensive approach to the evaluation of college teaching. *New Directions for Teaching and Learning*, no. 83. Fall 2000.

In addition to student rating forms assessing teaching effectiveness, institutions also employ practices like portfolios, peer review, consultation, self-reports, alumni ratings, document review, student focus groups and interviews, observation, and department review to evaluate teaching. Research advocates using different forms for evaluating different teaching methods and contexts for teaching. Employing different types of evaluation have helped better inform formative and summative decisions.

Research on the validity and reliability of these different forms of evaluation is scarce and especially in comparison to the amount of research available on the usefulness of student rating forms. However, studies do recognize that these other forms of evaluation can provide more qualitative and comprehensive information but may nonetheless be biased. In addition, these approaches require significant amounts of time and commitment on the part of the institution and faculty members. Institutions are continuing to incorporate more of these methods, and research has begun examining the effectiveness and utility of these approaches as well (Johnson & Ryan, 2000; Hoyt & Pallett, 1999; Hativa, 1995).

Student Evaluations of Teaching (SET)
Evaluation forms

Suggested Readings:

Abrami, P. C. & d'Apollonia, P. A. (1990). Validity of student ratings of instruction: what we know and what we do not. *Journal of Educational Psychology* 82(2): 219-31.

Centra, J. A. (1993). *Reflective Faculty Evaluations*. San Francisco: Jossey-Bass.

Hobson, S. M. & Talbot, D. M. (2001). Understanding student evaluations: what all faculty should know. *College Teaching* 49(1): 26-31.

Hoyt, D. P. & Cashin, W. E. (1977). Development of the IDEA system (IDEA Tech. Rep. No.1). Manhattan: Kansas State University, Center for Faculty Evaluation and Development.

Marsh, H. W. (1983). Multidimensional ratings of teaching effectiveness by students from different academic settings and their relation to student/course/instructor characteristics. *Journal of Educational Psychology*, 75, 150-166.

According to Hobson & Talbot (2001), the three most widely used student evaluation forms are published instruments based on specific and global, self-reported student rating questionnaires. Each system attempts to measure teaching effectiveness in terms of student learning and may be used for either summative or formative purposes.

The Instructional Development and Effective Assessment (IDEA) form was developed by Hoyt & Cashin in 1977 at the Center of Faculty Evaluation and Development, Kansas State University. The form was most recently revised in 1998, but adheres to its original premise that students rate instructors on a series of questions addressing their perceptions of the instructor's ability to teach and design courses. The form is paper and pencil based and has 46 self-reported questions that are both specific and global in nature.

Student Instructional Report was first developed by the Educational Testing Service in 1971. It was most recently revised in 1989 and is made up of "thirty-nine questions, plus space for responses to ten additional questions that may be inserted locally" (Centra, 1993, p.188). It asks both specific questions regarding the ability and behavior of the instructor, as well as global questions regarding the overall value of the course.

The Student's Evaluation of Education Quality (SEEQ) was first developed in 1976 by Marsh, who at the time was living in Australia. It was revised in 1991 to reflect American usage of the English language. The form uses a multi-dimensional approach to evaluation based on a nine-factor analysis. It is made up of thirty-five questions that pertain to nine areas of teaching: learning/value, instructor enthusiasm, organization/clarity, group interaction, individual rapport, breadth of coverage, examinations /grading, assignments/readings, and workload/difficulty. The instrument can be used for both summative and formative purposes, but Marsh cautions strongly that administrators must have sufficient expertise in the evaluation process if the summative purposes lead to personnel decisions (Abrami and d'Apollonia, 1990).

Other Widely Used Forms:

Suggested Readings:
Hobson, S. M. & Talbot, D. M. (2001). Understanding student evaluations: what all faculty should know. *College Teaching* 49(1): 26-31.

Instructor and Course Evaluation System, created by the University of Illinois at Urbana, uses a cafeteria based model that allows instructors to choose from over four hundred globally based questions. The form is computer-based and only recommended for formative use.

Student Instructional Rating Systems, created by Michigan State University in 1982, focuses on specific questions regarding instructor effectiveness, but includes one global question regarding the "general enjoyment of the course" (Hobson & Talbot, p.29).

Instructional Assessment System, created by the Educational Assessment Center at the University of Washington, uses both specific and global questions on several forms that are designed specifically for different class sizes.

Myths about Evaluation

Suggested Readings:
Aleamoni, L. M. (1999) Student rating myths versus research facts from 1924 to 1998. *Journal of Personnel Evaluation in Education, 13*, 153-166.

Bain, K. (2004). *What the Best College Teachers Do*. Cambridge, MA: Harvard Press.

