

TASK FORCE ON THE ASSESSMENT OF TEACHING AND LEARNING

On July 27, 2006, the Faculty Senate Executive Committee (FacEx) established a Task Force to consider how faculty and other teaching personnel at the University are evaluated for their teaching and for their related scholarship. FacEx determined that one important outcome of the committee's work would be a proposal for a set of blueprints or guidelines intended to aid academic units on campus when they evaluate teaching and student learning. The Task Force was asked to consider the following general issues:

- Identify the components of teaching and teaching scholarship that are valued and appropriate for consideration during this evaluative process.
- Determine who can appropriately contribute to the evaluation of each of these individual teaching components.
- Determine what evidence a teacher may present regarding his/her teaching effectiveness and teaching scholarship.
- Review comparable policies and procedures at peer institutions.
- Identify sustainable policies to ensure academic units document and regularly assess procedures for evaluating teaching effectiveness.

TASK FORCE RECOMMENDATIONS

The Assessment of Teaching

Teaching is an essential and integral component of universities' missions, and all members of the university community contribute to successful student learning. In contemporary institutions, opportunities for participating in the educational process continue to expand, and discovering how to acknowledge and reward the traditional and emerging elements that result in successful learning is complex. We approached the charge of assessing teaching and learning with the goal of providing opportunities to share the great diversity of possibilities for making positive contributions to the educational mission of the university. The perspectives of faculty colleagues, of students, and of the individual teacher can all inform the evaluation of teaching, and all three should be included in any process. The primary evaluation should come from faculty colleagues who look directly at the intellectual aspects of teaching, while also considering the views of students and the teacher's own observations. The materials we recommend below provide the evidence that faculty colleagues could use in making judgments for annual review and for promotion and tenure. No single source of evidence or information should predominate in the evaluation, and the relative weighting of these elements may vary within individual units.

Teaching is intended to generate skills, knowledge, and understanding in the students enrolled, so the evaluation of teaching should include some review of the learning asked of and demonstrated by students.

The teaching methods and practices that prompt learning can also be reviewed by faculty colleagues and by students to inform the evaluation. Also very important is the faculty member's own account of how teaching has developed over time in response to the effectiveness of prior teaching experience. Ultimately, faculty colleagues will consider the overall impact of teaching, the development of teaching skills over time, and the perceptions of students to form an evaluation of teaching quality.

The task force believes that assessment of teaching performance and student learning should be based on multiple forms of evidence, and accordingly we identified four basic kinds of evidence that can be used to assess teaching and learning. First, there is a wide range of materials that are routine products of teaching a class and require minimal work from teachers to document. Second, the student voice can be captured with a brief questionnaire focusing on the characteristics of teaching that students are best able to judge. Third, a reflective statement from the teacher provides insight into the development of teaching over time, allowing colleagues to observe what the teacher has learned from teaching and put into practice. Finally, faculty colleagues can observe and review the full context of teaching activities, including both the nature of the students and the kinds of courses taught and their place in the unit's overall goals.

Recommendation 1: Allow faculty members to report a broad range of teaching activities as the basis for peer review.

Provided below (Appendix 1) is a comprehensive list of activities that we think may be considered in the evaluation of teaching at KU. We recognize that this list is extensive and we urge faculty members to view it as an opportunity to showcase accomplishments and special talents rather than see it as a set of new expectations. Our intent is to generate ways to highlight ongoing teaching activities; the prompts are reminders of the many ways that faculty members can contribute to student learning and to their department's collective teaching mission. Our extensive list should open doors to a world in which the conventional student rating will not be the dominant form of evidence. Review committees and faculty peers should not expect to see contributions in every category; they should view the list as a broad menu from which individuals may include those activities that are a part of their contribution.

Key elements of the list are questions that focus on four essential facets of teaching:

- How does this teacher conduct courses?
- How does this teacher prepare for courses?
- What teaching work has the faculty member done in addition to teaching courses?
- Has the faculty member made progress over time in development of teaching and/or shared teaching work with colleagues?

Within each question, faculty members are provided with various prompts about ways they could demonstrate their work in each area and types of evidence that could indicate their success. Appropriate evaluators of the evidence are listed, as well. We have taken into account the resolutions of the Board of Regents from 1995 and 1996 that set some parameters for this process. Those resolutions state that there should be multiple sources of information and evidence about teaching, and that student ratings will be only one piece of a larger view of teaching. Materials relating to student learning and peer evaluation of teaching should be considered; students' ratings should not be a substitute for professional judgment about the quality of practices, materials, and learning.

These are the materials that could form the basis of a faculty peer review of the intellectual work in teaching.

Recommendation 2: Student ratings of teaching should be concise and focus on those specific aspects of teaching that students know best.

We affirm that student ratings are an important part of the feedback that faculty members receive about their teaching, and we have identified those aspects of teaching about which students have an excellent and useful perspective. Our recommendations have to do both with the questions that should be asked and with how the evidence should be handled and reported.

Information in the ratings should address five general areas around teaching: delivery of instruction, the assessment of learning, availability of faculty members, achievement of goals and expectations of teaching, and general support for learning. We have consulted with KU faculty members whose field of research includes the design survey instruments, and they are conducting test surveys during the current (Spring 2007) semester to identify the best possible wording for a concise survey that would gather student perception of these five key dimensions. We presume their work will identify the particular wording that would be best, and we include some possibilities as an example in Appendix 2.

