GUIDEBOOK

For Evaluating Teaching
PART IV:
Collecting Evaluative Information About Teaching
(Other Sources and Synthesizing Evaluative Information)

Larry A. Braskamp
Dale C. Brandenburg
Eileen Kohen
John C. Ory
Paul W. Mayberry

Abstract

In this concluding section of our four part series, we complete our discussion of specific strategies for collecting evaluative information about teaching by discussing self, alumni, and records as sources. Readers may wish to use Exhibit 3 in Part II to review our perspective. (See **NACTA Journal,** *March 1984, p. 19) Finally, we devote the latter sections to examining how all of the information collected can be synthesized in meaningful ways.*

Section 12: Self As A Source

Self evaluations range from informal self reflections to formal written appraisals to others. Instructors can benefit by systematically analyzing what and how they teach. Since a major goal of a formal evaluation program is to encourage faculty to become "monitors" of their own performance, self-evaluations are critical in getting instructors to reflect upon their own teaching. Self assessment can be both descriptive and judgmental. The following information is recommended for a self evaluation:

* Courses taught and enrollments
* Course materials, syllabus, and assignments.
* Course objective and goals.
* Course outcomes and student learning as measured by the exams and projects.
* Advising responsibilities.
* Involvement in curriculum projects.
* Evaluations by colleagues.
* Special teaching methods and techniques tried as ways to improve teaching competence.

Technical Quality

The generalizations based on self ratings and evaluations are summarized in Exhibit 14.

Examples

Self evaluations can be accomplished in a number of different ways. First, instructors can write their philosophy of teaching strategy, and their own strengths and weaknesses. Self appraisal can often serve as the initial step in an ongoing attempt by instructors to improve. A wirtten self evaluation then becomes a part of an integrated self appraisal which can include a description of the course objectives and goals, course syllabus, assignments, and method of teaching.

Second, instructors can rate themselves on a set of items in a checklist or rating scale. The **ICES Catalog** can be used for item selection. Third, instructors can use Instructor Self Evaluation Form (ISEF) (See 7). This survey consists of four scales: adequacy of classroom procedures, enthusiasm and knowledge for teaching, stimulation of cognitive and affective gains in students, and relations with students. Four items — one item from each scale — are presented in a set of forms. Instructors rank order the four items in terms of the items being "most" to "least" descriptive of their teaching. Because of this forced-choice, instructors are required to indicate their own relative strengths and weaknesses, and they can not specify strengths and weaknesses in some or all of the four areas represented by the scales. Thus this forced type of measurement may not accurately measure the absolute self ratings of strengths and weaknesses; instead relative strengths are to be interpreted.

Suggestions for Use

The value of self evaluations are highly dependent upon purpose. A statement of accomplishments and approach to teaching is ideal for course improvement.

Exhibit 14. Generalizations of the Technical Quality of Self Evaluations

1. Students and instructors generally show good relative agreement on overall ratings of the instructor; i.e., instructors rated highly by students rate themselves higher than instructors rated less highly by students. (2, 3, 5, 8)
2. Instructor self and student ratings on specific dimensions of student involvement, teacher support, and instructional skill are the most congruent. (3, 7)
3. Instructor self ratings are not unduly influenced by the instructor's age, sex, tenure status, teaching load, or years of teaching experience. (6)

Almost without exception, it is the recommended first step in an ongoing process of evaluation and adaptation. However, self judgments of overall ability to be used for personnel decisions are fraught with problems, most notably credibility. Self assessment in the form of description of course materials and philosophy of teaching are often more acceptable than self judgments of value or worth when used for personnel decisions. Thus self evaluations are a critical piece of information for each purpose but the type of information collected and communicated to whom varies.

Using Self Ratings For Personnel Decisions

1. Self ratings of overall teaching competence are not recommended for personnel decisions. Instead descriptions of teaching load, philosophy, and strategy are excellent types of information.

