A Systematic Approach to Documenting Teaching Effectiveness

Denise M. Anderson
Clemson University
Parks, Recreation, and Tourism Management
263 Lehotsky, Box 340735
Clemson, South Carolina 29634-0735
864-656-5679
dander2@clemson.edu

Francis A. McGuire
Clemson University
Parks, Recreation, and Tourism Management
263 Lehotsky, Box 340735
Clemson, South Carolina 29634-0735
864-656-2183
lefty@clemson.edu

All correspondence should be directed to Denise Anderson.

Abstract

Faculty in higher education are well-versed in their two primary job responsibilities of teaching and research with service often a distant third. Although the evaluation of research is reflected in one's publication record, the evaluation of teaching is often more haphazard and unreliable. While many departments rely solely on student evaluations of teaching and/or peer evaluations, there are a multitude of other evaluation tools that can be utilized in the development of a well-rounded teaching portfolio. Teaching evaluations are often used for personnel decisions ranging from promotion and tenure to merit increases. Therefore, it is important for a faculty member to take a proactive stance and develop a systematic and comprehensive evaluation plan that will serve to document his or her teaching abilities in a valid and reliable manner. The purpose of this paper is to identify several tools useful in documenting teaching effectiveness and present a way to build the case for teaching competence.

Keywords: teaching effectiveness, evaluation

Boyer's (1990) landmark work examining higher education emphasized the scholarship of teaching as one of faculty's primary roles. This recognition was accompanied by increased expectations that the evaluation of teaching should be more than an evaluation of the classroom act and should encompass the scholarship of teaching. This requires expansion of the components of teaching evaluations beyond those student ratings that originated in the 1960s (Knapper, 2001) to include an evaluation of the range of activities that comprise the teaching act. A reality for faculty today is that the evaluation of their teaching will be a factor in decisions about reappointment, promotion, tenure or post-tenure review. Therefore, it is crucial that faculty find ways to capture the scholarship of teaching in a manner that is reflective of their efforts as well as their competence.

An article by John Ory (2000) traced the evolution of teaching evaluations from the 1970s (primarily done for developmental reasons including faculty improvement) through the 1980s and 1990s (driven by administrative needs related to retrenchment and budget rather than student or faculty needs) through recent years (driven by renewed interest in improvement of education, demands for accountability, needs of a legal system seeking improved evaluations and demands by young faculty for more valid indicators of teaching effectiveness). According to Ory the present culture is marked by "collecting more information from multiple audiences and taking more care in doing so than ever before" (p. 14). The future of teaching evaluations should be driven by a systematic effort to find methods that will not necessarily yield more information but rather will yield better information.

A problem with evaluating teaching is that evaluations serve two distinct purposes, personnel management based on institutional needs and self- improvement based on individual need, and attempting to address these purposes in a single evaluation may not be effective (Casey, Gentile & Bigger, 1997; Mills and Hyle, 1999). In addition, difficulties in evaluating teaching arise when the focus is on the techniques of teaching rather than the substance of teaching (Pratt, 1997). The authors of this paper have over 30 years of teaching experience and evaluations of their teaching. Never have either of us received an evaluation that examined primarily the content of what we taught. Evaluations revolve around how we teach, typically evaluated by students, rather than what we teach and as a result evaluations are incomplete and possibly misleading. Another difficulty in many evaluations of teaching is that they are primarily student-centered, even though there are multiple stakeholders, including administrators, employers, parents, and professionals (Knapper, 2001), in those evaluations. The typical evaluation process is focused on the student evaluation of teaching, but largely ignores evaluation by academic peers, practitioners and the individual evaluated. The position we take in this paper is that evaluations of teaching must be comprehensive and systematic: formative as well as summative, qualitative as well as quantitative, technique based as well as content based, student centered as well as peer and self-centered.

