

UNIVERSITY OF WATERLOO

Review of the Faculty Annual Performance Evaluation Process

Submitted to:

The Provost

and

President, FAUW

Prepared by:

Working Group on Faculty
Annual Performance Evaluation

April 28, 2009

Table of Contents

Executive Summary	3
Key Recommendations	4
Introduction.....	5
Working Group Process.....	5
Findings and Recommendations	6
1. The Performance Evaluation Process.....	6
Structure.....	6
Performance Evaluation Committees	7
Expectations.....	7
Timing and Feedback.....	9
Rubrics and Dynamic Range.....	9
Communication	11
2. Teaching.....	11
3. Scholarship.....	13
4. Service.....	13
5. Related Issues	15
Conclusion.....	15

Executive Summary

In any large organization, the processes used to evaluate and provide feedback about the performance of the people working in it are extremely important. Carefully designed, these processes can help the organization thrive by providing appropriate rewards and encouragement for jobs well done, and guidance about how jobs can be done better. If the processes are poorly designed, however, the result can be lower morale, engagement and productivity.

The University of Waterloo has as its mission “the pursuit of learning through scholarship, teaching and research within a spirit of free enquiry and expression.” Both the University and the Faculty Association recognize the importance of a faculty performance evaluation process that is fair and that provides productive and appropriate incentives to faculty. As a result, it was agreed, as part of the Memorandum of Salary Settlement signed by the University and the Faculty Association on February 11, 2008, that the Provost, in consultation with the President of the Association, would set up a working group to review these processes.

The resulting working group on Faculty Annual Performance Evaluation reviewed the existing processes and systems that comprise and support the annual faculty performance evaluation. We invited feedback from every faculty member through email and paper announcements, considered this feedback as well as information from Chairs¹ and Deans on the existing processes, obtained further input through focus groups with representation from all Faculties, and took into account other pertinent research. Through this process, we identified recurring themes and developed recommendations for making changes to address problematic areas.

Our report is organized into the following sections:

- General procedural considerations;
- Evaluation of teaching;
- Evaluation of scholarship;
- Evaluation of service;
- Related topics that need further investigation.

For convenience, we have gathered the key recommendations from each section here. Where appropriate, we have provided a rationale for the recommendations in the body of the report. We have also tried, where it is not obvious, to indicate suggestions on how to implement the recommended change.

¹ The term Chair refers to both department Chairs and Directors of Schools and Programs.

Key Recommendations

THE PERFORMANCE EVALUATION PROCESS

- 1.1 Maintain the basic structure of the current performance evaluation process. Implement enhancements to strengthen the process, as recommended in this report.
- 1.2 Conduct the evaluations with a performance evaluation committee, advisory to the Chair.
- 1.3 Provide faculty members with a regularly updated document outlining the evaluation process and the performance expectations in their department² for scholarship, teaching and service.
- 1.6 Evaluate tenured faculty bi-annually, and evaluate tenure-track faculty annually, and provide detailed feedback.
- 1.8 Eliminate the rubrics attached to performance evaluation scores and increase the dynamic range of evaluation in all areas: scholarship, teaching and service.

TEACHING

- 2.1 Develop a common course evaluation instrument for use across campus, and establish a uniform method for collecting, analysing and disseminating the results.
- 2.2 Develop a document for Chairs and performance evaluation committees summarizing the best empirical evidence about how factors aside from quality of teaching affect student evaluations.

SCHOLARSHIP

- 3.1 Evaluate scholarship on the basis of a three year period, and evaluate teaching and service on the basis of work done since the previous evaluation.
- 3.2 Clarify that scholarly work outside of the usual peer reviewed venues is valued, but that the onus is on the faculty member to provide evidence of its quality, impact and relevance.

SERVICE

- 4.1 Establish internal service as an essential duty for faculty members.
- 4.4 For those in administrative roles that carry a stipend³, establish a portion of the stipend as an ongoing salary enhancement, at the end of the term, to be funded separately from the merit pool. Eliminate informal practices, such as automatic ratings of "2" or use of the Outstanding Performance Award in ways outside those specified in policy, for rewarding the carrying out of significant administrative duties.

