Extending the Domain of Instructional Effectiveness Assessment in Student Evaluations of Communication Courses

Ronald E. Rice, Lea P. Stewart, and Michele Hujber

Student evaluations of college-level courses are valuable tools for assessing and improving classroom teaching. However, they do not systematically capture information about the longer-term assessment of instructor and course effectiveness that could be provided by other stakeholders, such as alumni. To achieve this goal, an alumni teaching effectiveness survey was developed based on previous research findings, phone interviews, and a pilot survey, and then completed by one-year and 10-year graduates of the Department of Communication at a large public university. Results indicate that alumni believe that good instructors should demonstrate effective preparation, interest in course material and students, availability to students outside of class, effective pedagogy, practical application (e.g., "real-world" connections), classroom interaction with students, sociability, and a high level of scholarship. In general, according to alumni, being an effective communicator is an essential component of being an effective instructor. **Keywords:** alumni perceptions, classroom communication, course evaluations, educational accountability, instructional effectiveness

Numerical student evaluations of instruction are used in nearly 80% of postsecondary institutions (Trout, 1997). These reporting systems capture undergraduate students' perceptions of courses and their instructors based on the classroom experience. There is, by now, an extensive body of research on student evaluations and their relationships to learning, achievement, and instructor attributes (e.g., Scott & Nussbaum, 1981; Smart, 1997). Student evaluations of instruction can help a teacher target areas for self-improvement; they can be used for administrative decisions (but they must be used judiciously); and they can be used as information for students in their selection of courses. When done effectively, the process of creating, administering, evaluating, and using instruments for evaluation of teaching opens up within an institution a dialogue that can clarify the mission and goals of the institution.

As useful as undergraduate evaluations are, however, they do not systematically capture information relative to the longer-term assessment of instructor and course effectiveness that could be provided, for example, by gathering retrospective perceptions from alumni. As a recent editorial in *Change: The Magazine of Higher*

Ronald E. Rice and Lea P. Stewart are Professors in the School of Communication, Information and Library Studies, Rutgers University, New Brunswick, New Jersey 08901-1071. *Michele Hujber* is a Publications Coordinator in the Office of Communications and Public Affairs, Cook College, New Brunswick, NJ 08901.

This project was supported by a grant from the Rutgers Office of the Vice President for Undergraduate Education. We appreciate the participation by the survey respondents, the assistance and comments from Dr. Linda Lederman, Professor, Department of Communication and Dr. Brent Ruben, Professor and Director, Rutgers University Office of Quality and Communication Improvement, and Dr. Ed. Martin's MCIS course for participating in the pilot study. The web site www.scils.rutgers.edu/~rrice/teachcon.htm provides greater details on all aspects of this study.

Communication Education, Vol. 49, No. 3, July 2000, pp. 253–266

Learning (Marchese, 1997) notes, "we see the simple 'objectivity' of the numbers drive down–or out–richer forms of student and peer feedback for the evaluation and improvement of teaching" (p. 4). Plater (1997), responding to a study in the same magazine showing that simply improving one's level of enthusiasm in presentation (voice pitch variability and use of gestures), a content-free change in course delivery, can substantially raise students' evaluation of all items concerning both instructor and course, but without actually influencing any change in learned content, concludes that:

> . . . student evaluations can only be one part of a holistic assessment of teaching and learning. Clearly, students have a necessary role within this larger process, but surely not the primary one they have assumed . . . students, peers, administrators, and–most critically–the teachers themselves (in the form of reflective self-assessment) can complete the full picture of teaching effectiveness and student learning. (p. 17)

While student evaluations of instruction are a generally accepted evaluation tool at institutions of higher education, they do not fulfill all of an institution's needs for constituent evaluation. With increasing emphasis on accountability from government and society, some researchers are becoming concerned about whether or not student evaluations of instruction adequately reflect the quality concerns of all of their stakeholders. Changing notions of educational accountability, along with concepts from Quality Management, have extended our notion of who are the potential stakeholders in both obtaining and evaluating services from academic institutions (Fram & Camp, 1995; Ruben, 1997). While current students are and should remain the most important stakeholders to institutions of higher education, other important constituencies include alumni, faculty, review committees, university administrators, parents, state legislators, and those who employ university graduates (Glassick, Huber & Maeroff, 1997). Undergraduate alumni of an institution are unique among external stakeholders in that they have been participant observers in the classroom and have applied the results of that experience in a non-undergraduate-academic environment. Of all stakeholders, they are the most likely to be able to identify additional teaching dimensions upon which instructors could be evaluated (Pettit, 1991). Developing systematic assessment methods for alumni, and including their concerns on instructor evaluation forms, would add the perspectives of these important stakeholders in assessing instructional excellence.

The familiar student evaluation form and ratings of instructor effectiveness can be considered just one aspect of a wider teaching effectiveness domain. A more comprehensive matrix would include multiple evaluation criteria along one dimension and the various significant stakeholders along the other dimension. It would be valuable to assess the intersections between these domains and stakeholders. For example, we may find that alumni in the workplace now value or remember only some of the criteria for good teaching that current students rank and that they value other, different criteria as the most influential, valuable, or memorable. Faculty, evaluators, and other stakeholders should be more aware of the different overlapping sets of criteria that are applied to teaching effectiveness by different stakeholders. Some teaching approaches might well gain high current ratings, but low later ones, or even be assessed negatively, by other stakeholders.