2. Self evaluations of teaching provide contextual information for assessing competence for annual salary reviews. Descriptions of accomplishments and future goals can provide a useful framework for evaluating total instructional performance of a professor.

Using Self Ratings For Improvement

1. Self ratings should be compared with ratings of students if the same items are administered to the students.

2. Faculty can increase the value of their self evaluation by discussing the evaluation with faculty colleagues or staff members trained in faculty evaluation and development.

3. Instructors wishing to focus on specific classroom teaching behavior should consider videotaping a lecture or a discussion and using a rating scale such as one used for colleague review in Appendix E.

Section 13: Alumni As A Source

Alumni and graduating seniors have a unique perspective to evaluate individual faculty, courses in their major field of study, and curricular offerings. Alumni have the additional advantage of being able to judge the relevance of their courses to their present job demands and expectations. Unfortunately evaluations from graduating seniors and particularly alumni are relatively expensive to collect. Thus collecting evaluative information from these sources needs to be done after a comparison between need for the information and costs. Because of the lapse of time between a course and time of evaluation, assessments of highly specific aspects of a course or teaching style are generally not obtained. Instead evaluations about the sequence and depth of course material and support and advice faculty gave to the students during their college career are more valuable information to a department in its examination of its curriculum offerings and the role of its faculty in instruction.

Technical Quality

The research on the validity of alumni ratings is summarized in Exhibit 15.

Exhibit 15. Generalizations About the Technical Quality of Alumni Ratings.

1. Students who have evaluated instructors twice (first during the course and then one year after graduation) show good absolute agreement; i.e., their rating of the competence of the teacher was very similar. (9)

2. Alumni of five years or more and enrolled students show good relative agreement on their ratings of overall teaching effectiveness of instructors. (4)

3. Alumni ratings tend to be lower than ratings of enrolled students. (9)

Suggestions for Use

Collecting information from graduating seniors can be done by exit interviews, telephone, letters, and mail surveys, but only the last two are generally economically feasible for gathering information from alumni. Interviews with graduating seniors can provide considerable indepth information about professors and courses, but this method is expensive. The logistical problems of having students come for an interview and the difficulty in summarizing interview data also need to be considered before the interview as a method is adopted.

Information requested by graduating seniors or alumni needs to be specified in the directions. If information about long-term comprehension and relevance of the content, personal development, technical skills, and motivation to learn is desired, questions about these areas need to be included in the directions.

Alumni Evaluations For Personnel Decisions

1. General items about instructor competence or course organization are preferred over detailed items over specific aspects of a course.

2. General items included in a rating form administered to enrolled students can be included in an alumni form if comparisons are desired.

3. Someone other than the instructor needs to distribute the surveys, and the purpose for collecting the information must be made explicit to the alumni.

4. Alumni should have the option of returning or mailing surveys anonymously.

5. Alumni should be given the opportunity to write comments on a mail survey.

6. Alumni should be asked questions that take advantage of their perspective, such as relevance of course to current position, job demands, and suggestions for topics to be covered or relevance to their role in society.

7. If alumni with varying years of work experience are included in the sample, the evaluation will be more comprehensive.

8. If several faculty are to be evaluated in a single administration, listing names of the instructors reduces biases due to name difficulty and name recall.

9. Questions regarding the department/curriculum can easily be added to an alumni survey.

10. If the response rate is low (i.e., under 50 percent), generalizations must be made cautiously. The available information may be quite biased (i.e., too positive or too negative).

11. Instructors who teach a large number of students are likely to receive many positive and negative comments about their teaching.

Alumni Evaluations For Improvement

1. Since diagnostic information about specific components of a course is difficult to obtain from alumni, instructors are likely to be disappointed if they expect detailed critiques and specific suggestions for improvement.

2. Alumni in their evaluations may point out deficiencies that have been corrected, since they were enrolled in the course.