RELATED ISSUES

- 5.1 Further investigate two related issues: a) whether the merit evaluation process has features that generate gender-based anomalies; b) how to use the flexibility of the merit evaluation process, in conjunction with other policies, to make UW a more family friendly employer.

² The term department refers to all academic units including schools.

³ Department Chairs or Directors, Associate Deans, Associate Chairs or Directors.

Introduction

The Vice President Academic and Provost, after consultation with the President of the Faculty Association, established the working group on Faculty Annual Performance Evaluation to:

- (1) Review arrangements for annual performance evaluations for faculty, with emphasis on identifying the most significant limitations and weaknesses, and to determine whether opportunities exist to strengthen the process.
- (2) Recommend changes to eliminate or minimize major limitations or weaknesses.

The working group was asked to provide a written report to the Vice President Academic and Provost and to the President of the Faculty Association. Members of the working group are: Adel Sedra, Dean of Engineering, and Chair of the working group, David DeVidi, Philosophy, and President of the FAUW, Beth Jewkes, Chair of the Department of Management Sciences, Elizabeth Meiering, Chemistry and Mary Jane Jennings, Institutional Analysis & Planning, support staff to the working group.

This report will be used as the basis for further discussion at the Faculty Relations Committee prior to completion of performance evaluations for 2009.

Working Group Process

The working group met about fifteen times between June, 2008 and April 2009. Its initial step was to send an email to all faculty members inviting feedback on items (1) and (2) above and comment on any element of the performance evaluation process. The working group received approximately forty responses.

Members invited feedback and consulted individually with several department Chairs from each Faculty, and with the Deans to gain an understanding of current practices and to receive feedback on process and areas for improvement.

The working group held three focus group meetings with invitees from two Faculties at each two hour session. David DeVidi, on behalf of the group, met with Donna Ellis, Centre for Teaching Excellence, to obtain guidance on how best to run the focus groups. Invitations were sent to a diverse group of selected faculty members with a goal of having seven to ten attendees at each session. The individuals selected represented large and small departments, all career-stages, as well as both male and female faculty. Based on the feedback received up to that time, the working group used the following items as talking points:

- Evaluation of teaching, both undergraduate and graduate (including supervision);
- Evaluation of service: how it is weighted and how it is valued;
- Evaluation of scholarship;
- The transparency of the process and the feedback it provides. Technical issues: the rubrics (“Outstanding” “Excellent” “Satisfactory”, the range of scores, frequency of evaluation, etc.);
- What did we not talk about that we should have?

The focus group sessions were held on August 27th (Arts and Environment), September 3rd (Applied Health Sciences and Science) and September 4th (Engineering and Math) and were well-attended. The working group received extensive feedback and began to organize it into common theme areas. Attendees were appreciative of the fact the working group had been struck and for the opportunity to provide feedback and suggestions on such an important process. The working group would like to thank those faculty members, Chairs, and Deans who provided important input to this process through emails, interviews and focus groups,

At one of its fall meetings, the working group was given a detailed presentation on the faculty salary structure. We also took into account other information as deemed appropriate, including the 2007 FAUW survey about Annual Performance Evaluations and summaries of empirical research on student evaluations of university teaching⁴.

Findings and Recommendations

1. The Performance Evaluation Process

Structure

The first fundamental issue the working group needed to consider was whether or not the merit review process requires radical revision. Compared to other universities, the merit process at UW plays an unusually important role in determining annual pay increases for faculty. The salary structure for university faculty has low salaries early in careers, with provisions built into the structure for annual increases so that career average earnings are commensurate with qualifications. These structural increases at many universities are called “progression through the ranks” and are awarded automatically. At UW, the size of these increases is entirely dependent on the faculty member's merit evaluation, and how it compares to those of other faculty members in the same Faculty.

Some faculty members feel, not without justification, that there are important costs to a performance review system that puts colleagues in competition with one another---costs in willingness to behave collegially or to share credit, for instance. Others find the process stressful and demoralizing, and question its benefits. Yet others consider merit evaluations to be valuable as incentive for and recognition of contributions, and for providing feedback about where improvement can and should be made. After considering the balance of feedback we received and various alternatives, the working group feels that there are important advantages to the current system, and recommends that the University maintain its basic structure.