Regardless of the particular set of individual questions that emerge from the pilot testing of items, we offer the following general recommendations for student ratings of the KU faculty:

Student ratings should be gathered in a professional and respectful way according to common procedures that assure anonymity.

Consistent with Kansas Board of Regents resolutions, there should be no required form; there will be a recommended form that can be used directly or as a model for local development.

The number of questions asked should be kept to the minimum needed to adequately estimate the students' view of the five areas listed above.

Students should rate aspects of the teacher's performance that are clearly apparent to them; they should not be asked to make judgments that would require extensive extra investigation or professional knowledge.

The numerical data should be reported in a fashion that includes distributions of scores, not merely indicators of an average or mid-point.

The student ratings form should not ask for an overall summary judgment from students about either the instructor or the course. Such items actually report opinions outside the range of student expertise, and they may tend to be used as a substitute for a more complete analysis of the full range of teaching components. We strongly support the goal of providing a rich array of evidence about teaching; items that attempt to capture the student voice in a single number would overshadow students' comments on more appropriate issues.

The reports of numerical data should be put in a context of other teachers' ratings, but those comparison distributions of scores should reflect a restricted range of courses that are comparable in size, position in the curriculum, target student audience, student motivation, or other characteristics identified by departments.

The numerical reports should be given to faculty members in both paper and electronic formats that facilitate reporting for personnel matters without substantial processing or additional handling. Summary reports would show the range of student perception for the five general areas of teaching. The pilot study underway will offer an example of what this reporting sheet could look like.

Recommendation 3: Faculty members should collect open-ended student comments to guide improvement of teaching. Reporting these comments should remain optional, and if they are used in evaluations they should be reported systematically.

Many faculty members agree that a great deal can be learned from the observations offered by students in response to questions inviting commentary on features of instruction, and faculty members should be encouraged to seek open-ended commentary. It is especially valuable for faculty members to invite these comments during the semester at a time when they can still respond to the comments with adjustments in the course. Those adjustments are valuable to students, and faculty members can also include those changes as part of their teaching record.

We recommend continuing the policy jointly articulated by the Provost and by Faculty Governance that makes the use of such comments in personnel decisions optional, at the discretion of the faculty member or unit. Accordingly, any open-ended questions used by a faculty member or requested by an individual department should be on a separate sheet from the numerical rating items, so that they may be

collected and distributed independently of each other. Open-ended questions should be framed in ways that maximize the likelihood of constructive comments and specifically discourage irrelevant or disparaging comments about faculty members.

As a general rule, we recommend that these comments be used for the benefit of teachers working to improve their instruction. There may be times in the evaluation of an instructor when it would be useful to have such comments for the interpretation of ambiguous numerical information. In such cases, the faculty member should bring forward the entire sample of such comments to aid in interpretation. Additionally, an individual faculty member should always be free to bring the comments forward, as long as they are reported in a professional manner (see below).

Within KU there are distinct communities of evidence, in which knowledge is constructed in different ways and conclusions are evaluated based on varying criteria. For many in the humanities, as well as from other areas of campus, there is great value in close reading of text, done by an individual, using a combination of personal experience and professional conventions of reading. For people in the social and natural sciences, and many in professional areas, open ended verbal statements are data like any other, and the protocols and conventions of data analysis would be the preferred method of reading. The practices around using open ended student comments should be flexible enough that different local communities (such as departments) can adopt consensual practices that reflect their shared understanding of reading and evidence. This is especially true for the use of these comments as guides for continuous development of teaching.

When samples of comments are used in formal evaluation, the resulting judgment represents an estimate of the professional quality of a faculty member's teaching. In that context, people presume that the conclusions drawn are an accurate representation of the teaching in question and not an artifact of either the reading or the gathering of comments. For that reason, we recommend that professional standards of treating comments as data be followed whenever comments are included in an evaluative professional context beyond a department level. This practice would assure all faculty members that the information will inform the process in the fairest way possible.

The Task Force consulted with KU faculty members whose research includes the analysis of the content of comments. That group noted a series of careful, empirically justified steps that are required for such evidence to meet professional standards of validity. It would be impossible as a practical matter for review of teaching comments to meet the publication standards of the research community, but several general guidelines emerged from the conversation that we recommend be included in the handling of all open ended comments.

- The sample of comments should be as large as possible; small samples can be very misleading.

- Analysis should always include a record of non-responders. Summary descriptions place the responses in a context of the total possible population.
- The analysis begins with the framing of a specific question or questions to be answered by the comments.
- Analysis includes identification of categories of responses, as appropriate to the question being answered. Where possible these categories should be distinct from each other and include a full range of possible answers.
- Individual comments should be put into a single category when possible; if a comment is long or complex, it can be broken into components and counted as separate comments.
- The final product would report a frequency of comments in each category, always including the number of people who did not respond to the question.

The analysis produced will be only as good as the questions asked and the categories of responses that are identified. If those are well constructed and aligned with each other, the analysis can help answer the question posed.