

# What Function Can Colleagues Have In the Evaluation of Instruction?

Larry A. Braskamp

*According to newspaper accounts, national surveys, and personal experience, faculty are being evaluated more and more. Colleague appraisals and judgments about a professor's quality of research have a long and accepted tradition, but the evaluation of instruction has been based primarily on student assessments. Colleague evaluation of teaching has often been second hand, with much of it based on student comments to advisors. While students should have considerable input into the evaluation of instruction, they should not be the sole source of evaluative information (Brandenburg, Braskamp, and Ory, 1979).*

*This paper discusses the use of colleagues in the evaluation of instruction. Major issues involved in colleague evaluation of instruction are first discussed. Then issues and suggested strategies for colleague evaluation of in-class instruction and colleague evaluation of instructional materials and instructor involvement in instruction are presented.*

There is not yet one best colleague evaluation system. Evaluation is inevitably subjective, and peer evaluation incorporates to a large extent professional judgment. Peer evaluation cannot or should not eliminate another faculty member's preferences for and definitions of effective teaching. However, if colleague judgments can be collected systematically, the opportunity for cross-checking this type of evaluative information with evidence gathered from other sources (e.g., instructor self ratings, student ratings, student achievement) exists. Only through an iterative process of confirming and integrating the evaluative information collected from a variety of perspectives can a summary judgment of the faculty member's teaching performance be scientifically adequate, credible, and thus useful.

If colleagues are to evaluate their peers, the following questions are worth considering. The questions pertain to evaluations based on appraisals of written materials and instruction, classroom behavior, and instructor involvement in instructional developmental activities.

**1. What is the purpose of the evaluation?** Peer evaluation used for course improvement purposes should be clearly differentiated from colleague evaluation for promotion and tenure decisions. These differences should be reflected in the formality of the procedures used, the need for written documentation, and a set of guidelines regarding the selection of colleagues, observa-

tion practices, and reporting procedures.

In designing a colleague evaluation system for promotion purposes, each departmental faculty should develop its own system of colleague evaluation because the nature of instruction and the criteria to be considered for evaluating faculty performance may differ from department to department. For example, a department of agricultural communications may have different types of courses than a department in animal science. If faculty are involved in its development, the peer evaluation system will be more easily viewed as credible and useful. Without faculty acceptance the system will probably do more harm than good and could disrupt current collegial relationships.

**2. What can be evaluated?** The following are some areas that can be judged by colleagues who have the necessary expertise in the discipline of the faculty member being evaluated.

- a. Instructor's knowledge and expertise in major field as reflected by the course syllabus and the reading list
- b. Instructor selection of realistic course objectives
- c. Instructor assignments, group projects, and examinations
- d. Student achievement as indicated by performance on exams and projects
- e. Contributions to instructional efforts in the department
- f. Thesis supervision
- g. Involvement in instructional research
- h. Student-instructor relations within the classroom
- i. Instructor's style as a scholar and as a model teacher

**3. What standards and evidence for excellence should be used?** Development of an a priori list of detailed standards against which the evidence of faculty performance can be judged is probably of limited value. It is probably more prudent to list a number of areas to examine and then to allow the colleagues to use their own standards to judge the quality of the instruction and the organization and materials used in a course.

**4. Who should do the evaluations?** Practical and political considerations have to be weighed as well as the purpose of the evaluation. If the purpose is for course improvement, an informal relationship involving a pair of faculty members who trust and respect each other is most efficacious. Faculty members with considerable teaching experience and teaching competence were regarded as the best consultants for instructional improvement purposes by faculty development directors (Centra, 1977). However, faculty members are often reluctant to judge

Adapted from "Colleague Evaluation of Instruction" published in *Faculty Development and Evaluation in Higher Education*, 1978, 4, 1-9. Used by permission. Braskamp is Head, Measurement and Reach Division, Office of Instructional Resources, University of Illinois, Urbana, IL 61801.

their peers and offer assistance if not asked. Faculty members have a long history of respecting another's freedom, independence, and privacy in the instructional area.