Evaluating teaching effectiveness is a task faced by nearly all faculty and one often left to the whimsy of a cadre of undergraduate students. Few of us take the time to develop a systematic evaluation plan and carry out that plan. In fact, our experience is that in many departments the time taken by one faculty member to develop one manuscript for publication exceeds the time devoted by all faculty in the department to conduct an evaluation of their teaching. The simplest, and least time consuming, approach to evaluations is to use student evaluations of teaching (SET) as the sole device. While the literature on SET indicates they have value as evaluative measures (Algozzine, et al., 2004; Aultman, 2006) this value is primarily limited to student observation of what happens in the classroom, issues such as organization, fairness and teaching style. However, that falls far short of capturing the comprehensive nature of teaching. Hutchings (1994) stated that faculty must find methods of teaching evaluation that "capture the scholarly substance of teaching" and this includes not solely what happens in the classroom but all aspects of the act. The question is then how can faculty construct a representation of what they do that includes the multiple perspectives needed to capture the teaching process.

Ory (2000) uses a metaphor he identifies as "sitting beside" to describe the ideal use of teaching evaluations, a process marked by conversation and discourse and sharing perspectives prior to sharing judgments. The "sitting beside" process "promotes a developmental perspective. It is not a single snapshot but rather a continuous view" facilitating "development rather than classifying and ranking the faculty" (p. 16). The result of this process is the opportunity for faculty to "make their case" and document teaching effectiveness while obtaining help to improve teaching. There are multiple questions to address with this technique starting with *what would you want to include in making your case as an effective teacher if you had the opportunity to sit beside the head of your department during his or her review? What artifacts of teaching and information about your effectiveness would you want the chair of your departmental promotion, tenure and reappointment committee to have? Would you want your case to rely solely on SET? If not, what other forms of evaluation would you want to share with those making personnel decisions?* Donald French (2006) asks the same question but puts it in the framework of currency. He states that before effectively evaluating teaching we must answer the question: "what's the currency" meaning how and what should be evaluated. We believe that faculty should define their own currency.

The purpose of this paper is to identify several tools useful in documenting teaching effectiveness and present a way to build the case for teaching competence. The first step is to determine your university's policies on documentation of teaching competency. If there is a limit on what information is acceptable than what we suggest below may not be useful. However, your university may have few requirements and even fewer limits on establishing teaching effectiveness. Our faculty manual stipulates that student evaluations must be part of the faculty evaluation process. However, it also allows the use of other materials such as class visitations by peers, exit interviews with graduating students and peer evaluation of class materials in the process. If your campus is as open as ours to the use of a variety of artifacts to establish teaching effectiveness, the approach detailed below is an effective way to document your competence as a teacher..

Methods of Evaluation of Teaching

Selected authors (Baldwin & Blattner, 2003; Casey et al., 1997; French, 2006; Gastel, 1991; Macalpine, 2001) have recommended a multi-method approach to the evaluation of teaching and have identified multiple sources for evaluation information including self-assessment, disciplinarily colleagues, teaching experts, current students, graduates and employers. Panici (1999) identified three phases of teaching that could be included in a comprehensive evaluation: the pre-interactive phase including planning and preparation; the delivery phase when actual teaching occurs; and the post-interactive phase based on reflection and revision of course material. Based on this it is possible to identify four sources of information useful for the evaluation of teaching:

1. Student-centered sources
2. On-campus peer-centered sources
3. Off-campus peer-centered sources
4. Self-centered sources

Student-centered sources. The traditional method of student evaluation is the students' evaluation of teaching done at the end of a semester. Students are asked, usually anonymously, to rate their instructors on a variety of statements, typically related to organization, fairness and platform skills. At our university students also respond to open-ended questions related to strengths and weaknesses of the instructor, and these may be the most useful of all information provided on SET (Hodges & Stanton, 2007). There is a large body of literature related to SET and many authors acknowledge their value while others question their utility (Heckert, Latier, Ringwald & Silvey, 2006; Heckert, Latier, Ringwald-Burton & Drazen, 2006; Martinson, 2000, 2004; Panici, 1999; Sojka, Gupta & Deeter-Schmelz, 2002; Steiner, Holley, Gerdes & Campbell, 2006). However, students' perceptions of teaching are usually limited to what occurs in the classroom and may not reflect the broader aspects of the scholarship of teaching. For example, few student evaluations are accurate reflections of course content, the importance of material to future classes and careers and the recency of class material. It is possible to be an entertaining and compassionate teacher with effective classroom skills whose material is out of date or incorrect. There are additional student-centered evaluation methods available. During a workshop offered by one of the authors a participant from chemistry shared her frustration with SET. She believed that her evaluations were low because her introductory science class was difficult and students expressed their frustrations with her class on their evaluations. However, many of those students expressed their appreciation for her course after taking advanced level chemistry classes and realizing they were better prepared than their peers. The advice given to the instructor was to find a way to document the students' perceptions of their preparation.

