

**University of Puget Sound**  
**Faculty Meeting Minutes**  
**October 25, 2004**

1. President Thomas called the meeting to order at 4:08 p.m. Forty voting members of the faculty were present by 4:20 p.m.
2. Minutes of the September 14, 2004 faculty meeting were approved as distributed.
3. In response to President Thomas's call for announcements, Associate Dean Bill Barry reminded us that Connections core course proposals are due to the Curriculum Committee on November 1, 2004. President Thomas announced that November 17 would be Puget Sound night at the Tacoma Art Museum, with presentations by four of our colleagues. He asked us to watch for an upcoming announcement about the event, which will require a ticket (free) for attendance.
4. President Thomas had no report other than to thank faculty who had participated in fall faculty weekend. He said he had heard many complimentary comments and expressions of high regard in response to the various faculty presentations.
5. Academic Vice President Kristine Bartanen (1) asked us to participate in the upcoming faculty survey conducted for us for the third time by the Higher Education Research Institute of UCLA; (2) reported that trustees at their recent retreat approved a financial model for working with the graduate programs, accepted the report that both the occupational therapy and the physical therapy programs had met or exceeded their enrollment benchmarks, and agreed that the three-year benchmark process was now concluded; and (3) reported that at the same retreat trustees approved a January 2005 start of construction for the new science center.
6. Faculty Senate Chair Bill Beardsley had no report.
7. President Thomas turned the chair over to Bill Beardsley who launched our discussion of student evaluation of teaching. Beardsley reported that the Faculty Senate had voted at its last meeting to begin anew the process of review of the course evaluation form. He said that unlike previous efforts carried on by committee, this time the Senate would work on this with the help of the full faculty. He introduced a four-page handout (copy attached to these minutes) containing conclusions and recommendations from a 1998 report of a faculty survey on student evaluation of teaching, and he asked Eric Orlin and Suzanne Holland to lead discussion.

Holland began by reading from the handout four questions the 1998 ad hoc committee had asked in response to survey results: (1) How certain are we as an institution that our teaching evaluation processes, and the decisions based on them, are contributing to the improvement of teaching and student learning? (2) Are we unwittingly discouraging faculty risk-taking

which will result in different, but potentially more effective teaching methods? (3) To what degree does our current system of teaching evaluation hinder our ability to attract and retain effective and innovative faculty? (4) Do our decision-making processes place such strong conclusive presumption on data from student evaluations that other valid evidence of teaching effectiveness is not given due weight, or even ignored? Orlin reported that he had collected teaching evaluation forms from other colleges that could help to inform the review.

Holland called on Keith Maxwell, the only remaining member of the 1998 committee that conducted the survey. Maxwell gave a little history, saying that the Faculty Senate established the ad hoc committee fall 1997. The committee's 1998 survey of faculty produced the recommendations on the handout we received today. He said these recommendations were discussed with the Faculty Senate and that the Senate forwarded them to the Professional Standards Committee (PSC), but that nothing had resulted from them except for attention to recommendation #1 to "open the black box."

Mark Jenkins asked why the recommendations had not been acted on. No one was able clearly to explain why the PSC had not acted on them, although Beardsley did say the PSC had on its own worked on a revised teaching evaluation form only to have it "shot down" by the Senate.

Ted Taranovski asked if the 1998 committee had looked at other aspects of the evaluation process in addition to the evaluation form itself. Maxwell responded that the committee dealt with the content of the form and the use of the information on the form, but not at anything more in terms of the larger evaluation process. Holland pointed out that the handout makes it clear that the committee found problematic the use of a single form for both feedback and appraisal purposes.

David Droge asked whether the claim being made was that the specific recommendations of the 1998 committee were ignored or that nothing at all was done with them. He thought there may have been some changes that resulted from the committee's work, even though the recommendations were not implemented. Beardsley agreed that this was in fact the case. David Tinsley agreed that not all the changes that were actually made to the form were minor. Priti Joshi asked if the question we are taking up shouldn't be, "are we happy with the teaching evaluation form now?" Holland agreed.

Lisa Wood said that the problem is an imbalance in the evaluation of teaching effectiveness that relies too heavily on student feedback. She said that the weight given to student perceptions of teaching effectiveness is too high; that errors can be made; that we don't have the resources to measure teaching effectiveness any other way; and that this is a disservice to faculty since their evaluations may not be accurate.