Professionals with expertise in faculty and instructional development can also be valuable consultants. They can use their knowledge of instructional strategies to suggest alternative ways of teaching. Since they are not members of a department, they have an independence, which some faculty members consider less threatening. Confidentiality is also easier in this arrangement.

Senior faculty members with expertise both in content and in teaching are the most appropriate evaluators if the evaluations are to be used for personnel decisions. A number of methods have been used to select reviewers (Hoyt, 1977). The same faculty members could be selected to review instructional efforts of all departmental faculty seeking promotion and tenure. This plan maximizes consistency and comparability of ratings, but a different set of faculty may be needed to evaluate teaching, service, and research to distribute the workload.

Another strategy is a random selection of faculty members to serve as evaluators. However, if the department is large, the colleague may have little basis for judging the value of another's credentials as a scholar/teacher. A third is to allow the faculty members evaluated to select five or six colleagues they consider qualified and fair and then have a departmental executive committee select from the list.

**5. What type of judgment is desired?** The type of judgment is related to the specificity of the characteristics and behaviors of the instructor that are to be evaluated. Judgments can range from a global written assessment or selection of the "best" teacher to a detailed analysis based on checklists of specific teacher characteristics, skills, and student outcomes. The judgments desired should depend upon the system of colleague evaluation adopted by the department, especially if it is to be used for personnel decisions. The type of judgment selected depends on the importance one places on professional judgment and inter-judge agreement. Based upon his experience, Hoyt (1977) concluded that colleagues agree more often when detailed instructions are provided. Generally, the greater amount of inferences required by the evaluators, the greater the disagreement. But if human testimony, thoughtfully written and reasoned, is viewed as credible evidence in the evaluation process, then variability of judgments may be as important information as total agreement on behavioral and specific indicators of one's instructional efforts. Generally, summative global integrative judgments are required for personnel decisions, whereas diagnostic and evaluative information is more useful to faculty trying to assess their teaching behaviors and to test new approaches.

**6. What effect will the peer rating system have on the department?** Any system of evaluation will probably have some effect on the professional relationships and

collegiality among the faculty and the faculty chairman relationship. Although faculty members may consider a formal system fairer than an informal one, they still may be reluctant to engage in an elaborate formal system. For this reason a formal evaluation system may not work well. A pervasive judgmental climate can be counterproductive. Academic freedom and the ability to work in an environment which allows faculty members considerable independence is not only important but essential. If faculty members feel they are spending too much time evaluating others or feel that they are continually being evaluated, they may regard this as an intrusion on both their self image as professionals and their time to carry out their professional responsibilities. A "paralysis of analysis" can creep in which may eventually affect productivity and discourage an enthusiasm for experimentation. In designing an evaluation system, the professional integrity of the faculty must be given serious consideration; otherwise the unique role of the faculty as innovators and scholars may be threatened.

These six questions are intended to be guiding questions in the development of a formal colleague evaluation system. There are other considerations such as cost due to faculty time to implement the system, the value of peer evaluation as an aid to faculty self development, and the quality of the information obtained through a peer appraisal (Brandenburg, Braskamp, and Ory, in press).

### **Classroom Observations**

Classroom observations are intended for learning something about the teaching process and its relationship to student learning. The focus is on what can be observed in the classroom, the verbal and nonverbal behaviors of both the instructor and the students. The effects of instruction such as student learning are not studied per se.

Classroom observations are particularly useful in a program of self evaluation and improvement. Faculty members who desire to analyze their behavior in the classroom can find a record of their behaviors by an outside observer to be useful. If the purpose is to collect evidence for promotion and salary decisions, colleague evaluation as a method would need to be formalized to maximize fairness, reliability, and credibility. Since faculty members disagree on its value and appropriateness as an evaluative technique, considerable faculty discussion and input should precede any adoption.

Before classroom observation is used as an evaluative technique, some issues regarding its reliability and usefulness are worth noting. Following are suggestions for alleviating some of the problems in colleague evaluation based on classroom observation and maximizing its contribution.

