Nurturing and Evaluating Effective Teaching at UMBC

Alan T. Sherman
Computer Science Department
Member, SCEQ Committee
sherman@cs.umbc.edu

February 12, 1994

1 Introduction

There are, or ought to be, three reasons for evaluating teaching at UMBC. First, teaching evaluations can provide helpful feedback as part of a larger program to improve teaching effectiveness. Second, teaching evaluations are a component of faculty evaluations for tenure, promotion, retention, and merit raise decisions. Third, teaching evaluations provide one way in which UMBC can document, e.g. to the State Legislature, that it is carrying out its educational mission effectively.

In its teaching evaluations, UMBC relies heavily on the convenient, uniform, and quantitative *SCEQ (Student Course Evaluation Questionnaire)*. Unfortunately, the SCEQ measures teaching effectiveness only through a narrow dimension of student reaction; the accuracy of data in SCEQ reports is not completely reliable; and SCEQ results are inadequately analyzed and reported. UMBC does not have a comprehensive scheme for evaluating teaching effectiveness that goes beyond student reactions. Moreover, UMBC does little with the first and most important purpose of teaching evaluations—to improve teaching effectiveness.

I am writing this memorandum as a member of the *Ad Hoc Task Force Committee to Study the SCEQ (SCEQ Committee)* and as an educator who cares deeply about quality instruction. My views, however, do not necessarily reflect the opinions of other Committee members.

As I understand it, the purpose of the SCEQ Committee is to examine the SCEQ and to make recommendations concerning its use and improvement. More generally, I interpret the role of the Committee to examine teaching evaluation at UMBC and to make appropriate constructive recommendations for improving teaching effectiveness.

To make such recommendations, and to disentangle the web of complex issues surrounding teaching evaluations, it is helpful to review the current practice of teaching evaluation at UMBC, to understand what effective teaching means, to identify strategies for evaluating effective teaching, and to study means by which the University can support, encourage, and enhance effective teaching.

2 Teaching Evaluation at UMBC: Current Practice

Every instructor at UMBC undergoes a series of teaching evaluations, the most crucial of which is the tenure review. At each evaluation, the SCEQ plays a significant role. This section summarizes these evaluation events and their procedures. In addition, this section explains what the SCEQ is and summarizes its roles and shortcomings.

2.1 Evaluation Events and Procedures

Whenever a faculty member is evaluated for the purpose of tenure, promotion, retention, or merit raise, the evaluation includes an evaluation of teaching effectiveness. The nature of these evaluations varies depending on the event and on the department through which the evaluations are carried out. In addition to these formal evaluations, many instructors engage in informal self-evaluations and in dialogues with their students, colleagues, and department chairs.

The Faculty Handbook describes procedures for all major evaluations. In addition, each department has its own guidelines and practices. My impression is that there is considerable variation between departments in what actually takes place. In addition, the key players—especially the department chair and chair of the evaluation committee (if any)—have considerable flexibility in shaping the structure of the evaluation.

In the rest of this section I will summarize the procedures for evaluating teaching effectiveness that are used in the Computer Science Department during tenure reviews; I understand other departments follow similar procedures. Merit raise evaluations and first-year evaluations are much less involved and are typically carried out by the department chair, in consultation with senior faculty.

In the Computer Science Department, the teaching evaluation component of tenure review works as follows. First, the chair of the Tenure Review Committee collects the following data: (1) Statement from candidate. (2) SCEQ Reports for at least the last three years. (3) Statements from at least two students—typically one undergraduate and one graduate student. These students collect and interpret input from other students and present their findings to the Committee to provide the student's point of view. These students need not have had the candidate in any course. In addition, the candidate may submit additional evidence, such as course materials and Blue Sheets (see Section 2.2). Although not formally part of the evidence, individual impressions and scuttlebutt absorbed over time often influences committee members. The Committee, which includes two students, examines and discusses the evidence. A secret yes/no vote is taken and a report is written. The Department Chair writes a separate letter. Also, private discussions take place between the Department Chair and various administrators. The candidate is eventually given a summary tally of the Committee vote, the Committee report, and the Chair's letter.

Although the nature and duration of evaluation discussions varies with the details of each case, typically, the Committee spends a total of approximately four hours evaluating the teaching effectiveness of each candidate, representing approximately 20% of the total evaluation time. It is common for the Committee quickly to reach a consensus that a candidate's teaching performance is fine. It is uncommon for the Committee to make a big deal about lackluster teaching or even to make significant mention of dramatic differences in teaching quality. The same cannot be said about evaluations of publication records. Yet, it is my impression that there have been dramatic differences in the teaching effectiveness of recent candidates.