3. Comparisons of alumni and enrolled student evaluations can be examined for common themes and differences.

Section 14: Records As A Source

Records include grade distributions of students enrolled in courses, number of students enrolled in courses the first week and at end of semester, committee assignments, teaching load, types of course taught such as departmental core courses, generation of instructional units, student credit hours, and advising loads, etc. Some information can be obtained from faculty annual review reports; information such as grade distributions can be obtained from reports issued by an office of institutional research, whereas other information is contained in department records.

Suggestions for Using Records

1. A record of teaching workload over a number of years (e.g., 3-5 years) provides a better portrayal of a professor's efforts and contributions than a record of the latest year. A long-range perspective helps interpret the contributions of a professor with an unusual teaching load.

2. Comparisons of first week and final course enrollment can be helpful in pointing out unusually large decreases in number of students throughout the semester (as compared to others teaching the same course). This information can serve as a flag in interpreting end-of-course student ratings as well as a topic of discussion with the instructor regarding the reasons for dramatic enrollment shifts. Interpretation should be made cautiously, however, since students drop courses for several reasons and some may have little relevance to the instructor or course.

3. Grade distributions can be used to detect unusual grading practices by the instructor. Comparisons with other instructors teaching the same or similar courses can be used in interpreting grading practices. Again caution should be used in inferences about instructor effectiveness on the basis of grades. Instructors who have "honors" classes or grade according to an absolute set of standards may have unique but defensible grade distributions.

Section 15: Synthesizing And Using Evaluative Information

Evaluative information is collected to be used. As stated often in this guidebook, the purpose of the evaluation greatly influences the use that is to be made of the evaluation — how the information is disseminated to whom for what decisions? Once the evaluative information is collected, there are three major activities in which a "user" still needs to engage. The three activities are:

1. Cross checking all the collected information for emergent themes and contradictions.

2. Weighting information for a composite summative evaluation of the professor's instructional competence.

3. Disseminating and communicating the evaluations to the appropriate audiences.

Faculty conducting evaluations for their own purpose of self improvement engage in these activities routinely, since they collect, analyze, summarize, and weigh the information for their own personal use. They themselves are the primary audience. However, evaluations conducted for personnel decisions can involve many different audiences and the final summative evaluation of instructor competence is done by others than the instructors themselves. The individual faculty member does not make the overall assessment nor is the sole audience.

1. Cross Checking Information

If a multiple perspective approach to evaluation is undertaken, the information collected from the various sources by different methods should first be cross-checked for patterns, inconsistencies, and for detection of unique strengths and glaring problems. Often one single theme is not possible; rather a composite portrayal filled with an array of impressions, generalizations, and inferences is more accurate.

2. Weighting Information

Each piece of information needs to be weighed in terms of its importance in determining a summative evaluation of an instructor. This weighting process is one of the most critical phases in an evaluation. Unfortunately few explicit guidelines can be written, since this is heavily based on the professional judgments of those faculty, colleagues, and administrators examining the information. This weighting process, however, need not be secretive and done without a rationale.

One of the best strategies in weighting information is by the use of a set of accepted prescribed weights. This can be done by having the departmental faculty and administrators establish as policy the importance to be given to each criteria and their measured indicators (e.g., student ratings of instruction) in determining faculty competence. Thus weights given to each criterion are known to the faculty before evaluation takes place. Weights given to a faculty

member's contributions in teaching, research, and service can be noted; e.g., research receives a weight of 60 percent, teaching 30 percent, and service 10 percent. The types of evaluative information (i.e., the measures) to be used as indicators of the criteria can be specified so that faculty know the information from which quality is to be assessed. The department head/chairperson and the appropriate faculty committees still must ultimately use professional judgment in their overall summative evaluation, but faculty at least know the bases and the information used in the evaluations.

3. Dissemination and Communication of Information

The last and final step of an evaluation is the dissemination of the evaluative information. If teachers collect information for their own use, there are few problems to counteract. Motivation and interest in learning about teaching and a desire to change should both be high. If information is collected for personnel decisions, the communication can be done in many ways — salary increase figure, letter, and/or personal interviews. Feedback about performance is present in each way, but the salary figures included may be interpreted considerably differently by faculty given different contextual information. Faculty, especially those in their early professional years, often prefer an individual session with the department head/chairperson since they can then discuss their future and gain further understanding of the expectations of the department and institution. It also makes the evaluations more personal rather than mechanical or objective.