Given the need to extend the domain of teaching effectiveness beyond current students, this study used a literature review, telephone interviews, and a pilot survey to develop a comprehensive survey of current and alumni assessments of teaching effectiveness. This survey was mailed to graduates of the Department of Communication at a large public university on the east coast to assess their perceptions of the criteria of instructional effectiveness. The following sections contain a brief summary of previous literature on assessing instructional effectiveness, a summary of the research project conducted on alumni perceptions of instructional effectiveness, and a discussion of the results and implications for student evaluations of instruction.

Literature Review

In developing student evaluation of instruction forms, researchers have considered the question "What is good teaching?" from the perspectives of students, trained observers, and alumni.

Dimensions of Teaching Effectiveness According to Current Students

Students have played a major role in identifying teaching dimensions used to evaluate instruction. The results of numerous studies spanning a half-century have resulted in a general consensus among researchers about the dimensions that are important to students (Braskamp, Ory & Pieper, 1981; Feldman, 1997; Marsh & Dunkin, 1997; Murray, 1997; Perry, 1997; Solomon, Rosenberg & Bezdek, 1964; Wotruba & Wright, 1975). These include: teacher's stimulation of interest in the course and its subject matter; teacher's enthusiasm; teacher's knowledge of subject matter; teacher's intellectual expansiveness; teacher's preparation and organization of the course; clarity and understandableness; teacher's elocutionary skills; teacher's sensitivity to and concern with class level and progress; clarity of course objectives and requirements; nature and value of the course material; nature and usefulness of supplementary materials and teaching aids; perceived outcome or impact of instruction; teacher's fairness and impartiality of evaluation of students; personality characteristics of the teacher; nature, quality and frequency of feedback from the teacher to students; teacher's encouragement of questions and discussion, and openness to opinions of others; intellectual challenge and encouragement of independent thought; teacher's concern and respect for students; teacher's availability and helpfulness; teacher's ability to motivate students to do their best and require high performance; teacher's encouragement of self-initiated learning; teacher's productivity in research-related activities; difficulty of the course and workload; classroom management; pleasantness of classroom atmosphere; individualization of teaching; and pursuit or attainment of course objectives (Feldman, 1997, pp. 392–394).

Possible Biases in Teacher Effectiveness Evaluations

Researchers have analyzed potential biases of student evaluations and shed some light on the actual (or nonexistent) relationships between the bias as justified (defining it as a non-bias) or unjustified (defining it as a genuine bias) and student evaluations of instruction. There is a general agreement that class size, workload/difficulty, prior subject interest, expected grades, reason for taking a course, the instructor's research productivity and the length of time between taking the course and evaluating the teacher may have little or no relationship with student evaluations of instructional effectiveness (Beatty & Zahn, 1990). Potential biases considered to be minor are instructor rank and years of teaching experience, course level, sex of students and/or instructor, administration and stated purpose of ratings, academic discipline, and personality of the instructor (e.g., Cooper, Stewart & Gudykunst, 1982). But there does seem to be a link between immediacy and student motivation and learning Christophel (1990).

Dimensions of Teaching Effectiveness According to Alumni

To expand the domain of instructor effectiveness assessment beyond current students, institutions of higher education have been surveying alumni ever since the 1930s, when the economic trauma of the depression called for accountability of educational outcomes. These early surveys focused on how well education was preparing graduates for the job market, and alumni surveys have continued to highlight employability to the present time (Pettit, 1991). Researchers are becoming increasingly aware of the unique perspective that alumni can bring to studies of teaching effectiveness. Parkyn (1991) argues that while cognitive outcomes may be effectively measured during students' academic careers, the perceptions of alumni are a necessary component of the assessment of affective outcomes. He defines cognitive and affective outcomes as follows: Cognitive outcomes concern the acquisition of formal knowledge, the level of academic achievement, the refinement of critical thinking skills, and the development of vocational skills. Affective outcomes include the development of self-concept, values, and beliefs, as well as a concern for avocations, satisfactory mental health, participatory citizenship, and fulfilled interpersonal relationships. He concludes that:

> while those assessment endeavors which measure currently enrolled students focus on short-term outcomes, alumni-based research provides an appropriate context within which to measure long-term objectives. . . . We should not limit alumni-based studies to the cognitive realm. (Parkyn, 1991, p. 8).

Alumni are particularly well-suited to provide both cognitive and affective input to measurements of long-term objectives. As Clark et al. (cited in Donald & Denison, 1996, p. 25) suggest, recent graduates are establishing their career directions, so they are ultra-sensitive to the utility of the skills and knowledge they gained as undergraduates, in contrast with current undergraduates, who in general can only speculate about the utility or significance of various aspects of their educational experience.