2.2 The SCEQ and Blue Sheets

The *SCEQ (Student Course Evaluation Questionnaire)* is a one-page multiple-choice questionnaire that is filled out by students [2, 4]. Its stated purpose is "to provide measures of an instructor's teaching effectiveness."[1] For every UMBC course, it is supposed to be passed out during class by the instructor on an unannounced day during the last two weeks of the semester. A student in the class is supposed to collect the completed questionnaires and return them to the Administration. The machine-readable responses are scanned and analyzed. Many weeks later, a report is mailed to each instructor and to department chairs and administrators. The SCEQ was created in 1971

[1] Instructions for the SCEQ [2].

and substantially revised in 1976 [4]; it is administered through the Office of Institutional Research, under the direction of Nancy L. Ochsner. Wang and Schumann [4] and Nagaraj, Nabendu, Krass, and Shen [1] have statistically analyzed SCEQ data.

A separate questionnaire for free responses is also distributed along with the SCEQ. Known as "Blue Sheets" because they are printed on blue paper, these questionnaires are returned to the instructor after grades have been issued. These sheets are intended to give the instructor feedback and are not intended to be used by anyone other than the instructor. In particular, it is my understanding that the department tenure review committee may not demand to see the blue sheets and that the blue sheets may not be included in tenure-case documents beyond the department level. The rationale for these restrictions is to encourage student candor and to reduce the volume of tenure-case documents sent beyond the department.

2.3 Roles of the SCEQ

The SCEQ was created as a uniform, quantitative measurement of student impression of teaching effectiveness. It is extensively used in faculty evaluations and in documenting UMBC's teaching effectiveness. Although little seems to be done with using the SCEQ as a part of teacher improvement, it does provide data from which instructors could analyze and improve their teaching. Similarly, department chairs can use SCEQ results to identify problems and to offer constructive feedback.

The main advantages of the SCEQ are that is a simple, uniform, quantitative tool. It is simple in the sense that students fill out a one-page multiple-choice questionnaire that is automatically scanned. Also, results can be quickly interpreted—especially if the evaluator cares only how the numbers fall around the mean values. The SCEQ is uniform in the sense that all instructors are evaluated with the same questionnaire. Thus, each instructor can be compared against all other instructors. It is quantitative in the sense that student responses are integers in the range 1 to 5. Another feature of the SCEQ is that there is a mechanism in place for administering the SCEQ and for analyzing and printing its results. In addition, the State accepts quantitative SCEQ results as proof of UMBC's teaching effectiveness.

2.4 Problems with the SCEQ

There are numerous problems with the SCEQ. It is only one narrow measurement of teaching effectiveness that neglects many important aspects of teaching. SCEQ data say nothing about the course content. Moreover, these data are not completely reliable. For example, "instructor ranks" are not ranks but pseudorandom approximate ranks [3]. Its results are neither analyzed nor presented in statistically valid ways. For example, the reports include insignificant decimal places and exclude standard deviations. As a consequence, the current reports encourage misuse of the quantitative data. Furthermore, the analysis is not as informative as it could and ought to be. In particular, no trend analysis is performed—e.g. responses from same course over several years, and from the same instructor over several years. In addition, neither its construction nor its layout is done in a scientific way. For example, some students miss the key question on overall teaching effectiveness because this question is hidden in the questionnaire's confusing layout. Also, some useful factual data (e.g. hours spent on course) is not collected. Another problem is inflexibility: there is no mechanism through which instructors or departments can add individual questions that deal with special teaching considerations of their field. Most seriously, the SCEQ is not linked with any comprehensive system of encouraging, improving, and developing effective teaching.

These drawbacks of the SCEQ reflect deficiencies with its construction and analysis. More significantly, they place a disproportionately strong weight on student opinion, and reflect narrow views of teaching effectiveness and of the role of teaching evaluations.

As a tool for evaluating an entire course, the SCEQ needs to be administered at the end of the course. By then, however, it is too late for the instructor to change in response to student comments.

In addition, administering the SCEQ displaces approximately twenty minutes of valuable class time.

3 Effective Teaching

Before one can meaningfully suggest ways to improve teaching and teaching evaluations, one must first understand what is effective teaching. Effective teaching is a compound concept that makes sense only with respect to well-defined educational objectives and in the context of particular educational settings. To elucidate this concept, this section identifies the elements of teaching, the diversity of teaching, and some important (and sometimes conflicting) teaching objectives.