Hobson, S. M. & Talbot, D. M. (2001). Understanding student evaluations: what all faculty should know. *College Teaching* 49(1): 26-31.

Kulik, J. A. (2001). Student ratings: validity, utility, and controversy. *New Directions for Institutional Research* 109, 9-25.

McKeachie, W. and Svinicki, M. (2006). *McKeachie's Teaching Tips*. Boston, MA: Houghton Mifflin Company.

Theall, M. & Franklin, J. (2001). Looking for bias in all the wrong places: a search of truth or a witch hunt in student ratings of instruction? *New Directions for Institutional Research* 109: 45-56.

Many myths exist about the usefulness of student ratings. Some of these myths originate from faulty research studies, conflicting findings within the research literature, or reluctance on the part of administrators and faculty to evaluate and be evaluated, respectively. Some common myths of student evaluations of teaching include: students are not able to make informed and consistent judgments about their instructors; student ratings are essentially a popularity contest; students cannot make accurate judgments unless they have been away

from the course for a while; student ratings are negatively related to student learning; student ratings are based upon the expected grade in course.

While the above myths have been adequately disproved by research, some criticisms of SET's have been long-standing and not resolved. These criticisms center on issues of validity and reliability, and factors that may bias teaching evaluations, including, student, course, and instructor characteristics (Hobson & Talbot, 2001; Aleamoni, 1999; Theall & Franklin, 2001; Kulik, 2001; McKeachie, 2006; Bain, 2004).

Reliability of SET's

Suggested Readings:

Aleamoni, L. M. (1999) Student rating myths versus research facts from 1924 to 1998. *Journal of Personnel Evaluation in Education, 13*, 153-166.

Hobson, S. M. & Talbot, D. M. (2001). Understanding student evaluations: what all faculty should know. *College Teaching* 49(1): 26-31.

Marsh, H. W. & Roche, L.A. (1997). Making students' evaluation of teaching effectiveness effective: the critical issues of validity, bias, and utility. *American Psychologist* 52(11): 1887-1197.

Reliability refers to the consistency of ratings among different raters and also the stability of such ratings over time. Research has shown that student ratings show an acceptable level of consistency, or inter-rater reliability, given a class size of at least 15. The level of consistency among raters increases as class size increases. Longitudinal studies and studies of alumni ratings of an instructor/course have found that ratings show high levels of stability over time. Further, cross-sectional studies show that student ratings reliably reflect instructor versus course effectiveness (Hobson & Talbot, 2001; Aleamoni, 1999; Marsh & Roche, 1997).

Validity of SET's

Suggested Readings:

Aleamoni, L. M. (1999) Student rating myths versus research facts from 1924 to 1998. *Journal of Personnel Evaluation in Education, 13*, 153-166.

Hobson, S. M. & Talbot, D. M. (2001). Understanding student evaluations: what all faculty should know. *College Teaching* 49(1): 26-31.

Greenwald, A. G. & Gillmore, G.M. (1997b). Grading leniency is a removable contaminant of student ratings. *American Psychologist* 52(11): 1209-1217.

Greenwald, A.G. (1997). Validity concerns and usefulness of student ratings of instruction. *American Psychologist* 52(11): 1182-1186

Kulik, J.A. (2001). Student ratings: validity, utility, and controversy. *New Directions for Institutional Research* 109, 9-25.

McKeachie, W. (1997). Student ratings; the validity of use. *American Psychologist* v52, no.11, 1218-1225.

While researchers may argue about the degree of validity of SET's, most find that SET's are valid and can be used as a meaningful source of student input. The validity of ratings is difficult to confirm, however, since there is no agreed upon definition of what constitutes effective teaching. As a result, most researchers have adopted a construct validation or convergent validity approach. This involves examining the correlation between student ratings and other partial measures of teaching effectiveness, such as student learning, observer, peer, and alumni ratings. Studies in this vein have reported generally positive and significant correlations between these measures and student ratings. In addition, research on the validity of ratings has aimed to identify potential biasing factors not directly related to teaching effectiveness (Hobson & Talbot, 2001; Aleamoni, 1999; Kulik, 2001; Greenwald, 1997; Greenwald & Gillmore, 1997b; McKeachie, 1997).

Potential Biasing Factors of SET's

Suggested Reading:

Aleamoni, L. M. (1999) Student rating myths versus research facts from 1924 to 1998. *Journal of Personnel Evaluation in Education, 13*, 153-166.

d'Apollonia, P. A., & Abrami, P. C. (1997). Navigating student ratings of instruction. *American Psychologist, 52*, 1198-1208.

Marsh, H. W. & Roche, L. A. (1997). Making students' evaluation of teaching effectiveness effective: the critical issues of validity, bias, and utility. *American Psychologist* 52(11): 1887-1197.