Similarly, evaluations provided by recent graduates, often gathered during an exit interview, and alumni, are important pieces of information in building the case for effective teaching. Our university, and many others, contact alumni who are five years out. The data provided by the graduates are an important piece of information, and often contradictory to evaluations completed by students immediately after a class. The quality of teaching seen through the prism of five years experience in the field is often different since it is informed by the application in the "real world" of knowledge gained in a class.

On-campus peer-centered sources. Many campuses have an office devoted to teaching effectiveness. On our campus the Office of Teaching Effectiveness and Innovation will send a teaching expert to a class to complete an in-depth evaluation. Although this evaluation is not specifically required as part of a departmental evaluation process, faculty are free to include it with their artifacts intended to demonstrate effectiveness. This type of evaluation is effective in gauging the process of teaching; our director of the teaching center refers to these as platform skills, but it is not an effective evaluation of content.

The most common on-campus peer evaluation of course content is peer-review by faculty in one's department. This peer reviewer process, while convenient and traditional, should be used with caution (Yon, Burnap & Kohut, 2002). Concerns about reliability and validity of observation instruments, reviewer bias (positive or negative), and lack of observation training all contribute to doubts about the effectiveness of peer review (Yon, et al., 2002). Careful planning and preparation of observers, as well as those observed, are crucial to the success of peer-reviews (Mento & Giampetro-Meyer, 2000).

An additional source of evaluative material often overlooked is peers who teach courses in the department that require your course as a prerequisite. In the situation described above a junior faculty member from chemistry was decrying the negative feedback she received from students on the semester evaluation and the failure of the students to recognize the value of her class. When asked how she knew her class was effective she indicated that instructors in later courses indicated her students were better prepared than students from other classes for the advanced classes. The difficulty was capturing the phenomenon for her evaluation dossier. In addition to soliciting feedback from her students, she decided to contact faculty and ask them to provide an evaluation of her students' preparation for upper level classes, with the information sent directly to the departmental promotion, tenure and reappointment committee.

An additional approach to on-campus evaluations from peers is through the variety of experts whose area of focus is part of your classes. For example, a laptop class or an on-line class could be evaluated by the technology office on campus. Similarly, distance education staff can evaluate class web sites.

Off-campus peer-centered sources. Evaluations provided by colleagues who are in departments at other campuses or working in the field are rich sources of evaluative feedback. There are several options for soliciting evaluative comments from off-campus peers. A faculty member from another school who is known, particularly by members of your departmental review committee, for expertise in your course topic might provide feedback on your syllabus. The feedback could include whether your content reflects the current state of knowledge, whether your course objectives are appropriate and comprehensive, whether the tests and readings are reflective of the knowledge in the field and whether assignments, projects and exams are appropriate for the level and focus of your class. If an external reviewer is interested in giving additional feedback perhaps you could provide a video of one of your classes, although the information provided by the reviewer may be limited to that event.

An additional source of an off-campus peer reviewer might be practitioners. Practitioners could provide information about the content of a course, similar to that provided by faculty. However, they may also provide information about the professional capabilities of your students. Similarly, internship supervisors may be able to provide feedback about the preparation of the interns specifically related to the materials you taught in your classes. For example, if you taught the leadership class in your curriculum comments related to your students' abilities to lead groups would be crucial information in establishing the effectiveness of your teaching.