3.1 The Elements of Teaching

Teaching is an activity in which an instructor helps one or more students gain knowledge and skills. To accomplish this activity, the instructor must inspire the students, serve as a role model, offer feedback, and provide a structure in which students can learn efficiently. In addition, the instructor must identify the knowledge and skills to be taught. Some people think of this activity as taking place primarily with the instructor lecturing to the students, but there is no set format or style of teaching, and passive listening to lectures rarely develops skills or understanding. To succeed at teaching, the instructor must not only master the subject; the instructor must also understand human psychology, have a thirst for knowledge, be an excellent communicator, have good organizational skills, and base the instruction on a sound pedagogy. Which of these diverse skills is most important depends on the teaching situation and teaching objectives.

3.2 The Diversity of Teaching

There are many diverse teaching situations at UMBC. For example, teaching situations include lecturing to a large introductory calculus class, directing a Ph.D. dissertation, teaching a small junior-level discussion course on French literature, holding a chemistry laboratory, explaining a physics concept to a student in the hallway, supervising a master's project, organizing a summer research trip for students, running an honors seminar, supervising a physical fitness class, and advising a freshman who is undergoing a personal crisis. Moreover, students come from a variety cultures, backgrounds, and experiences and bring various outlooks, abilities, and goals. Instructors must deal with available materials, equipment, support staff, and classrooms, and they must balance varying teaching loads with their competing responsibilities in scholarship and and service.

Evaluators should recognize the many diverse types of teaching situations and evaluate instructors separately for their effectiveness in different situations. For example, some instructors may flourish in small classes but flounder in large lecture halls. It is not necessary for every instructor to excel at every teaching situation. Instead, UMBC needs a faculty that collectively can deliver a wide range of high-quality educational opportunities.

3.3 Perspectives on Effective Teaching

There are many different views of effective teaching. This section lists ten of these views which I have encountered in my teaching career. For dramatic effect I present each view in extreme.

1. *Behaviorist's view.* Set well-defined behavioral objectives and demonstrate high student performance, as measured by the high grades you assign them. Any understanding of the material is irrelevant.

2. *Businessman's view.* Maximize profit as defined by tuition revenue minus operating expenses. Seek large enrollments; grant many degrees; use many teaching assistants and part-time instructors; and consume as few instructor hours and other resources as possible. Course content is irrelevant.

3. *Content specialist's view.* Deliver a thorough, challenging, up-to-date, and insightful treatment of subject. Concentrate on the subject matter without any regard to how students react to it or to you.

4. *Elitist's view.* Concentrate on the extremely bright, highly-motivated, and well-prepared students without concern for the rest of the class. Assess your effectiveness by the performance of these elite students since they are more important than the others.

5. *Gardener's view.* Patiently support, encourage, guide, and nurture students to help them grow and achieve their potential. Concentrate on the weakest students who do not belong in the course. Document your success with displays of their work.

6. *Inspirationalist's view.* Inspire students to study your subject. Measure your success by the number of majors you attract and the number of students who continue their studies of your subject, regardless of how well they perform.

7. *Minimalist's view.* Minimize instructor time and effort subject to the constraint of retaining your job—an effective strategy for maximizing number of refereed publications.

8. *Popularist's view.* Attain high student approval ratings, especially on quantitative student questionnaires. Do not challenge students, but let them think that they are working hard and learning a lot. Tell lots of amusing stories in class and make students feel good. Give lots of A's and B's but rarely C's, D's, and F's.

9. *Researcher's view.* Generate Ph.D's, and publish research papers with students. This emphasis is appropriate since graduate programs are more important than undergraduate programs and since good researchers are necessarily good teachers.

10. *Steady sailor's view.* Don't rock the boat. Do not make changes and do not cause trouble. Above all, generate no complaints.

Taken to extreme, each of the foregoing perspectives is absurd. Yet there is also merit in each view when not taken to extreme. For example, instructors and administrators should be sensitive to complaints, which sometime serve as a barometer of teaching problems. But it would be unreasonable never to expect any complaint. For instance, a fundamentalist student might unreasonably complain about an instructor for discussing evidence of Darwin's theory of evolution. Likewise, although healthy enrollments are crucial to the University's existence, instructors who maintain high standards will typically attract fewer students than those with lower standards. Similarly, Ph.D. production might reasonably be the most important educational objective of a specialized research institute but irrelevant to a department that has no graduate programs. No single perspective is most appropriate for all teaching situations. Instructors and evaluators should recognize and balance these perspectives in the context of their educational and institutional settings.