⁴ Two sources that summarize the recent literature on which the working group has relied are: Pamela Gravestock and Emily Gregor-Greenleaf, *Student Course Evaluations: Research, Models and Trends*. Toronto: Higher Education Quality Council of Ontario, 2008.

Michael Theall, *Gathering, Analyzing, Interpreting, and Using Faculty Evaluation Data*, supporting documentation for an Academic Impressions Web Conference, November 16, 2007.

The working group also heard considerable sentiment that the existing system is not being used as effectively as it might be, that there is unfortunate variation in the way it is implemented across campus, and that it is poorly understood. Implementation is perceived by some to be unfair in various ways. Our recommendations are therefore aimed at enhancing the current process rather than radically changing it. Our view is that making these enhancements will go a long way towards eliminating perceived unfairness in the current process.

Recommendation 1.1 Maintain the basic structure of the current performance evaluation process. Implement enhancements to strengthen the process, as recommended in the body of this report. [Responsibility: Faculty Relations Committee]

Performance Evaluation Committees

The existing performance evaluation process is time consuming and stressful for those who are charged with evaluating the performance of their colleagues, especially in those units where the Chair conducts the evaluations alone. The working group therefore took it to be an important desideratum that its recommendations not increase workload, especially for Chairs. One factor that makes this job difficult for Chairs working alone is the difficulty of any one person having the breadth of knowledge required to confidently evaluate all aspects of performance, from the prestige of various scholarly outlets to the time commitment and impact of service contributions. Use of a performance evaluation committee, advisory to the Chair, while it does impose additional work on a small number of individuals from the department, will expand the breadth of knowledge brought to bear on equitably evaluating colleagues. In large or diverse units, the performance evaluation committee could include people with knowledge of the various areas of scholarship within the unit. A performance evaluation committee would also provide the Chair with a sounding board, and could reduce concerns about “arbitrary decisions” made by one person. In addition, a performance evaluation committee engages more individuals in the process, thereby making it less mysterious and promoting broader understanding. It can also have an important role in succession training.

Some departments explicitly choose to not use a performance evaluation committee for a variety of reasons. The working group heard concerns about imposing more work on faculty members and worries that the submitted reports would not be kept confidential. Storing the reports only in an official file in the Chair’s office, and reducing the frequency of performance evaluations for post-tenure faculty, should help mitigate these concerns.

It is the working group’s opinion that the advantages of a performance evaluation committee outweigh the perceived drawbacks.

Recommendation 1.2 Conduct the evaluations with a performance evaluation committee, advisory to the Chair. [Responsibility: Provost and Deans]

Expectations

A common source of dissatisfaction with the process is that it is often unclear to faculty why they have been assigned a particular merit score. Moreover, many faculty were unclear on what constitutes good work, and what evidence they needed to supply in their reports. The members of the working group recognise there is a great deal of variation from discipline to

discipline in terms of expectations in teaching, scholarship and service. However, the working group judges that it should be possible, at the department level, to make expectations clear, and to indicate what sorts of evidence are appropriate in the disciplines practiced in each department for documenting that good work is being done.

There is also some tension in trying to address two reasonable concerns: many worry about mere “bean-counting” replacing judgements of quality, since it seems clear that the University ought, for instance, to value one good article more than two poor ones. Others worry that moving too far from quantitative judgement opens up the door to arbitrariness. The working group feels that we cannot remove judgement from the process, but that a written statement of the relevant evaluation criteria, suitably worded to indicate that the criteria are general rules-of-thumb to which exceptions are possible, is the appropriate way to address this tension. For instance, while journal rankings are an important tool for estimating quality and impact in many disciplines, it should be possible for a faculty member to provide evidence that a particular work is “better than its venue.” In general, the onus must be on faculty members to provide sufficient information to allow for appropriate assessment of the quality and impact of their contributions.