In accumulating and summarizing information for assessing teaching quality, there are three different syntheses that are often done. They include the synthesis of information about:

1. instructional competence for a single course,
2. instructional competence of a faculty member for the promotion and tenure process, and
3. instructional competence for annual faculty salary adjustments.

Section 16: Synthesizing Single Course Evaluations

The usefulness in collecting information from various sources about a single course is the cross-checking of the evidence for patterns and dissimilarities. Integrating information from many perspectives also prevents undue importance to any one piece of information.

If the purpose of the evaluation is for improvement, a one page summary of all the information available for a course can be constructed. Information may be categorized around source (self, students, alumni, and peers) and major components of the course can be used to organize the synthesis of the available data as well as detect consistencies and conflicts. Instructors who complete this form for their own use should encounter few problems. If a form like this is used for personnel decisions, the credibility of this summarization can be increased easily by conducting a systematic analysis of the comments and by asking a colleague to verify the written summaries.

Section 17: Synthesizing Evaluations For Promotion and Tenure Review

Evaluation of instruction is required in the promotion/tenure process at most colleges and universities. Since many academic affairs offices issue guidelines each year, the requirements regarding documentation of instructional effectiveness should be reviewed annually.

Evaluative information about instruction is most useful if the following major guidelines are met.

1. Information collected from various sources about a number of courses increases the comprehensiveness and fairness of the evaluation. Evaluations from only students or from only a few or the most recent courses taught by an instructor present a biased view.

2. A report of students ratings should include the following information: title and name of course, number of students who returned survey, mean, and standard deviation of an item, and an index of relative standing in a defined comparison group. A composite computer-generated profile of evaluations of all courses over a several-year period taught by a professor is desirable.

3. Student responses to diagnostic items (such as from the ICES catalog) are seldom appropriate for inclusion in the documentation submitted for review by others.

4. If student written comments are used, it is preferred that they be randomly selected. Themes and highlights presented without actual comments have low credibility, unless they are written or verified by a colleague.

5. Colleague evaluations, if collected within a consultative arrangement for course improvement, are seldom justified for use in personnel decision making.

6. Colleague and alumni information collected for personnel decisions may be shared with the instructor but disclosure to the instructor should not be required.

7. Judgments in a number of different components of instruction enhance the veracity of the evaluation. An example of a college that takes into account a number of components of instruction and "sources of judgments" which are equivalent to the various ways of collecting information discussed in this guidebook can be obtained from the authors.

Section 18: Synthesizing Evaluations For Annual Salary Adjustments

Faculty members often complete an annual report describing their accomplishments (usually in the areas of teaching, research, and service). Other information

such as peer evaluations of instruction may be available to the administrators and committees respc,nsible for making annual summative evaluations of the faculty.

Department heads/chairpersons and their respective advisory/executive committees use a variety of methods to obtain a composite overall evaluation of each faculty member. The following fʒuidelines are given as guidance for assembling 'ɜvaluative information about instructional performarice.

1. The importance (weights) given to accomplishments in teaching, resear⌐ch, and service facilitate the process of determining a composite overall evaluation of a faculty membⱱr.

2. The weights may include a raɲge; e.g., teaching performance is 20-40 percent of the typical faculty member's contribution. The weigʰting scheme should be sufficiently adjustable for any one faculty member to take into account unusual c⁻ircumstances or contributions.

3. Suggestions listed unⱱder "Evaluations for Personnel Decisions" for each source are appropriate here. (See the relevant section.s in Parts II, III, and IV.)

4. A form to summarize the judgments based on the evaluative information can facilitate the review process. An example of su⌐ch a form can be obtained from the authors.