4 Evaluating Teaching Effectiveness

Given the diversity of teaching situations and the variety of perspectives on effective teaching, it is important that teaching evaluations be sensitive to these situations and perspectives. Also, to ensure an accurate, comprehensive view of a candidate's teaching, it is important that a wide array of measurements be considered. This section discusses several available measurement options and strategies for interpreting their results. In addition, this section discusses criteria for selecting these options and strategies.

Some people say that teaching effectiveness is difficult to evaluate but take no action to deal with this perceived difficulty. Although there are challenges to evaluating teaching effectiveness, the problem is not intractable: the perceived intractability is primarily a reaction to the multitude of teaching situations and teaching objectives. Given any situation and objective, often it is possible to measure (at least to some degree) the extent to which the objective has been achieved in that situation. Therefore, teaching evaluations must not be reduced to a single overall number but must address the gamut of educational challenges and objectives at UMBC. In particular, quantitative measurements must be carried out and interpreted in a way that is sensitive to these diverse challenges and objectives.

4.1 Criteria for Evaluating the Evaluation Strategies

Before considering evaluation particular strategies, it is important to examine criteria for selecting these strategies. To begin, these strategies ought to be accurate, informative, helpful, and fair. In addition, the cost of administering these strategies and of interpreting their results ought to be reasonable in relation to the nature of the evaluation.

4.2 Measurement Strategies

There are many useful strategies for measuring aspects of effective teaching. This section lists and discusses several of these strategies.

4.2.1 Objective Data

A variety of objective, descriptive data can be easily collected from the registrar and from questionnaires filled out by students and by the instructor. These data are useful in defining the teaching situation and in interpreting other data. They can be clearly summarized in the evaluation reports. These data include the following:

1. *Course data:* name, number, level, and type of course. Number of students enrolled in course (at each of several dates, including initial and final enrollments). Distribution of types of students in course (year, major, degree status). List of other instructors, if course is multi-sectioned or team taught.

2. *Course work:* number of class meetings delivered/attended, number of exams given/taken, number of homework assigned/handed in, number of pages of required reading assigned/read, total number of hours expected/spent on course. Separately measure student and instructor perspectives since they may differ; for example, some students may not hand in all homework.

3. *Student data:* distribution of grades assigned/expected.

4.2.2 Student Reactions

Student reactions to instructors provide valuable data. There are three main ways to measure student reactions.

1. *Quantitative questionnaires,* like the SCEQ. The main advantage of such questionnaires is that they offer a simple comparison among instructors. In particular, results can be statistically analyzed and easily visualized. The main disadvantage is that they measure a narrow aspect of teaching effectiveness (mainly, student contentment) and that their numerical results can be misleading.

2. *Free-response questionnaires,* like the Blue Sheets. These questionnaires contain a lot of useful data. The main difficulty is interpreting these data. Many instructors find it helpful to hand out their own additional free-response questionnaires part-way through the course so that they can change their teaching approach in response to student reactions. Although all measurements of student reactions tend to concentrate on student contentment, because the Blue Sheets contain more information, this limitation is less severe for the Blue Sheets than for the SCEQ.

3. *Solicited letters* from students several years after course. When students reflect on their educational experiences several years (say, five years) after graduation, they are free of immediate concerns of course work and grades and can explain how an instructor has affected their lives. Such letters provide an important perspective on teaching effectiveness. Problems include difficulties in locating students and bias in the selection of students solicited.

4.2.3 Classroom Observations

Competent classroom observers can gain useful insights into an instructor's presentation skills, classroom atmosphere, and interactions with students. Observations also provide a good opportunity for discussing teaching issues. Although the classroom experience is an important aspect of teaching, it is only one of several important aspects. An awkward aspect of classroom observations occurs when colleagues observe each other since some faculty members do not appreciate criticism by their peers. Therefore, consideration should be given to using neutral observers. On the other hand, when the Computer Science faculty routinely observed each other, the observations unexpectedly helped the faculty to get to know each other.

4.2.4 Course Materials—The Portfolio

Detailed course materials provide a wealth of information about an instructor's approach to teaching. These materials should not simply be limited to syllabi but should include all major course handouts, including all assignments and examinations. These materials reveal the amount of thought, care, and original research that goes into effective teaching. For example, from the materials, evaluators could determine much about the appropriateness and intellectual level of the course and if the instructor is updating the course in response to recent advances. One drawback of examining course materials is that it adds more work on the evaluators.