Droge added that it is hard to determine the extent to which the form is useful or not useful. Changing the form does not address the question of the extent to which the Faculty Advancement Committee (FAC) uses them.

Jenkins asked how this process would go forward. Beardsley responded that the Senate would “hash out” the information obtained from these broader discussions and would then take the information to the dean. (Later, in response to a similar question from Doug Cannon, Beardsley said it was the PSC’s responsibility to act on recommendations and then agreed with Cannon that the Senate and full faculty were ultimately responsible). Jenkins said that earlier this year he had emailed his own personal recommendations to Dean Bartanen, to the FAC, and to the PSC. He said he finds those recommendations reflected in the recommendations made by the 1998 ad hoc committee, and in addition he would require the student to indicate the expected course grade. Mott Greene said that expected student grade was struck from an earlier version of the form because students found it to be intimidating. He said the current evaluation form was the result of competing interests. He argued it was more important for us to concentrate on how the form is to be used and weighted, rather than on what specifically it says. He said we need to ensure these student evaluation data do not outweigh colleague letters, for example. He said we need rules for how to read it.

Joshi asked Greene what he meant when he said the form was the result of competing interests. Greene said that students, faculty, and administration all had concerns. Joshi responded she found it ironic that students were involved in shaping the form. Greene responded that it didn’t seem ironic at the time.

Taranovski said that the current form was approved by the Senate after much work in 1998 and he agreed with Greene that we should not spend a lot of time revising the form. He said that if we were to review all of the forms we have used over the years we would find that we have never been happy with any of them. He added that we should look at differences in evaluations across different types of courses, for example core courses and major courses. He argued that students evaluate professors differently in core courses than they do in other courses. Is it good to be popular in core courses but not in major courses, he asked?

Carolyn Weisz asked if there was any way of separating the feedback function from the appraisal function. Orlin responded that Macalester College doesn’t do any in-class teaching evaluations; that teaching evaluations for tenure are based on surveys of students already out of the course. Robin Foster said that social scientists would recommend that separate forms be used for each function.

Taranovski said that, although the form has over time become increasingly a form for the administrative evaluation of faculty, that was not the case in the old days when faculty used their own forms for feedback. He said that now he would hesitate to submit a feedback form to his own class because he couldn’t then submit the results to the Faculty Advancement Committee. Would he even be allowed to administer his own separate feedback form, he asked? Dean Bartanen responded that there is nothing to preclude this and that she actually encourages faculty to use self-designed feedback forms along the way in their classes.

Keith Ward, returning to a point made by Greene, argued that we must look at how the forms are read. Juli McGruder, a recent member of the FAC, reported that FAC members do weigh

rationally what students say about things students are equipped to evaluate. She said that what students say is not taken at face value and, furthermore, it is no less necessary to weigh what colleagues report and what candidates themselves say.

It was at this point that, in response to a question from Cannon, that Beardsley declared that the PSC had responsibility for moving forward with any recommendations for change. Cannon responded that it was his understanding that, as a “creature of the faculty,” actions of the PSC could be overruled by the full faculty, and Beardsley agreed. Cannon said that he was troubled by the practice of the FAC to require inclusion in the file of the copies of the student teaching evaluation forms (the yellow copies) that the faculty member had used in creating the file. What sense does this make, he asked, given that any divergences between the yellow copies and the white copies (which the FAC already has in hand) can only be minor. He said he understood that differences between the sets was treated as “evidence in a court of law,” implying that a comment or two could make a big difference in the evaluation. He said he thought this suggested something was seriously wrong.

Dean Bartanen responded that the FAC uses the white copies and that departments work with the yellow copies. She said that to her knowledge the FAC is not routinely looking at the differences between the white and yellow stacks of evaluation forms. Cannon asked why then it was necessary for the candidate to submit the yellow copies. McGruder said that the white copies are more readable than the yellow copies and she said it is sometimes necessary for the FAC to make sure the number of white and yellow copies equate.

Wood said she was concerned about how teaching evaluation forms over time can shape the teaching culture. She suggested that we should change the forms periodically just to shake the culture up a bit. Joshi agreed with Wood that the specific questions asked on the form do make a difference in faculty behavior and that the actual questions do create a culture. Wood added that this isn't all bad, but that over time students may be missing opportunities to learn in different ways.