### **Issues**

**1. What is the effect of having an observer present in the classroom?** Does the instructor perform better or differently and do the students respond differently? It will depend on the physical setting of the class, its size, and

the frequency of observations. The first time, most instructors are anxious and uncomfortable. The threat of the intrusion of an outsider is partially alleviated if a trusting relationship has been established. Many faculty members have difficulty allowing others to view their teaching. Teaching is regarded as a private matter, involving only the students and the instructor.

**2. How reliable are colleague ratings based on classroom observations?** How trustworthy is one colleague's appraisal of another? Will colleagues agree on one's instructional effectiveness from classroom observations? Can a colleague generalize about another's effectiveness and competence after observing one or two class sessions? The accuracy of generalizations is based on the representativeness of the sample of behavior observed. One classroom visit is rarely adequate. The trustworthiness of the observations increases as the visits, if done randomly or systematically, increase. Accuracy and agreement in observations is related to the extent that the observer needs to make qualitative judgments in classifying the behavior observed. If a well developed behavior oriented classification scheme is used, reliability of judges will likely be higher. However, a record of specific units of behavior may not be very beneficial if assessments of appropriateness of material covered, instructor style, etc., are needed. Centra (1975) concluded that colleague judgments of instructor skill and general teaching effectiveness based on two classroom visits was not sufficiently reliable to warrant their use for promotion purposes.

**3. How valid are ratings based on classroom observations?** The relationship between what an instructor does in the classroom as observed and student learning has not been very strong. Thus we do not know which instructional behaviors are indicative of student learning. From classroom observations, a colleague cannot definitively determine if the instructor demonstrates the most effective instructional behaviors. However, the colleague can use the observations to judge the instructor's ethical behavior, scholarship, and teaching approach.

**4. Are colleague ratings related to student ratings?** In one study (Centra, 1975), peer ratings based on two classroom visits by faculty members within and without the department were not highly related to student ratings on the instructor's effectiveness in the course, whether *class time* was well spent and whether the instructor was open to other viewpoints, but colleagues and students were reasonably agreed on items relating to specific instructional practices. Colleagues were even more generous in their ratings than students with 94 percent of the teachers rated excellent or good by their peers.

**5. Of what value is videotaping?** Instructional and professional development directors or coordinators at 756 colleges and universities consider analysis of in-class videotapes one of the most effective practices although only a few on any one campus are using it (Centra, 1977).

The same problems of validity and reliability present in classroom observation emerge in analyzing video tapes, but there is one important distinction. Instructors have an opportunity to view themselves; self confrontations can not be avoided in video taping. Feedback which accompanies the playback can be given by some colleague and can range from straightforward questions to opinion-giving.

The use of video taping has received considerable attention, and not everyone agrees on its merit. Fuller and Manning (1973) argue that this method can benefit but also harm an instructor. The initial reaction of teachers in pre-service programs is one of stress, disturbance, and anxiety. Self esteem usually will decrease or not increase. Instructors who are physically attractive and self assured receive some benefit from viewing themselves, but others are not prepared to be confronted with how they may appear to students. Since the medium focuses on the physical attributes of the instructor, more than one videotaping is needed before the instructors can concentrate on the content of the lecture and the relationship they are establishing with students. Initially, teachers tend to underrate themselves but improve over time, once the "confrontation" becomes a little more natural. Despite the intensity and realism of the confrontation, its effects on teaching performance or competence are complex. Effects depend on the personality characteristics and teaching competence of the teacher, the nature of the accompanying feedback, and the playback itself.

Colleague appraisal based on classroom observation or video tape seems especially useful in a continuous program of evaluation for course improvement purposes. The issues of confidentiality, authenticity of the behavior observed, obtrusiveness of the observers, accuracy and objectivity in recording the behaviors rated, can more easily be dealt with if the instructor has the opportunity to respond and discuss the ratings. Video taping may have the most impact because of its confrontation value.

There are some judgments of instructional effectiveness that colleagues can make only by observing. One is a judgment of the ability of the instructors to present up-to-date and accurate content in their discipline to their students in a classroom setting. Of course, the confidence of this judgment is enhanced by the observer's knowledge of the subject matter area. Another is an assessment of teachers as scholars in their field. Do faculty members demonstrate an approach that incorporates research into teaching? How do they solve a problem, discuss an issue, integrate conflicting points of view, respect those with differing viewpoints? In sum, what kind of model are they to students, prospective teachers, and researchers?