4.2.5 Student Performance

The instructor's effect on student performance is a crucial aspect of teaching effectiveness. As such, it is very helpful to measure student performance. For example, an instructor could show samples of student work, invite external examiners to evaluate the students, administer standardized tests, and encourage students to participate in scholastic competitions (e.g. science fairs). Unfortunately, all of these measurements are highly problematic and their results must be interpreted against some comparison group.

4.2.6 Instructor Input

To interpret student reactions and other data, it is useful to understand the instructor's perspective. This perspective can be gained through instructor comments written before, during, and after the course. In addition, such comments can be useful to the instructor as a means of self-evaluation. For example, for each course, the instructor might focus on ways in which the instructor can improve his or her teaching effectiveness.

4.2.7 Peer Evaluations

Another valuable source of information is peer evaluations from other members of the department. These evaluations can examine all available data concerning an instructor's teaching effectiveness. Colleagues are often well-qualified to carry out such observations because they can observe an instructor's attitude and approach of teaching over many years and because they are experts in the general subject matter. On the other hand, colleagues may know little about the specialized course content and there may be significant personal bias and political conflict within a department.

4.3 Interpreting the Results

Interpreting the results of teaching measurements is a crucial aspect of the evaluation process. If the candidate is to improve, the candidate needs helpful feedback. For promotion and tenure reviews, recommendations and decisions must be made.

Typically, for major evaluations, results are interpreted by a department committee and by the department chair. This framework is reasonable and workable. It can break down, however, when the department is extremely small (in which case too few people participate in the process) and when few participants care about effective teaching.

For each objective, it would be desirable to measure what effect the instructor had on the students in comparison with what effect other instructors, or no instructor, would have had. Because carefully administered case-control or prospective studies would be expensive, teaching evaluations typically avoid such direct comparisons. Nevertheless, some steps can be taken toward such comparisons. In particular, for all measurements, comparisons can be made against other sections, against similar courses, and against the same course in previous semesters. To the extent that quantitative measures are used, their analyses should include these types of comparisons.

5 Improving Teaching Effectiveness at UMBC

Although many UMBC faculty members are outstanding educators and although the quality of instruction at UMBC is high, concrete steps can be taken to enhance the quality of instruction at UMBC even further. Specifically, I recommend that UMBC undertake the following four steps:

1. Create an atmosphere in which effective teaching is encouraged, nurtured, and rewarded.

2. Recruit and retain faculty members who care about teaching and excel at it.

3. Improve the teaching evaluation process so that: (a) Effective teaching is examined more closely and plays a more significant role in faculty evaluations, and (b) The evaluation process is intimately linked with a comprehensive plan for improving teaching effectiveness.

4. Establish a Center for Faculty Development that will serve as a resource center for excellence in research and teaching and that will play a role in teaching evaluations.

The rest of this section explains each of these suggestions.

5.1 Create an Atmosphere that Encourages Effective Teaching

Although many people on campus do care deeply about teaching, there are also many people who project an attitude that teaching—especially undergraduate teaching—is not very important and that it is much less important than research. Also, little recognition seems to be given to scholarship in teaching. This attitude is reflected in comments, faculty evaluations, and in how some faculty members approach their own teaching. Unfortunately, some people infer that a strong interest in education somehow demonstrates a lack of commitment and dedication to research.[2] This attitude, which is more prominent in some departments than in others, discourages effective teaching.

It is not the role of the SCEQ Committee to define what the balance of research, teaching, and service at UMBC ought to be. And this balance need not be the same in all departments or for all faculty members. To help UMBC fulfill its objectives, however, this balance ought to be made clear and explicit. For example, when I was hired, I was told that my professional responsibilities would be scholarship, effective teaching, and service and that these responsibilities would be weighted scholarship-40%, teaching-40%, and service-20%. In the next sentence, however, I was also told

[2] For example, I fear that some people will think less of me for spending time to write this memorandum and for expressing some of the views therein.

that publishing papers in refereed journals and obtaining external funding are most important. Which is it? And if we believe that research is more important than teaching, then why do we not say, for instance, that the weights are scholarship-80%, teaching-15%, and service-5%. To whatever extent teaching plays in this balance, the suggestions in this section apply.

The teaching atmosphere is created primarily by faculty evaluations for retention, promotion, tenure, and merit raise, by other incentives, and by the attitudes and examples of department chairs and senior faculty. Of these factors, tenure decisions carry the most influence.