Tinsley asked that we not rule out looking at the form and trying to improve it as we move forward. He said we could at the same time look at the process, but he suggested that we should give this generation of faculty the opportunity to make changes to the form.

ASUPS President Ryan Cunningham reported that students agree with many of the things being said today. He said students say the form is too long and that they get frustrated toward the end. He said the form may be too specific in what it asks for and that it doesn't always give students the chance to comment on what they want to evaluate about a professor. He added that students are unsure whether they're writing to the professor or to the person evaluating the professor and that knowing which it is could affect their answers.

Jenkins said that, at other colleges he's been at, teaching evaluation forms are administered the last day of the class, and that he was astounded to learn that here they can be given as early as the tenth week of class. He said we must revisit this to eliminate manipulating results by manipulating timing.

Dean Bartanen said that looking at the form is appropriate. She said that some other questions faculty have raised have to do with whether the form reflects the range of teaching on our campus now, for example team-taught or interdisciplinary courses. She suggested that we should talk first about what are the issues we are trying to address, and then determine whether review of the form addresses those issues.

Bill Breitenbach suggested that we consider collecting student evaluations on-line, with the proper protections. He said that students feel the onslaught of these forms at the end of the semester, that we could save class time if the forms could be administered electronically, and that we would no longer be at the mercy of just those students who happen to be in class the day it was administered.

Nancy Bristow suggested that collecting information on student experience with the forms could help us decide how to look at the forms.

Cunningham added that students are very much interested in teaching evaluations that can be reviewed by other students and that ASUPS has been exploring web-based options.

Taranovski said the form seems to be slanted toward certain teaching styles and that his main concern is with what exactly the form is measuring. He said that it is a conceit for us to assert that the forms measure teaching effectiveness at all. He said we haven't defined teaching effectiveness, and that this was a sign of the faculty's general resistance to any kind of outcomes assessment. He said the forms may measure the degree to which the faculty is liked, or perhaps whether a professor has a decent repertoire of teaching skills. But, he asked, do we really want to measure teaching effectiveness? If so, the forms may be no good.

Ward asked if we were agreeing that we should take a look at the form. The response by general consensus was yes.

Joshi asked if the teaching evaluation forms Orlin had collected could be put on line for faculty review. Orlin responded that he would have them put up on the Faculty Senate web site.

President Thomas had to depart the meeting at about 5:00 p.m. Dean Bartanen, a few minutes later, received the motion to adjourn.

We adjourned at 5:10 p.m.

Respectfully submitted,

John M. Finney  
Secretary of the Faculty

## STUDENT EVALUATION OF TEACHING AT UPS

### VALIDITY CONCLUSIONS

<u>Validity Type</u>	<u>Problem</u>
<i>face validity</i>	<p>Survey results reveal a disturbingly large proportion of the faculty who view the <u>process</u> negatively and with distrust. The evaluation <u>form</u> is viewed somewhat more positively than the process, but the difference is not significant.</p>
<i>construct validity; content validity; convergent validity; criterion validity</i>	<p>At UPS, each of these three validity-types is confounded by our attempt to use the same instrument for multiple purposes. There is extensive research literature to show that a single instrument is inappropriate for both feedback and appraisal purposes. Also, there are differences in the construct definitions of effective teaching in large vs. small classes, team-taught classes, and laboratory classes.</p> <p>The problem of feedback vs. appraisal is aggravated by the fact that immediately prior to completing the evaluation students are expressly told that the form will be used for both purposes; therefore, the responses on a set of evaluations for a particular class will likely be addressed to different audiences for different purposes, i.e., to faculty for feedback, and to the FAC, departments, and the administration for appraisal. Currently, there is no possible way to determine the purpose the student had for his or her responses.</p> <p>An additional problem is presented when the evaluation form asks questions about "teaching style." These questions are argued to lead to invalid stereotyping by virtue of exclusion or inclusion in a group, since good teaching may be associated with a variety of styles. That is, it is fallacious to argue that, "Professor X does not use a particular style of teaching, ergo, X is an ineffective teacher." Or conversely, that X's membership in a group which uses a particular style of teaching means that X is an effective teacher.</p>

*discriminant validity*

The research literature identifies a number of biases in student evaluations of teaching. These are:

- absolute and relative grades
- student motivation and prior interest
- workload

In fact, these biases have been shown to contribute from 16% to 35% of the variation in student evaluation ratings.