Conclusi.on: Part IV

Throughout the Guɭidebook, we have attempted to follow an applicati·on-oriented perspective in presenting our views orɪ evaluating teaching. In a sense, we have strived to desⱼcribe the "state-of-the-art" as we apply these concepts , principles, and strategies to our professional work with faculty and administrators. Evaluating teachirɪg is a difficult task and improvements in our processes will continue to be made. Although we prof.ess a comprehensive perspective, we welcome additio·ns and modifications from our readers.

P·art IV - Bibliography

1. Batista., E.E. and Brandenburg, D.C. 1978. The instructor sⱼelf-evaluation form: development and validation o'f an ipsative forced-choice measure of self-perceived faculty performance. **Research in Higher Educatior,ɪ, 9,** 319-332.

2. Bɭackburn, R.T. and Clark, M.J. 1975. An assessm⌐ent of faculty performance: Some correlates betweeɼ:n administrators, colleagues, student, and self-ratinp,s. **Sociology of Education, 48,** 242-256.

⌐3. Braskamp, L.A., Caulley, D.N., and Costin, F. 197⌐Cɪ. Student ratings and instructor self-ratings and their relationship to student achievement. **American Eɾlucational Research Journal, 16,** 295-306.

4. Centra, J.A. 1974. The relationship between 'student and alumni ratings of teachers. **Educational and Psychology Measurement, 34,** 321-326.

5. Doyle, K.O., Jr., and Crichton, L.A. 1978. Student, peer, and self evaluation of college in-

struction. **Journal of Educational Psychology, 70,** 851-826.

Doyle, K.O. and Webber, P.C. 1978. Self ratings of college instruction. **American Educational Research Journal, 15,** 467-476.

7. Marsh, H.W. 1982. Validity of students' evaluations of college teaching: A multitrait-multi-method analysis. **Journal of Educational Psychology, 74,** 264-279.

8. Marsh, H.W., Overall, J.U., and Kesler, S.P. 1979. Validity of student evaluations of instructional effectiveness: A comparison of faculty self evaluations and evaluations by their students. **Journal of Educational Psychology, 71,** 149-160.

9. Overall, J.U. and Marsh, H.W. 1979. Midterm feedback from students: Its relationship to instructional improvement and students' cognitive and affective outcomes. **Journal of Educational Psychology, 71,** 856-865.

BOOK REVIEWS

Donald M. Elkins
Book Review Editor
Southern Illinois Univ.
Carbondale, IL 62901

The NACTA Journal Book Review policy encourages the academic freedom of peers in the constructive criticism of unsolicited books submitted by publishers for review. The peer reviewers are persons who teach and/or conduct research in the subject matter area in which the book is written. A given review reflects the opinion of only the reviewer, and such opinion is not necessarily the opinion of NACTA and/or NACTA Journal.

Donald N. Riemer. **Introduction to Freshwater Vegetation.** AVI Publishing Co., 1984. 207 pp. Hardbound. $35.00

Introduction to Freshwater Vegetation is intended for use in an introductory course for students of ecology, botany, wildlife, and fisheries management and other related fields. It is also considered by the author to be useful in non-degree training courses such as those required for certification in the area of pesticide application.

The text is presented in three parts, with several chapters included in each part. Each chapter contains information on a given area of aquatic vegetation and is documented with bibliographic citations.

Part I is an introduction to aquatic environments and includes five chapters covering the diversity of aquatic environments, factors affecting plant life in standing and flowing waters, light in aquatic environments, and plant nutrients. The chapter on diversity of aquatic environments includes a brief discussion of the forthcoming subjects of standing and flowing waters and factors affecting the aquatic environment including geology, topography, local climate, and human activities.

The chapters on standing and flowing waters include sections on succession, classification, and productivity; lake stratification is covered in the chapter on standing waters. An entire chapter is devoted to light in the aquatic environment which includes topics such as surface loss and qualitative and quantitative light transmission by water. Light quality and intensity in relation to submersed