To improve the atmosphere about teaching, a variety of steps can be taken, including the following:

1. *Encourage department chairs to value effective teaching, and appoint chairs who value effective teaching.*

2. *Allocate all merit raises separately for scholarship, effective teaching, and service, with the percentage weights made explicit for each department or faculty member.* For example, in a department where teaching counts 25%, 25% of the merit raise money should be allocated for effective teaching. Allocating merit raises in this way would help ensure that appropriate attention and value be placed on scholarship, effective teaching, and service.

3. *Offer teaching reductions, research reductions, and summer support for instructors who undertake special teaching projects.*

4. *Recognize and reward effective educators with teaching awards and special titles.* For example, recognize about 5% of the faculty with the additional title of "Outstanding Maryland Educator."

5.2 Recruit and Retain Outstanding Educators

The single most effective way to improve teaching on campus is to hire and retain faculty members who care about teaching and excel at it. It is easier to hire an excellent, motivated educator than to motivate and to improve the teaching effectiveness of an ineffectual teacher. Unfortunately, it is sometimes difficult to assess the teaching effectiveness of fresh Ph.D.'s. Hiring excellent educators cannot be done unless people who make the hiring decisions value teaching.

Ideally, we would like every UMBC faculty member to be an outstanding scholar, educator, and community member. A more realistic goal would be to recognize that UMBC needs a variety of types of researchers, teachers, and community members and to ensure that, collectively, each department has an appropriate mix of faculty members. This view would allow and encourage a department to hire some faculty members primarily for effective teaching. Also, some faculty members may wish to stress different activities at different stages of their careers.

With respect to retention, I have heard it said that every faculty must be a good researcher. Fine. But why can we not impose the same standard for effective teaching? The reality, however, seems to be that departments will often excuse lousy teaching for good researchers, but departments will seldom excuse tolerable research for outstanding teaching. If effective teaching is to be required of all faculty members, faculty evaluations must include an honest and thorough examination of teaching effectiveness.

5.3 Improve the Teaching-Evaluation Process

To improve the teaching-evaluation process at UMBC, UMBC should integrate teaching evaluation with teaching improvement, include in each teaching evaluation a variety of measurements of teaching effectiveness beyond the SCEQ, and fine-tune the SCEQ instrument and its analysis programs. This section discusses these three recommendations, which aim to improve teaching effectiveness.

5.3.1 Link Evaluation with Improvement

Teaching evaluations provide an opportunity for instructors to improve their teaching effectiveness. But if this improvement is to take place, the evaluations must go beyond the numerical reporting of student reactions. Moreover, the evaluations must involve more feedback and discussion among instructors, department chairs, and teaching mentors. In addition, the University ought to provide more resources—including workshops, reading materials, and consultations with master teachers—to help faculty members improve their teaching skills. Some faculty, however, would not take advantage of these resources unless poperly motivated.

Each assistant professor ought to meet with the department chair at least once a year to discuss research, teaching, and service. For the first year or two, these discussions ought to take place at least once a semester. During these discussions, the chair can clarify University and department goals, identify how the faculty member contributes to these goals, review the faculty member's performance, and suggest concrete areas to work on. These discussions should be constructive and substantive, yet relatively simple and informal; they can be continued informally at appropriate moments throughout the year.

There is also value in encouraging informal, confidential discussions between instructors and master teachers from other departments. By keeping the discussions confidential and nonjudgemental, instructors might feel more receptive to constructive criticism and self-improvement. The "teaching table" at the Faculty Dining Hall was one positive step in this direction, but one-on-one private discussions would be even more effective.

Instructors should be encouraged to engage in more self-evaluation. For example, instructors should be encouraged to hand out informal midterm course evaluations and to discuss the results with a master teacher. Instructors should also be encouraged to invite observers into their classrooms and to watch video-recordings of their own classroom performances.

In addition, the University should provide more resources, including an extensive collection of reading lists and materials, videos, and workshops on effective teaching. For example, more could be done with orientation workshops for new instructors and with continuing workshops and lectures for instructors of all levels. There should be at least one workshop every semester.

5.3.2 Look Beyond the SCEQ

Teaching evaluations—especially the formal evaluations in tenure reviews—ought to examine a wider range of measurements of teaching effectiveness than they currently do (see Section 4.2). Specifically, retention, tenure, promotion reviews of teaching effectiveness at the department level should be expanded to include each of the following mandatory measurements in addition to the SCEQ reports:[3]

[3] Most Departments already include some but not all of points 1 and 2.