Another way to establish your teaching effectiveness is through the success of your students on certifications exams. We receive an annual report from the National Council on Therapeutic Recreation specifying the pass rates of our students as well as areas of the test where they did well or poorly. This information is a powerful way to make the case for teaching effectiveness as it measures performance on a national test and allows comparison to the performance of other students. Of course the information is useful only if it can be related to your classes. In a department with only one or two faculty members in an area such as therapeutic recreation the link is easier to make.

Self-centered evaluation. Faculty should have an opportunity to participate in their own evaluations. There are many levels of participation and faculty should choose those that best assist in making their case for effective teaching. In many departments at our university faculty are required to include a teaching philosophy with their material for promotion or tenure and this philosophy is part of the teaching component of the evaluative process. Similarly faculty are permitted to place rejoinders to the student evaluation forms in their promotion or tenure materials. Faculty may include materials documenting their improvement over time in the teaching area, often in response to evaluative feedback. For example, one faculty member indicated steps she had taken, from attending workshops to interviewing her students, to improve her classroom instruction. These types of reflective materials, as well as various artifacts of teaching such as samples of students' work, videotapes of teaching, bibliographies of class materials, a record of engagement in teaching seminars, or a journal detailing class activities are traditionally artifacts included in teaching portfolios, an evaluation device popularized in the 1990s (Appling, Naumann & Berk 2001; Babin, Shaffer, & Tomas, A. M., 2001; Baldwin & Blattner, 2003; Reece, Pearce, Melillo & Beaudry, 2001).

Making the Case for Teaching Effectiveness

The sources of data provided above are one component of making the case for teaching effectiveness. Issues such as how the data are collected and how they are used are also crucial. At the end of this article there is a template (Figure 1) for designing a plan to evaluate teaching. Figure 1 provides a worksheet reflecting the other components in a systematic teaching evaluation process. The system includes five components:

1. What are the aspects of effective teaching you want to evaluate?

2. Who is most capable of determining effectiveness for each aspect?

3. When should the appraisal occur?

4. How should the evaluation be done?

5. What evidence is necessary to document effectiveness?

What to evaluate. Teaching is a multi-faceted activity and deciding what components to evaluate is critical. There is evidence (Hativa, Barak and Simhi, 2001) that there are multiple components and paths to teaching effectiveness. Polya (1957) stated, "I can give you no rules [for effective teaching], for there are as many good ways of teaching as there are good teachers" (p. 725). Therefore the first task in designing an evaluation system is to identify those areas that best define your teaching and your effectiveness. One approach is to determine what constitutes good teaching and evaluate your efficacy related to those components. According to Knapper (2001) there is seldom such an attempt and therefore the criteria of good teaching are often in the eye of the beholder. Students, peers and review committees may have divergent opinions of what constitutes good teaching and therefore have equally divergent opinions of what reflects good teaching. Given such a chasm it is incumbent of faculty undergoing evaluation to clearly define what it is that is being evaluated. Certainly there is a focus on content, whether it is the history of parks, the budgeting process, medical protocols or marketing strategies. However, there are other indicators of teaching effectiveness that you may want to evaluate. If organization is a component of your effectiveness it should be part of the evaluation. If teaching students to be creative thinkers is important than that should be measured and evaluated. The key is staking your claim and then providing evidence your are reaching your goals.

Who should evaluate. Once the decision of what to evaluate is made, the next step in the process is determining the best evaluator. Students are capable of evaluating instructor enthusiasm, willingness to work with students and fairness in grading. However, as discussed above, other faculty, practitioners, teaching experts or others may be appropriate evaluators for other dimensions of teaching. For example, the director of our Center for Ethics on our campus could evaluate whether students have developed a sense of ethics.