Alumni surveys also benefit from the inclusion of open-ended questions in addition to standardized measures of instructional effectiveness. Responses to these types of questions can illuminate concerns of alumni that have been unanticipated by the designers of student evaluations of instruction. For example, research shows that, in many cases, open-ended responses reveal that the skills that alumni name as being most important to their careers today are the same skills that they believe were most lacking in their undergraduate education (Pettit, 1991). Upon reviewing comprehensive alumni studies from the State University of New York (Albany), Tufts University, and Georgetown University, Pettit (1991) notes the agreement among alumni responses to open-ended questions that point out that, while effective oral communication ranked high on the list of important abilities and skills to have in the workplace, it always ranked very low in alumni lists of abilities and skills enhanced in college. Other skills that were consistently named as highly important in the workplace, but also not well learned as undergraduates, were functioning independently and thinking analytically and logically.

Donald and Denison (1996) also noted the unique types of responses alumni gave to open-ended questions:

> As manifested in the coding categories developed from the content analysis of graduates' comments, the features cited by the graduates were at a different level of specificity and concerned different topics from those they had been asked to rate. The most frequently cited features related to students' academic development, both general, for example, to acquire knowledge and the opportunity to improve themselves generally, and specific, for example, to acquire an in-depth knowledge of an academic discipline. (p. 31)

Various dimensions of teaching effectiveness or learning outcomes identified by alumni surveys include the ability to think critically, opportunities to interact with faculty, a balance between theory and practice, good oral and verbal communication skills, interest in lifelong learning, intellectual passion for the subject, clarity and organization, good discussion moderation skills, ability to introduce real-life examples into the class, acknowledging the student's worth and ability, challenging students to achieve, making connections to jobs, caring and respect toward students, and career advising (Adams & McClelland, 1997; Allen, Armstrong, & Gutierrez, 1990; Belcher, 1996; Carson, 1996; Donald & Denison, 1996).

Parkyn (1991) asserts that alumni-based evaluation of instructors is important for at least four reasons: (1) it provides the opportunity for "detached objectivity"; (2) it can ask for the graduates' perceptions concerning the extent to which their undergraduate experience cultivated responsible citizenship; (3) it permits alumni to differentiate between developmental outcomes which were the result of their undergraduate experience and environment and those which were promoted principally through contributing forces beyond the educational setting; and (4) it provides an appropriate context within which to measure long-term objectives (p. 7). Overall, Carson (1996) and Parkyn (1991) believe that, given the opportunity, former students will contribute data that current students are unable to give.

A comparison of the teaching dimensions used in evaluation forms intended for alumni with those used for student evaluations of instruction highlights the sharp contrast in the amount of research that has been done in this area. The information that can be gleaned from the literature cannot be readily categorized into a systematic list of alumni-specific teaching dimensions. Other than the surveys which are used to ask about institution-wide indicators of excellence, there is no organized body of knowledge about what alumni have to contribute to an evaluation of the actual instructional practices of the faculty of their alma mater. Thus, faculty are missing a valuable assessment that could help them become better teachers.

Linking Teaching Effectiveness to Learning Outcomes

If alumni have different information to offer than do currently enrolled students, it makes sense to reach out to them and obtain that information. However, regardless of who has identified dimensions of instructional effectiveness, that knowledge is useless unless there is evidence that certain dimensions really do have an effect on learning. How can we be sure that, when students or alumni say that they have learned something, they actually have?

When teachers are deciding how to maximize their students' learning, they must consider the extent to which each dimension will actually affect student learning. In the words of Abrami, d'Apollonia, & Rosenfield (1997), "We wish to know not only the extent student ratings reflect what instructors do when they teach but also the extent to which students learn course content" (p. 336).

Feldman (1987) analyzed existing student evaluations of instruction to draw conclusions about the correlation between teaching dimensions and student outcomes. However, there is general agreement among researchers, including Feldman, that such correlational studies are not sufficient to prove that certain teaching behaviors cause learning; instead, Abrami et al. (1997) argue, effective instruction needs to be evaluated in terms of the links between what teachers do and whether and how students change as a result.

Research that investigates the links between teaching practices and how those practices affect learning focuses on a number of underlying processes, including the psychological constructs of self-worth and perceived personal control (Perry, 1997), cognitive theories of information-processing and learning (Murray, 1997), teacher immediacy, solidarity and communicator style (Andersen, Norton & Nussbaum, 1981), and motivation (Christophel, 1990). Feldman (1997) concluded that the five dimensions that most affected student achievement were: teacher's preparation/organization of the course; clarity and understandableness; teacher pursued and/or met course objectives; perceived outcome or impact of instruction; and teacher's stimulation of interest in the course and its subject matter.

Thus, research shows that there are links between the teaching dimensions that are currently used in student evaluations of instruction to evaluate teaching and the actual learning of students. Feldman's conclusions about the relationship between teaching practices and student achievement, therefore, can be used to inform instructors about what teaching practices are most likely to encourage student achievement.

Given these conclusions and the importance of including multiple stakeholders in the teaching evaluation process, this study seeks to understand the contributions to our understanding of effective teaching that can be made by graduates of a communication program in the assessment of instructional effectiveness.