1. *List of courses taught, together with objective data including enrollments and complete grade distributions.*

2. *Statement from candidate evaluating his or her teaching, and describing what steps the candidate has taken to improve his or her teaching effectiveness.*

3. *Blue Sheets from all courses.*

4. *Detailed course materials (not just syllabi) from all courses.*

5. *Letters from students, both present and past.* Letters from students several years after graduation should be included whenever possible.

6. *Classroom observation reports.*

7. *Peer evaluations, in the form of letters and free-response questionnaires from department members and other colleagues.*

Of these measurements, the Blue Sheets and course materials provide the core primary evidence. For bottom-line judgements, I consider most reliable the peer evaluations from those faculty members who are informed and objective and whose opinions on teaching effectiveness I respect highly.

Inclusion of the Blue Sheets would require a change in the currently stated function of the Blue Sheets. I recommend such change. In addition, each instructor should be encouraged to use informal midterm evaluations for direct feedback.

Expanding the measurement strategies beyond the SCEQ will create some additional work, but the extra burden is reasonable at the department level. Also, it should be noted that, due to the diversity of teaching and the numerous perspectives on teaching effectiveness, there is great subjectivity in evaluations of teaching effectiveness. Consequently, there is great potential for abuse (e.g. trying to hurt (or help) someone by framing a case of bad (or good) teaching). In addition, there is a danger of creating divisiveness within a department, especially when faculty members observe and evaluate each other. Nevertheless, adding the aforementioned measurements ought to enhance the accuracy, objectivity, and descriptive power of teaching evaluations.

Other current procedural aspects of the review process seem reasonable. In particular, it is reasonable for the most careful review of teaching effectiveness to be carried out by a departmental committee. However, I favor mandatory inclusion of an external committee member (from outside the department), to help ensure objectivity and balance. This external person could file a separate written report with a separately tallyed vote.

In tenure reviews, separate votes on scholarship, teaching, and service are crucial to help enforce the appropriate balance of these considerations. This practice should continue. To the extent to which teaching effectiveness counts, an honest and thorough evaluation of teaching effectiveness ought to be made. Thus, if UMBC wishes to promote excellent researchers who are mediocre teachers, it should at least do so openly and honestly. Separate votes help ensure the integrity of the review process.

Some people have suggested the idea of taking ternary votes (no/yes/emphatically-yes) rather than binary votes (no/yes) in tenure reviews. I see little advantage in doing so given that a descriptive letter accompanies the vote.

5.3.3 Fine-Tune the SCEQ

There are several ways in which the SCEQ and, more importantly, the analysis and presentation of its results can and ought to be improved. This section briefly outlines some of these ways. In principle, it ought to be relatively simple to implement these changes; unfortunately, given the horrendous state of the SCEQ software, considerable effort will likely be required.

1. *The reliability of the data must be ensured.* The current errors and uncertainties in SCEQ data are unacceptable, especially when SCEQ data are used in crucial personnel decisions. For example, instructor ranks should be correctly computed.

2. *The analysis, presentation, and interpretation of SCEQ data must be done in mathematically valid ways.* For example, insignificant decimal places should not be printed or used. Also, since every numerical score (e.g. overall teaching effectiveness) is interpreted with respect to the distribution of such scores in various comparison groups (e.g. section, dept, level, school, University), for each question and comparison group, a detailed description of the distribution of the score for the group ought to be given. This description should include a detailed histogram, together with its range, median, and mean. Here, the histogram is especially important since scores tend to be asymmetrically tightly clustered around the mean.

3. *SCEQ reports should include more analysis of scores in comparison with scores from previous years from the same and similar courses.*

4. *The questionnaire should include a few special questions to be set by the instructor, department, and school.* Doing so would enable the SCEQ to become more flexible with respect to the diversity of teaching situations.

5. *The questionnaire and its report should include more objective data (see Section 4.2.1).*

6. *The selection, wording, and layout of the questionnaire should be improved.* For example, questions should be worded and layed out in a uniform fashion.

5.4 Establish Centers for Faculty Development and Educational Technologies

Finally, I suggest that two separate centers be created at UMBC—a *Center for Faculty Development* and an inter-departmental *Center for Educational Technologies.* The Center for Faculty Development would provide a variety of support services for researchers and educators; the Center for Educational Technologies would be a research and support organization for incorporating emerging educational technologies into UMBC courses. These centers would be separate and have a cooperative relationship.