### Suggestions

Colleague appraisals based on classroom observations should be conducted systematically; faculty members are more apt to accept them as a credible source of evaluative information. The following guidelines can be used for that purpose, but they are only suggestive.

1. **Select observers from similar academic disciplines.** Colleagues in a discipline are the most appropriate persons to assess another's ability to present his/her scholarship to students. Judgments regarding appropriate level of difficulty of material presented, relevance of examples, integrations of topics, structure of the lecture, and congruence between what was intended and accomplished by the classroom activity are easier to make if the colleague has knowledge of the subject matter presented and some experience in teaching. If the purpose is for course improvement, a consultant from an office of teaching resources can be valuable in describing the classroom interactions, pointing out strengths and weaknesses in instructional strategies, and discussing with the instructors their approaches to teaching.

2. **Assign observers and instructors who respect each other.** If colleague evaluation is for course improvement and is voluntary, this usually is not a problem. A voluntary arrangement of reciprocity pre-supposes this trust. For personnel decision making, it may be advisable to avoid having a faculty member with a strong personal dislike or lack of respect for a colleague serve as the evaluator. If faculty members disagree considerably on a definition of effective instruction, the colleague may reflect only his philosophy and biases, which may reduce the validity of the judgments.

3. **Have more than one colleague rater and have as many visits as possible.** By having more than one judge, both the faculty member and the department have evaluation information that summarizes the judgments of faculty members with different perspectives. Since any observation must be interpreted, the raters (colleagues) will quite naturally rely on their own experiences, values, and definitions of effective teaching in making judgments. Cross-checking of interpretations and judgments is a crucial strategy in the establishment of reliable and credible information. In instructor evaluation, agreement may not be even possible because of the observer's varying perspectives and thus multiple judgments are critical if the evaluative information is to be used for personnel decisions.

4. **Have colleague(s) and instructor meet together for a session before the observation(s).** In this meeting the colleague(s) can receive copies of the course materials, learn the overall goals of the course and the intent of the class period(s) to be observed, discuss a method of observation (checklist rating form, behavioral observational schedule), and arrange for post observation meetings if appropriate. If the purpose is for improvement, the instructor may suggest concerns and course dimensions on which he/she would like some feedback.

5. **Determine which aspects or dimensions of the course are to be observed.** In class observations, an observer cannot simultaneously record every transaction or behavior, but instead must focus on specific areas. General areas to observe are listed below. A set of rating scales that can be used by an observer are presented in

the booklet, **Improving Your Lecturing**, published by the Office of Instructional Resources, University of Illinois, Urbana. IL 61801. (Diamond, Sharp and Ory, 1978).

- A. **Content Suitability**  
 Relationship to course syllabi, assigned readings.  
 Content is worth knowing  
 Content represents current thinking in discipline  
 Presentation of divergent viewpoints  
 Level of difficulty of material
- B. **Organization of the Content**  
 Logical sequence of topics  
 Pace of the lecture, discussion topics  
 Provision of summaries and syntheses  
 Appropriate use of class time
- C. **Instructor's Clarity of Presentation**  
 Definition of new terms, concepts, principles  
 Relevance of examples  
 Relationship to lab and discussion group assignments
- D. **Instructor's Questioning Ability**  
 Asks variety of types of questions (rhetorical, open-ended, short answer)  
 Engages class members in discussion  
 Allows and encourages students to respond to each other  
 Directs discussion that is centered on the intended topic
- E. **Instructor's Style**  
 Stimulates student thinking  
 Engages student in problem solving activities  
 Appropriate modeling behavior  
 Professional and ethical behavior
- F. **Instructor-student Interactions**  
 Reinforces and encourages student contributions  
 Creates stiff or relaxed atmosphere  
 Demonstrates mutual respect  
 Personal mannerisms and teaching style (voice, vocabulary)

6. **Decide on the method of recording the classroom observations.** There are three major types of "instruments" that can be used. First, the colleague may observe and make clinical interpretations and judgments based on observation. This open-ended approach gives the colleague maximum freedom in deciding on what to observe, how to construct a picture of what is viewed, and how to interpret the observations. Second, rating scales can be used. An observer has a listing of behaviors and attributes on which to rate the instructor. This approach is a common one, with the items often being a set of items that students complete. Since these items are often general in nature, considerable judgment and inference are required. (Items referring to the dimensions listed above are typical examples of this type.)