Research Questions

This discussion leads to the following research questions:

RQ_1: Which of the teaching evaluation criteria used in standard instructional evaluation forms provided to undergraduate students, and identified by prior research as generally relevant to stakeholders, are considered most important by alumni?

RQ_2: What are the underlying dimensions of all these criteria?

RQ_3: Which of these underlying dimensions are rated as most important by alumni?

RQ_4: Do these underlying dimensions differ by gender, or by recent versus older alumni?

RQ_5: What are the most important criteria, according to alumni, including any new ones?

RQ_6: Which of the important and additional criteria would alumni recommend be added to the standard instructional evaluation forms?

Methodology

The project consisted of three phases: telephone interviews, pilot study, and mailed survey. All stages were evaluated and approved by the University's Institutional Review Board. Along with the literature review, the telephone interviews and pilot study were conducted in order to develop a rigorous, comprehensive, and meaningful large-scale survey that reflected criteria used by current students and alumni in evaluating teaching effectiveness.

Telephone Interviews

To determine initial perceptions of alumni concerning criteria for effective instruction, the researchers conducted and recorded interviews over the telephone and transcribed the interviewees' comments. We were able to obtain comments and insights from a useful sample of recent and older, male and female, graduates from the Department of Communication at a large public university on the east coast. Alumni were asked to describe faculty members who they thought were excellent teachers. They were also asked what constituted good teaching from their perspective.

Sampling

Using the roster of graduates from 1989/90 and 1996/97, we created a quota sampling strategy to obtain five male and five female respondents from each of the two time periods. This procedure was not random, because we did not impose a call-back procedure to avoid potential selection biases involving those who were, or were not, available for the phone call. However, as the goal was to generate likely themes to complement the criteria identified by the literature review, it was deemed sufficient.

Telephone Interview Results

Comments represented four general dimensions of teaching excellence, listed in declining order of total comments:

- *balances practice and theory* (14 comments): presents practical information, has practical experience; balances theory and research, able to put theory into action; understands the 'working world'; offers solution-based instruction; balances teaching and research effort.
- *effective interaction* (9 comments): makes time for students, spends time with individual students, available for office hours; very approachable; listens to students' interests; has a good personality; gets along well with students.
- *good communication* (7 comments): effective interpersonal communication skills; excellent presentational skills; effectively communicates information; able to explain material clearly.
- *course management* (6 comments): makes class/lectures interesting and fun; willing to adapt the class to suit the types of people in it; explains expectations of course; tests assess students' ability accurately.

Pilot Study

Based on the instructional effectiveness criteria identified from the literature survey and from the telephone interviews with Communication graduates, we developed an initial survey asking people to rate a range of instructional criteria, to rank those they felt should be added to the current Student Instructional Rating Form, and to list any additional criteria.

We then conducted a pilot assessment of our initial survey in a master's level summer research course with 16 students. We analyzed the descriptive statistics, factor and reliability analyses, and frequency listings of added and ranked teaching criteria. The first author then returned to the class the following week to summarize the results, discuss the implications for, and ask for comments on, the next stage of the study. The survey instrument was revised based on this analysis.

Mailed Survey

Preparation of the Survey. Based on the telephone interviews and pilot study responses, we revised the final survey in the following ways: (1) dropped one item that had no variance; (2) extended the range of item values from 1–5 to 1–7 to increase the variance; (3) reformatted and reworded some questions to avoid ambiguities (such as concerning graduation and graduate programs) and converted some open-ended questions to close-ended questions (such as listing Communication specialization) to simplify coding; (4) used the results from the factor analyses to drop some redundant items, and rearranged the order of some questions to avoid a sequence of similar responses; (5) added clearer instructions; (6) added final information about the incentive for responding; and (7) developed and included a cover letter.

The final survey asked respondents to rate the importance of six criteria appearing on the standard University teaching effectiveness survey and of the 44 criteria developed from the literature review, telephone interviews and pilot survey. It then provided a section asking for their four most important criteria (taken from the 44 or new items they could write in); a section asking which of those the respondent might recommend be included on a revised standard University evaluation survey; and a section asking several demographic questions such as gender and graduation class year.

Sampling and Response Rate. A sample was drawn from a listing of Department of Communication graduates from 1989/90 and 1996/97 provided by the University Alumni Office. The total number of available graduate mailing addresses and the resulting sample from each year were: for 1989, 290/245; for 1990, 250/210; for 1996, 230/192; and for 1997, 120/102, for a total of 750 sampled from a population of 890.

Survey packets (including a cover letter explaining the purpose of the survey and incentive for participation) were mailed out near the end of June, with a request to return them by the end of July. Respondents were requested to return the survey anonymously in an envelope provided in the packet. Respondents wishing to enter a drawing for movie tickets were directed to write their names and addresses on a postcard provided and mail it separately to the researchers. In this manner, respondents were offered an incentive to participate while maintaining the anonymity of their responses. By the end of September, we received 188 complete surveys for a survey response rate of 28%. We recorded the date we received each survey so we could later test for a potential bias between early and late responders and, by implication, non-respondents.