McKeachie, W. (1997). Student ratings; the validity of use. *American Psychologist* v52, no.11, 1218-1225.

Theall, M. & Franklin, J. (2001). Looking for bias in all the wrong places: a search of truth or a witch hunt in student ratings of instruction? *New Directions for Institutional Research* 109: 45-56.

Researchers and critics of SET's have suggested numerous factors which may bias student ratings of teacher effectiveness including: class size, grade leniency, instructor personality, gender, course workload, time that class meets, and type of class, including the academic discipline and required/elective status of class. For each of these factors, research has been somewhat inconclusive, with some studies asserting a positive, negative, or null relationship between variables. Understanding the potential relationships, however, institutions and researchers have begun controlling for certain student and course characteristics before examining student ratings (Aleamoni, 1999; Theall & Franklin, 2001; Marsh & Roche, 1997; d'Apollonia & Abrami, 1997; McKeachie, 1997).

Grading Leniency

Suggested Readings:

Chambers, B.A. & Schmitt, N. (2002). Inequity in the performance evaluation process: how you rate me affects how I rate you. *Journal of Personnel Evaluation in Education* 16(2): 103-112.

Greenwald, A. G. & Gillmore, G.M. (1997a). No pain, no gain? The importance of measuring course workload in student ratings of instruction. *Journal of Educational Psychology* 89(4): 743-751.

Greenwald, A. G. & Gillmore, G.M. (1997b). Grading leniency is a removable contaminant of student ratings. *American Psychologist* 52(11): 1209-1217.

Marsh, H. W. & Roche, L.A. (1997). Making students' evaluation of teaching effectiveness effective: the critical issues of validity, bias, and utility. *American Psychologist* 52(11): 1887-1197.

McKeachie, W. (1997). Student ratings; the validity of use. *American Psychologist* v52, no.11, 1218-1225.

Many researchers have focused on the positive relationship between grades and student ratings as a potential biasing factor. The danger with this relationship is that it may encourage faculty to inflate their grades for the purpose of garnering higher ratings. Several theories exist that posit why a grades-rating relationship exists. Prominent among them are: students infer course quality and own ability from received grades; students give high ratings in appreciation for lenient grades; and teaching effectiveness influences both grades and ratings. The last of these, also termed the validity hypothesis, finds that effective teachers promote student learning (a measure of effective teaching) which results in higher grades. However, studies examining differences in ratings and grades across instructor and course have been used to support each of the theories posited above (Marsh & Roche, 1997; Greenwald & Gillmore, 1997a; McKeachie, 1997; Chambers & Schmitt, 2002; Greenwald & Gillmore, 1997b).

Global versus Specific Items

Suggested Readings:

Bain, K. (2004). *What the Best College Teachers Do*. Cambridge, MA: Harvard Press.

d'Apollonia, P. A., & Abrami, P. C. (1997). Navigating student ratings of instruction. *American Psychologist, 52*, 1198-1208.

Gallagher, T.G. (2000). Embracing student evaluations of teaching: a case study. *Teaching Sociology* 28, 140-147.

Young, S. Y. & Shaw, D. G. (1999). Profiles of effective college and university teachers. *The Journal of Higher Education* 70(6): 670-686.

Student rating forms generally contain both global (or overall rating) items and specific items, which assess specific aspects of the instructor and course. Research is split on the value of both of these types of items. Some argue that teaching is multi-dimensional and therefore requires specific items to accurately assess different facets of teaching. Others show that when specific items are factor analyzed, they essentially reduce down to one or two items that are global in nature. Studies also reveal that responses on specific and global items are highly correlated.

With regard to the uses of these types of items, researchers warn against making summative decisions based solely on ratings on global items. In addition, formative purposes seem better informed by having data on specific areas that faculty can target in order to improve their teaching (Gallager, 2000; d'Apollonia & Abrami, 1997; Young & Shaw, 1999; Bain, 2004).

The first edition of the Guide to the Evaluation of Instruction *was developed by Professor Morton Friedman (Psychology) and printed in 1983, following the recommendations of the 1980 Systemwide Task Force on Teaching Evaluation. A revised edition by Professor Friedman and Ann Futterman was produced in 1986. Dr. David Unruh expanded and updated the third edition, printed in 1993. The current edition builds on the previous efforts and includes a summary of evaluation literature and research and suggested readings for further information. The Teaching Committee, the Undergraduate Council, and the Graduate Council of the Academic Senate have independently reviewed the text in accordance with the University mandate to evaluate and monitor the quality of instruction.*

Larry L. Loeher
Kumiko Haas
Joanne Valli-Marill
June Chang

2006