The last approach incorporates the use of systematic observation schedules whereby predetermined behaviors, concretely and specifically defined and directly observable, are recorded in a systematic way on a schedule. The observer records the unit of behavior observed at every interval, usually only a few seconds. For example, one could record the number of student questions in a class. If the observational schedule is highly behavioral and specific, the observer will merely need to record frequencies of the behaviors occurring which generally increases reliability. However, a listing of specific behaviors may not be very useful for assessing instructor competency, especially in areas related to scholarship.

**7. Schedule a post observation meeting.** A meeting following the observation is especially valuable if the purpose is for course improvement. If the classroom session was videotaped, a meeting with a colleague can be used to discuss the rationale for the evaluation and to point out strengths and areas for improvement. If no video tape is available, then a meeting to discuss a written report should be arranged as soon as possible to take advantage of the recency of the experience. If the written report is used for personnel decisions, the faculty member may be given an opportunity to respond to the appraisals. The instructor's response can become a part of the self evaluation report that the instructor completes for personnel decisions.

#### **Peer Evaluation of Materials**

Colleague evaluation based on the instructor's syllabi, assignments, testing and grading practices, and advising, as well as involvement in the improvement of instruction through research and development can provide an important source of evaluative information. Evaluation in these areas is often a part of the self evaluation report submitted as part of a faculty member's defense for promotion, but the information is usually not course specific.

The evaluation of out-of-class instruction such as instructional and curricular development, academic and vocational or professional advising and instructional research has been heavily judgmental, often based on little empirical evidence. The following checklist is intended to suggest items that may be evaluated, but its use does not preclude the necessity to make judgments. Rather its use may make the assessment more comprehensive and standardized if more than one colleague is evaluating.

#### **Course Organization**

1. The syllabus adequately outlines the sequence of topics to be covered.
2. The stated course objectives are clear.
3. *The outline* and sequence of topics is logical.
4. The course duplicates or is not an adequate prerequisite for other courses.
5. The difficulty level is too high or low for the enrolled students.

#### **Readings, Projects, and Laboratory Assignments**

1. The reading list (required/recommended) is up to date and represents the work of recognized authorities.

2. Readings are appropriate for level of course.
3. The texts used in the course are well selected.
4. The assignments require busy work.
5. Students are given ample time to complete the assignments/take home exams.
6. Readings are well selected.
7. The amount of reading and homework is appropriate.
8. The written assignments and projects are carefully chosen and reflect course goals.
9. A variety of assignments is available to meet individual student needs.
10. Laboratory work is integrated into the course.

#### **Exams and Grading**

1. The exam content is representative of the content and course objectives.
2. The exams are graded in a fair and consistent manner.
3. The grade distribution is appropriate for level of course.
4. The standards used for grading are communicated to the students.

#### **Concern and Interest in Instruction**

The Instructor:

1. Seeks advice from and discusses with colleagues how to be an effective instructor.
2. Is interested in how other colleagues teach.
3. Encourages cooperative teaching arrangements.
4. Is sought by colleagues for advice on instruction.
5. Is knowledgeable about current developments in teaching in his field.

#### **Participation in University Community**

The Instructor:

1. Is involved in student organized and sponsored activities.
2. Participates and attends activities in which students are involved.
3. Participates in departmental seminars, activities, projects involving students.

#### **Vocational and Personal Advising**

The Instructor:

1. Takes an interest in and advises students in their future vocational and professional careers.
2. Helps students in their selection of courses.
3. Meets with students informally out of class.
4. Helps students obtain job related experiences that are beneficial to their professional careers.

#### **Academic and Thesis Advising**

The Instructor:

1. Takes committee membership seriously.
2. Is dependable.
3. Is constructively critical and supportive of students' progress.
4. Provides opportunities for students to conduct publishable research.
5. Is accessible to students.