Results

Demographic Characteristics of the Respondents

Seventy-three percent of the respondents were female, only slightly higher than the average proportion of women undergraduate Communication majors at the university. Nearly all respondents had graduated from the two listed time periods, although 10 respondents had graduated one or two years before or after that. In terms of the emphasis in their Communication major, 29% of the respondents had focused on public relations, 23% on interpersonal communication, 22% on mass communication, 10% on organizational communication, and a small percentage on international or health communication. Two-thirds of the respondents either planned to attend, were attending, or had already graduated from, graduate school. The current job positions of the respondents were widely distributed across professions.

RQ_1: Rated Importance of Criteria

According to alumni respondents, the 15 most important of the 50 individual criteria for judging instructional effectiveness were, in declining order: Knowledge of subject matter; Effectively communicates information to students; Explains material clearly; Responds effectively; Has effective communication skills; Is prepared and organized; Assists students in understanding material; Provides clear instruction; Is prepared; Assigns grades fairly; Generates interest in the material; Uses instructional methods that encourage learning; Has enthusiasm for subject teaching; and Has good speaking and presentation skills. Table 1 lists all 50 items in decreasing order of importance.

TABLE 1

RANKING OF CRITERIA BY RATED IMPORTANCE, IN DECREASING ORDER

Criteria	Mean	Factor
has knowledge of subject matter	6.76	Effective Comm
effectively communicates information to students	6.74	Effective Comm
is able to explain material clearly	6.63	Effective Comm
responds effectively to student comments and questions*	6.62	Availability
has effective communication skills	6.62	Effective Comm
is prepared for class and presents the material in an organized manner*	6.58	Effective Comm
has a positive attitude toward assisting all students in understanding course material*	6.55	Availability
provides clear instruction	6.55	Effective Comm
is prepared	6.50	Effective Comm
assigns grades fairly*	6.48	–
generates interest in the course material*	6.46	Interest
uses instructional methods that encourage learning*	6.41	–
has enthusiasm for subject or for teaching	6.39	Interest
has good speaking and presentation skills	6.36	–
provides appropriate feedback both in-class & through grades	6.35	–
assigns useful and relevant course material	6.33	–
makes class and lectures interesting and fun	6.31	Interest
stimulates interest in course & subject	6.31	Interest
stimulates interest in the course & subject matter	6.30	Interest
is very approachable	6.27	Availability
evaluates student progress fairly	6.23	Fairness
shows concern, enthusiasm, respect and tolerance for students	6.20	Availability
understands the "working world"	6.16	Practical App
maintains a classroom climate conducive to learning	6.15	Fairness
encourages questions, discussions, & openness to others' opinions	6.14	Interaction
explains expectations of the course	6.13	Effective Comm
motivates students to greater effort, overall learning, higher achievement	6.13	–
is able to put theory into action	6.12	Practical App
organizes course well	6.12	Effective Comm
presents practical info	6.11	Practical App
assesses students' ability accurately	6.09	–
directs and moderates class discussions well	6.09	Interaction
provides instructor/student in-class interaction discussions	6.08	Interaction
is available for office hours	6.07	Availability
uses/chooses appropriate methods/materials	6.03	Fairness
makes time for, spends time with, students	5.97	Availability
has practical experience	5.93	Practical App
offers challenging course workload	5.90	–
is sensitive to class progress	5.88	–
emphasizes students' oral, written, and presentation skills	5.84	Interaction
gets along well with students	5.73	Accessibility
listens to students' interests	5.69	Accessibility
demonstrates intellectual range	5.59	Research Exp
balances theory & research	5.54	Practical App
encourages group activities & discussions	5.53	Interaction
offers solution-based instruction	5.49	Research Exp
balances teaching & research	5.37	Practical App
provides opportunities for interaction & projects with faculty	5.05	Interaction
has good research productivity & reputation	4.85	Research Exp
uses audio-visual materials	4.78	Research Exp

Note. The survey asked, "In your opinion, how important are these as criteria for evaluating teachers?" The question portion began with "The instructor . . .", followed by the criteria. The rating scale was 1 = Very unimportant, 2 = Unimportant, 3 = Somewhat unimportant, 4 = Neutral, 5 = Somewhat important, 6 = Important, and 7 = Very important. – = Item did not load highly on any factor.

*One of the current University evaluation form questions.

The most important criteria in terms of frequency of open-ended mention, taken from either the 44 items or new entries listed by the respondent (each receiving from 2% to 7% of all responses) were: Is prepared and organized; Understands the working world; Motivates effort toward learning; Emphasizes student oral and written presentation; Shows concern, enthusiasm, and respect towards students; Shows enthusiasm toward the subject; Has and relates practical experience; Makes class lectures interesting; Moderates discussions and group interactions well; Stimulates interest in the course; Has knowledge of the subject; Presents practical information; Encourages questions and discussion; Presents useful course material; Offers solution-based teaching; Assists in understanding the material; Explains material clearly; Uses appropriate methods and audio-visuals; Responds effectively to questions; Effectively communicates information to students; Gets along well with students; Puts theory into action; Offers challenging course work; Generates interest in the material; and Emphasizes good speaking and presentation skills.