*empirical validity*

Our instrument does not collect adequate information to identify and adjust for these biases.

Lack of "norming"

*consequential (use) validity*

In the opinion of the Ad Hoc Committee, the survey results reveal a serious problem with the way a large proportion of our colleagues perceive the method we use to evaluate teaching effectiveness. The degree of negativity reflected in the scale-responses and the written comments is disturbing and we believe this represents a threat to the continuing quality of faculty at UPS.

Even those of us who do not share these negative views must not dismiss them, for they come from a group that is not a small vocal minority. It is a group of sufficient size and importance to the mission of the university that it constitutes a critical core of the faculty. Without their loyalty and commitment we will be unable to sustain the progress toward excellence.

The extent and the tone of the survey responses of our colleagues is a matter of deep concern to the Ad Hoc Committee. It has prompted us to ask these questions among many: How certain are we as an institution that our teaching evaluation processes, and the decisions based on them, are contributing to the improvement of teaching and student learning? Are we unwittingly discouraging faculty risk-taking which will result in different, but potentially more effective teaching methods? To what degree does our current system of teaching evaluation hinder our ability to attract and retain effective and innovative faculty? Do our decision-making processes place such a strong conclusive presumption on data from student evaluations that other valid evidence of teaching effectiveness is not given due weight, or even ignored?

<b><u>RECOMMENDATIONS</u></b>	<b><u>PROBLEMS ADDRESSED</u></b>
<p><i>(1) Open the "black box"! This means to assure the faculty and the students, in an openly demonstrable way, of the accuracy and fairness of the process employed in the interpretation and use of evaluation data.</i></p>	<p>Negativity and distrust of the process by faculty (face validity and consequential validity)</p>
<p><i>(2) We must design different forms for feedback and appraisal.</i></p> <p>[According to most research findings, using a single form for both feedback and appraisal is inappropriate. Since teaching improvement and teaching appraisal are, by definition, different constructs, construct validity cannot be assured with a single instrument.]</p>	<p>Dual purposes of the evaluation form. (construct and content validity)</p>
<p><i>(3) Students must be clearly instructed about the purpose of the particular form they are completing for a particular class.</i></p> <p>[Currently, students are advised that their responses are used for both feedback to the instructor <u>and</u> for appraisal of the instructors teaching effectiveness for purposes of tenure and promotion. As a result, student responses are ambiguous as to whether they are for feedback or appraisal.]</p>	<p>Dual purposes of the evaluation form. (construct and content validity)</p>
<p><i>(4) Questions on the form should take into account whether the class is small or large, team-taught, multidisciplinary, lab, etc.</i></p> <p>[The qualities of effective teaching will differ depending on the class type.]</p>	<p>Failure to recognize differing constructs of effective teaching for different types of classes. (construct and content validity)</p>
<p><i>(5) Questions on the evaluation form should not ask about a particular teaching style.</i></p>	<p>Stereotyping by teaching style (construct validity)</p>
<p><i>(6) The evaluation form should obtain information which assist in identifying and adjusting for the influences of students' grade expectations, motivation and prior interest, and workload. Each of these have been shown to have at least a moderate confounding effect evaluations. In order to have discriminant validity, the impact of these biases must be accounted for.</i></p>	<p>Contamination of evaluation information by external variables, or biases. (discriminant validity)</p>



<p><i>(7) Serious consideration should be given to establishing norms for-each criterion of effective teaching we measure. The norms should be determined for each of the course types, i.e., small, large, team-taught, etc.</i></p> <p><u>[This is not a recommendation for the use of a strict, or absolute, numerical standard which must be attained or exceeded to receive a positive decision. Rather, these norms should be used to give an instructor a basis for determining the quality of their own performance relative to their colleagues.]</u></p>	<p>Lack of "norming" (empirical validity)</p>
<p><i>(8) An external consultant should be retained to conduct a comprehensive validity study of the, student evaluation of teaching process including the design of the form, the manner in which the form is administered to students, and the process by which the information is interpreted and used in arriving at decisions.</i></p>	<p>General validity</p>