Recommendation 1.3 Provide faculty members with a regularly updated document outlining the evaluation process and the performance expectations in their department for scholarship, teaching and service. [Responsibility: Provost, Deans and Chairs]

Best practice will be for Chairs to be open to suggestions for improving the process and that the document provides discipline-specific commentary on scholarship expectations and unit-specific guidance on teaching and service expectations. The document should be updated regularly to reflect current practices and should be supplementary to and consistent with Faculty guidelines and templates for merit reporting (Recommendation 1.11).

It is clear in policy, but unfortunately sometimes not clear to those doing evaluations, that when a person’s weightings depart from the standard 40/40/20, the expectations for quality of work are unchanged, but the expectations for quantity do change. This avoids the unfortunate situation where the same change is compensated for twice. For example, someone whose teaching is to be weighted at 20% should not be given a lower score on account of teaching half the regular load. A similar problem can occur in the case of part-time or reduced -load appointments.

Recommendation 1.4 Chairs and all members of the performance evaluation committee should be informed of, and take into account, all relevant information, including whether there has been any adjustment in a faculty member’s weightings. [Responsibility: Deans]

Recommendation 1.5 Insert suitable wording in the Memorandum of Agreement to clarify that for faculty on fractional load or leaves of absence, expectations for quality remain the same, while expectations for quantity change. [Responsibility: Faculty Relations Committee]

Timing and Feedback

The merit evaluation process is an important vehicle for communicating to both tenured and tenure-track faculty recognition for the quality and impact of their contributions to their department, the university and the academic community. It is therefore important that the outcome of the process is not just a numerical score, but informative feedback on the individual's performance. It is particularly important that tenure track faculty annually receive detailed information about their progress towards tenure and specific recommendations about what to change, if anything, to improve the likelihood of tenure and promotion.

Providing thoughtful feedback to each faculty member is a time consuming task, and not current practice in all departments on campus. While it is critical for tenure-track faculty to have this guidance on an annual basis, it may not be necessary to provide the same frequency of feedback to tenured faculty and continuing lecturers. Maintaining the annual evaluation for tenure-track faculty, but moving to a bi-annual schedule for tenured faculty members will reduce the total number of evaluations conducted annually. It will thus allow a performance evaluation committee more time to construct detailed feedback to the individuals evaluated without an overall increase in workload.

Recommendation 1.6 Evaluate tenured faculty bi-annually, and evaluate tenure-track faculty annually, and provide detailed feedback. [Responsibility: Faculty Relations Committee]

Recommendation 1.7 Provide every faculty member with histograms showing how their merit scores rank in their department and Faculty. The current histograms need to be redesigned to make them easier to read and more meaningful. [Responsibility: Faculty Relations Committee]

Rubrics and Dynamic Range

A frequent theme heard in the consultations is that while officially scholarship and teaching are equally valued, and officially service is an essential part of being an effective faculty member, the process rewards excellent scholarship much more than excellent teaching or service. Insofar as this is so from a financial point of view due to the merit process, it is because there is a much wider dynamic range in scores for scholarship than for the other two areas. However, we also heard many expressions of concern that while officially we have a scale that runs from 0 to 2.0, in practice it tends to run over a much smaller range (which varies a bit from Faculty to Faculty), with scores of 1.0 or lower being quite rare.

There is also good reason to worry about the existing rubrics attached to each score in the Memorandum of Agreement and their link to the tenure process. The rubrics for 1.0 is "satisfactory," but if that score is very rare (except for first year faculty, who get it as a matter of policy when there is insufficient evidence to give them any other score), then such a score may not, in fact indicate satisfactory performance. This is especially problematic for pre-tenure faculty, for whom this process must be a guide to how well they are progressing towards tenure; it seems to us unfair that the department would indicate to someone that they are doing "satisfactory" or even "good" work each year, then fail to make the case that she or he deserves tenure when the time comes.

Recommendation 1.8 Eliminate the rubrics attached to performance evaluation scores and increase the dynamic range of evaluation in all areas: scholarship, teaching and service. [Responsibility: Faculty Relations Committee]

The range of scores should continue to run from 0 to 2.0, but all multiples of 0.25 in that range should be possible scores. The University expects very strong performance in all three areas from its faculty---that they be productive scholars, effective teachers, and diligent and effective in the service roles they undertake---and the typical UW faculty member is a very strong performer. We recommend that the score of 1.25 be recognized as indicating performance at the high level typical of UW faculty, and so that both the median and the mean score will be at this level in each of teaching, scholarship, and service. Scores of 2.0 will be rare, indicating “career years” for the best among us; in any area only 3-5% of faculty should have such a score. A score of 1.75 will also be unusual, with 7-10% of faculty having that score in a given year. Some scores will be 1.0 or below; but since 1.25 indicates very strong performance, a score of 1.0 represents strong performance.