RQ_2 and RQ_3: Underlying Dimensions of Teaching Effectiveness Criteria and Their Importance

Principal components analysis (with varimax rotation) identified eight underlying dimensions, or potential evaluation scales. We can assess the importance of these factors or sets of related instructional effectiveness criteria in two ways: the amount of variance among all the 50 rated items that each factor represents, and the average importance of each scale formed from the mean of the highest-loading items of each factor (items that load .60 or higher on one dimension and .40 or lower on any other dimension). Table 2 provides the factor labels, variance explained, eigenvalue, overall reliability, and mean of each of the scales. Table 1 indicates which specific items indicated which factors.

The first factor, labelled "Effective communication," had the highest importance rating of the teaching criteria mean scales, and explained nearly half the variance of the 50 items. It includes items indicating effective communication skills, good organization and preparation, and explains material and expectations. The second factor, labelled "Interest," represents criteria such as generating interest in the course material and has enthusiasm for the subject or teaching. The third factor, labelled "Interaction," emphasizes activities such as fostering instructor/student interaction, within-class discussion groups, and students' communication skills. The fourth factor, labelled "Availability," was indicated by criteria such as responding effectively to students' questions, assisting students, is available, and shows concern and respect for students.

TABLE 2

SUMMARY OF FACTOR ANALYSIS OF TEACHING EFFECTIVENESS CRITERIA

Factor	Variance Explained	Eigenvalue	Scale Mean	Scale S.D.	Alpha Reliability
Effective communication	44.9%	22.4	6.51	.61	.91
Interest	5.6	2.8	6.36	.74	.91
Interaction	4.1	2.1	5.71	.86	.86
Availability	3.4	1.7	6.30	.73	.87
Practical applications	2.9	1.5	5.82	.84	.82
Fairness	2.6	1.3	6.15	.74	.89
Research expertise	2.5	1.3	5.18	.90	.75
Accessibility	2.1	1.1	5.71	1.06	.81

Note. Results from principal components analysis, varimax rotation. Scale means constructed from items loading strongly on the corresponding factor. Reliability is Cronbach's Alpha.

The fifth factor, labelled "Practical applications," included criteria such as understands the "working world," and balances theory, research, and practice. The sixth factor, labelled "Fairness," was represented by evaluates student progress fairly and fosters a learning climate. The seventh factor, labelled "Research expertise," included items such as shows intellectual range and productivity, and offers solution-based instruction. The eighth factor, labelled "Accessibility," included gets along well with students, and listens to students' interests. Eleven items (identified in Table 1) did not load highly on any factor.

RQ_4: Differences Across Gender, Class Year, and Reply Date

The eight mean scales, based on the high-loading criteria of the respective factors, were then tested for three possible influences or biases. Results from independent pairs t-tests on gender (male/female) and class year (recent or older) showed there were no significant differences for any of these eight mean scales except a slightly greater importance placed on the first dimension ("Effective communication") by older alumni (6.6 vs. 6.4; $t = 2.1$, d.f. $= 122$, $p < .05$). Given this lack of difference in responses from more recent and older alumni it might be inferred that alumni in general do not value the dimensions differently than current students, though, as noted below, there were considerable differences in additional criteria suggested by recent versus older alumni.

There was no significant correlation between the number of days from when the first survey was returned until each other survey was returned, and any of the eight mean scales (the average correlation was $r = -.05$, n.s.). This lack of relationship between survey receipt time and the mean responses is an implied indication that there would not likely be a systematic difference between respondents and non-respondents. Given that the surveys were anonymous, so that we could not develop a small sample of non-respondents, there was no other easy way to assess non-response bias.

RQ_5: Most Important Named Criteria

After the section on the survey where respondents were asked to rate the importance of the 50 effectiveness criteria, the survey then asked the respondents to first mention up to four of their most important criteria for teaching effectiveness evaluation (either from the 44 non-University items, or any other ones they wrote in), and then to recommend any items (from any of the 44 or the additional items) that they felt should be added to the standard University teaching evaluation survey.

Responses were examined to determine if the alumni listed criteria for effective teaching which had not been included on the survey. Some of the items mentioned had been reported in prior studies but had not been specifically listed on prior instructional effectiveness surveys, and so had not been included on our survey. Older alumni did mention many more new criteria (32 compared to 9) than recent graduates, and they suggested a nearly completely different set of new criteria for instructional evaluation. Recent graduates mentioned teachers should not overgeneralize their opinions; Have a passion for teaching; Not sacrifice learning for entertaining classes; Understand student situations; Have a positive attitude toward assisting students; Assign a variety of assignments, and Be willing to reverse their position. Older alumni mentioned items such as Promotes presentation and writing skills; Clearly explains and instructs; Deals with sexual harassment questions; Relates subject to real world; Balances theory and applications; Is not opinionated; Has hands-on experience related to the course; Teaches practical materials; Be prepared; Be approachable; Motivates students; Have more enthusiasm; Assign-

ments provide practical experience; Be aware of current events in communication; Experience is connected to communication instruction; Shows a genuine interest in students; Can relate to students; Treats students with respect; Stays current and up-to-date; Has guest speakers with field experience; Fosters career networking in the field; Understands relevant job market; Understands the pros and cons of related careers; Emphasizes students' strengths; Periodically assesses the course objectives; and Emphasizes writing skills.