## Conclusion

If colleague evaluation is used, peer appraisals of course materials and exams should be at least a part of the process. Peer evaluation of course materials is less threatening than classroom visitation; course organization and reading lists are also easier to change than are personality traits of an instructor. Evaluation of course materials also requires less time. In addition course materials may be reviewed by faculty in other universities. It is a common practice to have outside reviews of faculty for promotion. Course syllabus and exams could easily be a part of the materials reviewed.

Colleague visitation of classrooms should be used judiciously; they seem most appropriate for faculty development purposes. Faculty members can play an important part in the evaluation of the instruction of their peers. However, the evaluation must be carefully and judiciously planned and implemented. Otherwise, it may result in a working environment that would decrease

rather than promote the pursuit of excellence, faculty autonomy, and faculty diversity, a goal that has been the landmark of American higher education.

## References

- Brandenburg, Dale C., Larry A. Braskamp and John C. Ory. 1979. Considerations for an evaluation program of instructional quality. *CEDR Quarterly*, 12, 8-12.
- Centra, John A. 1975. Colleagues as raters of classroom instruction. *Journal of Higher Education*, 46, 327-337.
- Centra, John A. 1977. The how and why of evaluating teaching. In J.A. Centra (ed), *Reviewing and Evaluating Teaching*. San Francisco: Jossey-Bass.
- Diamond, Nancy, Greg Sharp and John C. Ory. 1978. *Improving Your Lecturing*. Urbana, IL: University of Illinois, Office of Instructional Resources.
- Fuller, Frances F. and Brad A. Manning. 1973. Self-confrontation review: A conceptualization for video playback in teacher education. *Review of Educational Research*, 43, 469-528.
- Hoyt, Donald P. 1977. *Background on the Uses and Misuses of Peer Evaluation — Perspectives from Another Campus*. Iowa City, Iowa: Unpublished manuscript.

# Professional Faculty Development Program At An Agricultural Technical Institute

Arnold Mokma and  
Roger Baur

"Faculty Development" activities have traditionally concentrated on remaining current in one's subject matter area. This "development" was, and continues to be, delivered through such programs as conferences, workshops, seminars, advanced degrees, and sabbatical leaves. These programs were, and in many cases are, uncoordinated and not available to all faculty.

However, because scholarly competence in a technical area does not necessarily translate into teaching effectiveness, the emphasis of faculty development changed at the Agricultural Technical Institute, The Ohio State University, to include the improvement of instruction. The basic assumption is that the teaching process is made up of knowledge, skills, attitudes, and motivations which can be learned. Under this concept, it is important to take a holistic approach to the teaching-learning process.

It becomes essential to emphasize the individual and the interrelationships of faculty, administration, staff, and students — the institutional environment — if we are to improve the effectiveness of teaching.

The current concept of faculty development — more appropriately labeled professional development — places an emphasis on the areas of technical competence, im-

provement of instruction, and development of skills and sensitivities needed to work with students and colleagues in an ever-changing society.

## Our Program

The Teaching Committee in 1977 surveyed faculty regarding a "faculty resource and support program." The results of this survey (Table 1) identified the following items as high priority goals/outcomes:

1. To develop greater competency in the area of instructional skills and techniques in the classroom
2. To increase ability to use appropriate, varying teaching strategies
3. To increase knowledge about ATI's operational procedures, personnel policies, facilities, etc.
4. To enhance the personal confidence and self-worth of each staff member
5. To develop greater competency in subject areas
6. To increase understanding of the characteristics and needs of our students
7. To increase ability to use human relations skills in order to communicate more effectively in your area

These goals/outcomes reflect the professional development concept: technical competence, teaching methodology, the institutional environment, and interpersonal/human relations skills.

The key outcome of this survey was the hiring of a "Faculty Development Coordinator" who works with individuals and groups of faculty members on their

Mokma, recently named Assistant Director for Academic Affairs, was Faculty Development Coordinator at the Agricultural Technical Institute and is Assistant Professor in Agricultural Education, The Ohio State University. Baur is Assistant Professor at The Agricultural Technical Institute.