When the evaluation should occur. Typically evaluation occurs at the completion of a class. These summative designs are useful but we suggest that a more comprehensive strategy is more effective. The evaluation process should occur at every step of the teaching process, beginning with syllabus development and continuing through the professional career of students. Syllabus reviews completed prior to the first day of class may result in an improved class for students. Evaluations gathered from alumni may provide richer information about the students' experience in a class than SET completed during the last day of a semester. The answer to when evaluative materials should be collected depends on the purpose of the evaluation (formative or summative), the nature of the information gathered and the source of the information. For example, if data of student success on national exams is part of the evaluation package, the information will typically not be available until after students have graduated. Peer review of teaching obviously must occur during a semester.

How evaluative materials should be gathered. The validity of an evaluation requires rigor and structure in gathering materials. For example, letters from students, or their parents, attesting to how wonderful a class was, or how miserable, do not provide useful information for an evaluation designed to reach decisions about promotion, reappointment or tenure. The users of the information, including department chairs or promotion committees, should receive evaluation materials directly. A faculty member from another campus, who is completing a syllabus review as part of tenure or promotion decision, should not send the review to the faculty member under evaluation. We suggest that evaluation materials have merit if they are gathered in an unbiased way. For example, if outside reviewers are examining materials such as syllabi or tapes of lectures, promotion and tenure committees should follow the same process they use in soliciting external peer review letters. At our campus that involves the faculty under review providing a list of potential reviewers and then the promotion and tenure committee selecting the individuals who are asked to provide a review. Reviewers send review letters directly to the committee and all letters are included in the dossier. If the purpose of the evaluation is to improve teaching rather than for administrative decision making then information material should go directly to the faculty member. For example, if the purpose of a syllabus review is to upgrade the syllabus then the review is logically sent directly to the faculty member.

What evidence should be gathered. The nature of the artifacts used for evaluations will depend on what is being evaluated. In the example above regarding teaching ethics, student papers or projects related to ethics might be provided to the Center Director. They would guide his or her evaluation of whether students developed an ethical perspective. If part of the evaluation focused on knowledge, perhaps faculty members teaching advanced classes in a discipline might evaluate final projects submitted by students. If the focus was on the accuracy of class content, a syllabus review by faculty at other programs might provide part of the comprehensive evaluation.

Conclusion

At the end of this article there is a template (Figure 1) for designing a plan to evaluate teaching. The first row is completed and illustrates how the template could be used as a formal statement of the evaluation plan, thus providing the faculty member with a clear plan of action. Through the development of a systematic and comprehensive plan of evaluation, a faculty member will have the resources with which to put together a well-rounded representation of his or her teaching competence. The stakeholders for teaching excellence are diverse, ranging from students to practitioners to departmental review committees and chairs, therefore an inclusive evaluation plan can help meet the needs of these varied groups. It is also important to note that as evaluations continue to be used for important personnel decisions, including tenure and promotion as well as merit increases, the most relevant stakeholder may be the individual faculty member who is being evaluated. It would behoove the faculty member to ensure that valid and reliable measures of evaluation are being used for such decisions, a well thought out evaluation plan can both help ensure that the faculty member is treated fairly and more importantly, assist the individual faculty member in improving his or her teaching skills. As Boyer (1990) has emphasized, teaching is one of a faculty member's most important responsibilities, therefore an effective evaluation of this responsibility is necessary to ensure that it is being met.

References

Algozzine, B., Beattie, J., Bray, M., Flowers, C., Gretes, J., Howley, L., Mohanty, G., & Spooner, K. (2004). Student evaluation of college teaching. *College Teaching, 52*(4), 134-141.

Appling, S. E., Naumann, P. L., & Berk, R. A. (2001). Using a faculty evaluation triad to achieve evidence-based teaching. *Nursing and health care perspectives, 22*, 247-251.

Aultman, L. P. (2006). An unexpected benefit of formative student evaluations. *College Teaching*, *54*(3), 251.

Babin, L. A., Shaffer, T. R., & Tomas, A. M. (2001). Teaching portfolios: Uses and development. *Journal of Marketing Education, 24*, 35-42.

Baldwin, T., & Blattner, N. (2003). Guarding against potential bias in student evaluations: What every faculty member needs to know. *College Teaching, 51*(1) 27-33.