RQ_6: Criteria Recommended to be Added

The criteria most frequently mentioned (by at least 4% of the respondents) as ones that might be added to the standard student evaluation form included: Class lectures are interesting; Motivates effort toward learning; Spends time with students; Is approachable; Understands the working world; Effectively communicates information; Explains material clearly; Shows enthusiasm for course material; Has knowledge of the subject matter; Has good speaking and presentation skills; Seeks feedback in class; and Emphasizes oral and written presentations.

Discussion

While student evaluations of instruction are a generally accepted evaluation tool at institutions of higher education, they do not fulfill all of an institution's needs for constituent evaluation. Undergraduate alumni of an institution are unique among external stakeholders in that they have been participant observers in the classroom and have applied the results of that experience in a non-undergraduate-academic environment and, therefore, may provide valuable insight into the dimensions of effective instruction.

As summarized earlier, several reviews of the literature and analyses identified a variety of dimensions of instructional evaluation and, therefore, good teaching which leads to student learning. Our results indicate that alumni respondents believe that by far the most comprehensive and important criterion for teaching effectiveness is teachers' effective communication, organization and preparation. This primary evaluation dimension is followed by seven other dimensions: (2) generates, and has, interest in course material and topic, (3) interacts with students and fosters students' discussion and communication skills, (4) is available and open to questions, (5) has practical application of content and "real world" connections, (6) evaluates students fairly and promotes a classroom learning climate, (7) demonstrates good research expertise, and (8) listens to students' interests. Communication abilities and processes, both on the part of the teacher, and on the part of student groups and interactions, permeate most of the effectiveness dimensions.

Communication instructors often receive some of the highest teaching evaluations in a university. Beatty and Zahn (1990) conclude that students' perceptions of communication instructors as more sociable and qualified than other instructors leads to their higher evaluations. They stress that this is an indication of healthy student-teacher interaction which is in the best interest of the instructor, department, and the field. The present research presents several factors perceived by alumni as major components of this "healthy student-teacher interaction."

This research has practical applications for instructors in their day-to-day teaching practice. Scott and Nussbaum (1981) argue that communicative behaviors that are associated with higher learning and higher instructor evaluations are teachable. Thus, it should not be difficult for instructors to learn skills for classroom communication that will be more positively evaluated and contribute to greater student learning. From the perspective of graduates of communication programs, these skills should include: organization, interaction with students, sociability, interest, and clarity of presentation. It is notable that alumni support the finding that being an effective communicator is the most significant dimension in evaluating instructional effectiveness, in addition to an increased emphasis on real-world applications of theory and research. Alumni in this study favorably evaluate instructors who are clear, prepared, and organized. These are important goals for all educators to strive to achieve and maintain in their classrooms.

Part of the motivation of this study, as noted initially, was to begin to expand the domain of teaching effectiveness in our own department's evaluations. To that end, we selected several of the highest-loading items on each factor, giving some preference to the factors rated overall more important by the respondents, and to avoiding redundancy with the current six teaching evaluation criteria on the standard University form. As there were only nine extra, blank question lines on the University form, we have added the following items to our departmental survey: The instructor. . .

- Has good knowledge of the subject matter
- Explains material clearly
- Shows enthusiasm for the subject
- Moderates discussions well
- Fosters professor-student interaction and discussion
- Understands the working world
- Uses appropriate methods and materials
- Displays a wide intellectual range
- Gets along well with students.

Using these items on student teaching evaluations should lead to a fuller assessment of teaching that could be linked to student learning. Our department will assess the utility of these additional teaching effectiveness criteria, and will be among the first to be able to refer to an expanded domain of effectiveness evaluation.

These results could also be used to develop and test an overall teaching effectiveness scale, by using the high-loading items, which constitute eight reliable scales, augmented by those additional named items considered appropriate by the researcher.

When evaluating teaching, it is preferable to rely on multiple perspectives offered by a variety of institutional stakeholders (Ruben, 1997). One reason for considering these perspectives is that evaluations of instruction are used for different purposes, each requiring its own type of data. Data gathered for personnel decisions would most appropriately focus upon answers to general questions that elicit general levels of satisfaction, whereas data gathered for faculty improvement should be more specific, so that instructors can identify their specific strengths and weaknesses.

Evaluations of instruction should be based on criteria that have been agreed upon by an institution and its stakeholders. Each criterion can be measured by several appropriate indicators, such as student and/or alumni evaluations of instruction, colleague or trained observer assessments through observation, and reviews of lecture notes. Adding alumni evaluations to this process can provide helpful information to be used in the development and maintenance of effective instruction.