5.4.1 Center for Faculty Development

The Center for Faculty Development would serve the following three functions in support of research and teaching excellence:

1. To provide resources for improving and nurturing effective teaching.

2. To provide resources for improving and nurturing scholarship and grantsmanship.

3. To provide resources in support of community service, including coordinating UMBC's participation in variety of teaching activities and organizations.

The Center could provide a variety of practical resources, including reading materials on effective teaching, video-taping services, and consultations with teaching experts. In addition, the Center could run workshops for teaching assistants, Ph.D. students, and instructors of all levels, and the Center could host educational conferences. For example, winners of the annual Presidential Teaching Award could be invited to give lectures and workshops on teaching. Other workshops could be held on effective time management and on grantsmanship. The Center could also coordinate all teaching evaluations.

As a part of the Baltimore community, the Center could help coordinate educational activities between UMBC and other educational organizations, including other schools in the Maryland system, Baltimore schools, and Teach for America. For example, the Center could run special summer programs for inner-city youth and the Center could serve as a center of expertise for educating teachers in the use of educational computing.

5.4.2 Center for Educational Technologies

The Center for Educational Technologies would serve the following three functions:

1. To facilitate educational computing at UMBC.

2. To generate new educational software.

3. To carry out research, including hands-on experiments, on using new technologies in all aspects of teaching in higher education.

The major activity of the Center for Educational Technologies would involve research, development, and application of computer technology to education. The Center could have a staff of experts in educational computing who could support UMBC instructors who use computers in their teaching. For example, an instructor might envision some educational software tool, and the Center could help the instructor realize that tool. The Center could spawn companies and have a close relationship with them to develop high-tech educational products—especially innovative educational software. The creation of such software could provide jobs for our students.

Such a Center for Educational Technologies would have a synergetic relationship with the University. The University would benefit from resources provided by the Center, and the Center would benefit by using the University as a laboratory. Moreover, the center might be able to attract large external grants, and there is commercial potential for its products (especially in the area of educational software). In turn, the center could attract outstanding researchers who would enhance the reputation of UMBC.

There is a danger that a large, heavily-funded Center for Educational Technologies might overshadow the primary activities of a smaller Center for Faculty Development. Therefore, I recommend that these centers be kept separate.

6 Conclusion

The SCEQ—that is, the questionnaire—performs reasonably well at its intended job of providing a uniform, quantitative, measurement of student reactions to teaching. It is a simple and convenient tool for documenting teaching effectiveness and for justifying tenure-review decisions. The problem is that numerical SCEQ reports say little about teaching effectiveness and that the SCEQ is not linked with a comprehensive plan for improving teaching effectiveness.

In this memorandum, I have suggested several steps for improving teaching effectiveness and teaching evaluation at UMBC. These steps include integrating teaching evaluation with teaching improvement, expanding teaching evaluations beyond the SCEQ, creating an atmosphere that supports and encourages effective teaching, recruiting and retaining excellent teachers, fine-tuning the SCEQ and its analysis, and creating a Center for Faculty Development. None of these suggestions will help, however, unless the instructor, department, and University care about teaching effectiveness.

In light of the many problems with the SCEQ, some people have advocated discontinuing its use. Doing so would be a mistaken overreaction. Despite its problems, the SCEQ has a role, and its simple, uniform, quantitative nature ought to be preserved. It analysis and use, however, ought to be improved, and teaching evaluations need to go beyond the SCEQ.

Throughout I have focused on teaching evaluation of regular (full-time) faculty members. In many departments, however, a large percent of undergraduate instruction is provided by adjunct (part-time) faculty and by graduate students. Given this reality, additional thought ought to be given to special issues of evaluating teaching effectiveness of graduate students and adjunct faculty.

References

[1] Nagaraj, Neerchal; Nabendu Pal; Katherine Krass; and Wei-Hsiung Shen, "The SCEQ project," Department of Mathematics and Statistics, UMBC (November 28, 1988), unpublished document. 68 pages.

[2] Office of Institutional Research, "UMBC Student Course Evaluation Questionnaire (SCEQ) Handbook," UMBC (September 1993), unpublished document. 24 pages.

[3] Sherman, Alan T., letter to Bruce Walz (April 21, 1992). 2 pages.

[4] Wang, Marilyn D.; and Barbara C. Schumann, "Student course evaluation guide," UMBC (October 1980), unpublished document. 49 pages.

Acknowledgments

I thank Timothy Finin and James Sasaki for helpful comments.

Contents