Boyer, E. (1990). Scholarship reconsidered: Priorities of the professorate. Princeton, NJ: The Carnegie Foundation for the Advancement of Teaching.

Casey, J. R., Gentile, P., & Bigger, S. W. (1997). Teaching appraisal in higher education: an Australian perspective. *Higher Education*, *34,* 459-482.

French, D. P. (2006). What currency should we use? (Teaching evaluation). *Journal of College Science Teaching, 35*(4), 60-61.

Gastel, B. (1991). A menu of approaches for evaluating your teaching. *BioScience, 41,* 342-346.

Hativa, N., Barak, R., & Simhi, E. (2001). Exemplary university teachers: Knowledge and beliefs regarding effective teaching dimensions and strategies. *Journal of Higher Education, 72*, 699-729.

Heckert, T. M., Latier, A., Ringwald-Burton, A., & Drazen, C. (2006). Relations among student effort, perceived class difficulty appropriateness, and student evaluations of teaching: Is it possible to "buy" better evaluations through lenient grading? *College Student Journal, 40*(3), 588-596.

Heckert, T.M., Latier, A., Ringwald, A., & Silvey, B. (2006). Relation of course, instructor, and student characteristics to dimensions of student ratings of teaching effectiveness. *College Student Journal*, *40*(1), 195-203.

Hodges, L. C., & Stanton, K. (2007). Translating comments on student evaluations into the language of learning. *Innovation in Higher Education*, *31*, 279-286.

Hutchings, P. (1994, November). Peer review of teaching. *AAHE Bulletin.* (Available from the American Association for Higher Education, One Dupont Circle, Suite 360, Washington, DC, 20036).

Knapper, C. (2001). Broadening our approach to teaching evaluation. *New Directions for Teaching and Learning, 88,* 3-9.

Macalpine, M. (2001). An attempt to evaluate teaching quality: One department's story. *Assessment and Evaluation in Higher Education, 26*, 563-578.

Martinson, D. L. (2000). Student evaluations of teaching and their short term validity: Looking for dialog with impacted parties. *Journalism & Mass Communication Educator, 54*(4), 77-82.

Martinson, D. L. (2004). A perhaps “politically incorrect” solution to the very real problem of grade inflation. *College Teaching, 52*(2), 47-51.

Mento, A. J., & Giampetro-Meyer, A. (2000). Peer observation of teaching as a true developmental opportunity. *College Teaching, 48*(1), 28-32.

Mills, M., & Hyle, A. E. (1999). Faculty evaluation: A prickly pair. *Higher Education, 38,* 351-371.

Ory, J. C. (2000). Teaching evaluation: Past, present, and future. *New Directions for Teaching and Learning, 83,* 13-18.

Panici, D. A. (1999). Methods of assessing teaching: Investigating the how and why. *Journalism & Mass Communication Educator, 54*(2), 61-72.

Polya, G. (1957). *How to solve it.* New York: Doubleday.

Pratt, D. D. (1997). Reconceptualizing the evaluation of teaching in higher education. *Higher Education, 34,* 23-44.

Reece, S. M., Pearce, C. W., Melillo, K. D., & Beaudry, M. (2001). The faculty portfolio: Documenting the scholarship of teaching. *Journal of Professional Nursing, 17,* 180-186.

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

Steiner, S., Holley, L.C., Gerdes, K., & Campbell, H. E. (2006). Evaluating teaching: listening to students while acknowledging bias. *Journal of Social Work Education, 42*(2), 355-377.

Yon, M., Burnap, C., & Kahut, G. (2002). Evidence of effective teaching: Perceptions of peer reviewers. *College Teaching, 50*(3), 104-110.

WHAT	WHO	WHEN	HOW	EVIDENCE
Imparting Knowledge	Students Peers National Exams Content Experts Class Evaluation	On-going On-going After Class Before Class	Class Evaluation Class Evaluation Score Reports Mail	Forms/Summaries Written Reports Test Documentation Written Report

Figure 1: Teaching Evaluation Matrix for Decision-Making