References

Abrami, P. C., d'Apollonia, S., & Rosenfield, S. (1997). The dimensionality of student ratings of instruction: What we know and what we do not. In R. P. Perry & J. C. Smart (Eds.), *Effective teaching in higher education: Research and practice* (pp. 321–367). New York: Agathon Press.

Adams, S., & McClelland, L. (1997). *CU-Boulder alumni: Reflections on the college experience and post-graduation activities of 1991 graduates.* Unpublished manuscript, University of Colorado at Boulder.

Allen, M. J., Armstrong, C. A., & Gutierrez, D. M. (1990). *Alumni vs. faculty opinion on undergraduate psychology programs.* Paper presented at the Annual Meeting of the American Psychological Association, Boston, MA.

Andersen, J. F., Norton, R. W., & Nussbaum, J. F. (1981). Three investigations exploring relationships between perceived teacher communication behaviors and student learning. *Communication Education, 30,* 377–392.

Beatty, M. J., & Zahn, C. J. (1990). Are student ratings of communication instructors due to "easy" grading practices? An analysis of teacher credibility and student-reported performance levels. *Communication Education, 39,* 275–282.

Belcher, M. J. (1996). *BSU's impact on skills valued by graduates.* Boise, ID: Boise State University.

Braskamp, L. A., Ory, J. C., & Pieper, D. M. (1981). Student written comments: Dimensions of instructional quality. *Journal of Educational Psychology, 73,* 65–70.

Carson, B. H. (1996, Nov./Dec.). Thirty years of stories: The professor's place in student memories. *Change: The Magazine of Higher Learning, 28,* 10–17.

Christophel, D. M. (1990). The relationships among teacher immediacy behaviors, student motivation, and learning. *Communication Education, 39,* 323–340.

Cooper, P. J., Stewart, L. P., & Gudykunst, W. B. (1982). Relationship with instructor and other variables influencing student evaluations of instruction. *Communication Education, 30,* 308–315.

Donald, J. G., & Denison, D. B. (1996). Evaluating undergraduate education: The use of broad indicators. *Assessment & Evaluation in Higher Education, 21,* 23–39.

Feldman, K. A. (1987). Effective college teaching from the students' and faculty's view: Matched or mismatched priorities? *Research in Higher Education, 28,* 291–344.

Feldman, K. A. (1997). Identifying exemplary teachers and teaching: Evidence from student ratings. In R. P. Perry & J. C. Smart (Eds.), *Effective teaching in higher education: Research and practice* (pp. 368–395). New York: Agathon Press.

Fram, E., & Camp, R. (1995). Finding and implementing best practices in higher education. *Quality Progress, 28*(2), 69–73.

Glassick, C., Huber, M. T., & Maeroff, G. I. (1997). *Scholarship assessed: Evaluation of the professoriate.* San Francisco, CA: Jossey-Bass.

Marchese, T. (1997, Sep./Oct.). Student evaluations of teaching. *Change: The Magazine of Higher Learning, 29,* 4.

Marsh, H. W., & Dunkin, M. J. (1997). Students' evaluations of university teaching: A multidimensional perspective. In R. P. Perry & J. C. Smart (Eds.), *Effective teaching in higher education: Research and practice* (pp. 241–320). New York: Agathon Press.

Murray, H. G. (1997). Effective teaching behaviors in the college classroom. In R. P. Perry & J. C. Smart (Eds.), *Effective teaching in higher education: Research and practice* (pp. 171–204). New York: Agathon Press.

Parkyn, D. L. (1991). *The role of alumni in the assessment of affective outcomes.* Grantham, PA: Messiah College.

Perry, R. P. (1997). Teaching effectively: Which students? What methods? In R. P. Perry & J. C. Smart (Eds.), *Effective teaching in higher education: Research and practice* (pp. 154–168). New York: Agathon Press.

Pettit, J. (1991). Listening to your alumni: One way to assess academic outcomes. *The Association for Institutional Research, 41,* 1–10.

Plater, W. (1997, Sep./Oct.). In response. . . . *Change: The Magazine of Higher Learning, 29,* 16–17.

Ruben, B. (1997). *Excellence in higher education: A guide to self-assessment, strategic planning and improvement.* Dubuque, IA: Kendall-Hunt.

Scott, M. D., & Nussbaum, J. F. (1981). Student perceptions of instructor communication behaviors and their relationship to student evaluation. *Communication Education, 30,* 44–53.

Smart, J. (Ed.) (1997). *Higher education: Handbook of theory and research* (vol. 1). NY: Agathon Press.

Solomon, D., Rosenberg, L., & Bezdek, W. (1964). Dimensions of teacher behavior. *Journal of Experimental Education, 55,* 23–30.

Trout, P. (1997, Sep./Oct.). What the numbers mean: Providing a context for numerical student evaluations of courses. *Change: The Magazine of Higher Learning,* 25–30.

Wotruba, T. R., & Wright, P. L. (1975). How to develop a teacher-rating instrument: A research approach. *Journal of Higher Education, 46,* 653–663.

Received June 7, 1999
Final revision received October 27, 1999
Accepted December 14, 1999