When these changes are made, the University must emphasize that the value of the average merit award in a Faculty will not be affected. If, under the current process, the average merit score is 1.6, then under the recalibrated process a score of 1.25 in that Faculty will be worth as much, monetarily and otherwise, as a score of 1.6 is now.

Recommendation 1.9 Chairs and performance evaluation committees should be prepared to use a significant range of scores, and a similarly wide range for all three areas. [Responsibility: Provost, Deans and Chairs]

There are two other process-related issues that the working group has considered. The first is the practice of assigning a score of 1 in all three performance categories in the first year after hiring. The working group believes that this practice penalizes new faculty and has the potential of creating salary anomalies. To put it differently: If department Chairs do their job well and place new faculty at appropriate salaries commensurate with their years of experience and scholarly work since the Ph.D. and consistent with the salaries of departmental colleagues, an artificial generic score in the first year and its associated effect on salary, undoes some of that good work. Although, the rationale for the across-the-board score of 1 is that little is known about the performance of the new faculty, that in fact is rarely true. New faculty arrive with scholarly work in progress and do not simply reset and start anew. Thus one can view their first year at UW as a continuation of their ongoing work, and so there is information on which to base an evaluation of performance. In the relatively rare case where there really is too little information to evaluate performance, we believe a more rational and fairer approach would be to assign our new colleagues a score equal to the average for their rank in their department.

Recommendation 1.10 Stop the practice of assigning a score of 1 in all categories to new faculty in their first year and, instead, assign a score based on actual performance or, when too little information is available, assign a score equal to the departmental average for their rank. [Responsibility: Provost, Deans and Chairs]

The second point relates to an observation the working group has made: The average of the merit scores is correlated with rank; that is, the average for Professors is higher than that for

Associate Professors which in turn is higher than the average score for Assistant Professors. Although one can perhaps justify this pattern on the basis of improved performance that typically comes with experience, the working group is of the view that faculty should be evaluated on the basis of their actual performance and that one's performance should be placed in the context of the expectations appropriate for peers at roughly the same point in their careers. Although we are not making a formal recommendation on this issue, we think it appropriate for departments to address the way expectations change with career stage in the document setting out performance expectations.

Communication

In our consultations, it became clear that understanding of the merit evaluation process is currently quite poor among faculty in general, and sometimes even among academic administrators. We therefore make a number of recommendations that will help ensure consistent, fair, and better understood evaluations. The use of a performance evaluation committee, advisory to the Chair, will be beneficial in this regard.

While the criteria for evaluating performance will vary naturally among departments, eventually the Dean must be involved to ensure an appropriate level of consistency across the departments in a Faculty, since the merit pool is determined at the Faculty level.

Recommendation 1.11 Establish in each Faculty a standard template to be used by all faculty for presentation of performance evaluation materials. It will require three years of information for teaching, scholarship, and service. Recommendation 3.1 suggests that Scholarship be evaluated over a three year period. While evaluations for teaching and service are for the period since the previous evaluation, the additional information can provide useful context for the Chair and the performance evaluation committee.
[Responsibility: Deans]

Recommendation 1.12 Provide Chairs and performance evaluation committee members with access to training sessions and support materials that allow them to perform this difficult task as effectively as possible. [Responsibility: Provost and Deans]

Recommendation 1.13 Establish a centralized, website for faculty, Chairs and Deans to access pertinent documents, policies and updates related to performance evaluations.
[Responsibility: Provost]

2. Teaching

In each Faculty, the main tool used for determining merit scores for teaching is feedback from student questionnaires. The working group heard many expressions of concern about this practice. Some doubt the usefulness of such questionnaires in general; some criticize too heavy a reliance on the questionnaires in the absence of other sorts of information; many feel that the information is used unfairly because there is insufficient attention paid to factors besides quality of teaching that can influence results (such as the gender of the teacher, or whether the instructor is perceived as an easy marker, or whether a class is an elective or a required course).

There is, in fact, a large body of empirical research on student evaluation of university teaching⁴. It suggests that while factors besides teaching quality can influence ratings, with a well designed instrument the main determinant of scores is teaching quality. There is growing consensus about best practices for teaching evaluations, and increasingly good evidence about how much various irrelevant factors influence ratings.

The academic literature on teaching evaluations makes clear that there are other legitimate sources of information about quality of teaching beside student evaluations. Some examples are self and peer evaluations, alumni ratings, evaluation of course materials such as syllabi, assignments and exams, and various measures of student outcomes (e.g., for graduate students, time to completion or successful publication might be used, depending on discipline). Some of this information is more appropriate for “formative” purposes (i.e., to help faculty improve their teaching) than for “summative” evaluation of current teaching quality. For example, some experts argue that peer evaluations of in-class performance, while a good source of feedback for formative purposes, is not much use for summative purposes, since it is rarely very critical. On the other hand, these experts say, peer evaluations of *course materials* such as course outlines and exams, can be useful for summative purposes. The literature we have seen, however, is consistent that it must be clear in advance of performance evaluations what sorts of evidence will count for summative purposes, and what is to be used only for purposes of improving teaching. The absence of clear rules gives faculty reason not to gather information that might help improve their teaching, because it may be perceived as negative evidence about the quality of their teaching.

Recommendation 2.1 Develop a common course evaluation instrument for use across campus, and establish a uniform method for collecting, analysing and disseminating the results. [Responsibility: Deans’ Council]

Recommendation 2.2 Provide to faculty members, Chairs and performance evaluation committees a ten year history of scores on key questions (e.g., overall evaluation of instructor, overall evaluation of course) for each course the faculty member has taught during the period under evaluation. [Responsibility: Deans]

Recommendation 2.3 Departments should be open to other sorts of evidence of teaching quality besides student questionnaires and have clear rules about what counts for summative evaluation and what is to be used only for formative purposes. [Responsibility: Department Chairs]

Recommendation 2.4 Develop a document for Chairs and performance evaluation committees summarizing the best empirical evidence about how factors aside from quality of teaching affect student evaluations. [Responsibility: Associate VP Academic with the support of the Centre for Teaching Excellence]

Recommendation 2.5 Evaluate graduate supervision primarily as part of teaching. Develop, at the department level, appropriate methods for evaluation of the quality of supervision, not merely its quantity. [Responsibility: Department Chairs]

3. Scholarship

In the focus groups and submissions to the working group, there were relatively few concerns expressed about the evaluation of scholarship compared to service or teaching. However, this observation must be considered alongside the remark of a Dean of a large Faculty who estimated that 80% of the complaints he gets about merit scores have to do with evaluation of scholarship.

The most frequently voiced concerns have to do with the difficulty of properly valuing scholarly contributions of a non-traditional nature (e.g., works of art, knowledge transfer, and social innovations). Another common concern is that the criteria used to estimate the worth of scholarly contributions are not appropriate, either because they amount to valuing quantity over quality, or, as some feel, because the size of a faculty member's grant is taken to be a reliable indicator of quality of contribution. Still another frequent concern is that a one-year "snapshot" is not an appropriate way to evaluate scholarly contributions in fields where the most important methods of publication are ones where projects take years to reach fruition.

The working group feels that Recommendation 1.3, which requires a clear statement of the criteria to be used in the evaluation of merit and states that the process must leave room for judgement and so for estimations of quality rather than mere quantity, is very important with respect to the evaluation of scholarship. In addition, we make the following two recommendations.

Recommendation 3.1 Evaluate scholarship on the basis of a three year period, and evaluate teaching and service on the basis of work done since the previous evaluation. [Responsibility: Faculty Relations Committee]

Recommendation 3.2 Clarify that scholarly work outside of the usual peer reviewed venues is valued, but that the onus is on the faculty member to provide evidence of its quality, impact and relevance. [Responsibility: Chairs and Directors]

4. Service

The working group heard in both written and focus group feedback that there is more serious discontent about the evaluation of service than either of the other two areas. This is also consistent with the FAUW survey.

Many faculty members who carry a heavy service load feel that it is unrewarded when they observe that colleagues who avoid service duties often do not get appreciably lower service scores. This indicates that the full dynamic range of service scores is not being used, and that the relative contribution of hard-working individuals is not being recognized. It is a particular aggravation to some that even poor service is not seen to be a barrier for promotion. That service is undervalued also seems evident to us from the difficulty the University has convincing people to become Chairs, Associate Chairs and Associate Deans. The working group heard concerns that the Outstanding Performance Awards are being used as a

mechanism to reward individuals who are in positions with heavy administrative duties. Another means of rewarding those in positions with heavy administrative duties is to assign a score of “2” in all three categories. Presumably these practices are due to a perception that our merit system alone does not provide sufficient recognition.

There is also a common feeling that important informal contributions to the University---things like being available to undergraduate or graduate students, mentoring new faculty members, or being willing to teach unfamiliar courses when the department needs someone to do it---are not being recognised sufficiently in the evaluation of service activities. Some faculty feel that colleagues who are rarely on campus avoid doing such informal service and, in effect, leave it to others.

There was also wide consensus among those who have done evaluations that service is the hardest of the three areas to evaluate. The extent of someone’s contribution to a committee is not evident merely from membership on it and, especially for external service, the requirements may not be clear from the title of the job. One Dean used the amusing term “Committee Hawk” for people who are good at spotting and being appointed to committees with impressive titles but involving little work. In order for the performance evaluation committee to properly evaluate service contributions, faculty members need to provide sufficient information about the nature and scope of their activities.

The University cannot function without internal service. This is especially true within small units where the effective operation of the unit depends on participation of all its members. It is therefore vital that a faculty member’s contribution is fully reflected in his or her evaluation. Using the full dynamic range of scores should mitigate some of the aforementioned concerns about contributions going unrecognized (by comparison to others).

Recommendation 4.1 Establish internal service as an essential duty for faculty.
[Responsibility: Deans’ Council]

Recommendation 4.2 Require faculty members to describe the nature and scope of their service contributions so that they can be properly evaluated. [Responsibility: Deans’ Council]

Recommendation 4.3 Include “good citizenship” among the criteria for evaluation of service, so that faculty members can be recognized for activities such as mentoring new faculty, being available in the department, being willing to take on hard-to-cover courses, and being available to students. [Responsibility: Deans’ Council]

Recommendation 4.4 For those in administrative roles that carry a stipend³, establish a portion of the stipend as an ongoing salary enhancement, at the end of the term, to be funded separately from the merit pool. Eliminate informal practices, such as automatic ratings of “2” or use of the Outstanding Performance Award in ways outside those specified in policy, for rewarding the carrying out of significant administrative duties.
[Responsibility: Provost]

Recommendation 4.5 Inform faculty when they take on a significant administrative role that they can have a temporary adjustment of their weightings for purposes of merit evaluation and that the adjustment involves a change in expectations for quantity of work in each area, but not for quality. [Responsibility: Deans and Chairs]

5. Related Issues

During the course of our deliberations, it became clear to the working group that there are some issues closely related to merit evaluations that we are not in a position to address appropriately at this time, but that we feel warrant further investigation.

It is not clear to us whether there are features of the merit evaluation process that are likely to generate anomalies, for instance on the basis of gender. There is some prima facie evidence that there may be some, including the fact that the University has needed, more than once, to employ ad hoc mechanisms to address gender based inequities. The current working group on Pay Equity for Women may shed some light on this matter when it has completed its work. We think this complex issue deserves further scrutiny so that any systemic problems can be addressed.

There are some intriguing possibilities for making use of existing policies in new ways to address problems faced by faculty trying to combine academic careers with modern family life. For example, a temporary reduction in teaching weighting for merit evaluation with a reduced load appointment could be a mechanism to help faculty in Science or Engineering maintain their research lab while their children are very young. This may benefit many young faculty, and help with recruiting and retention, especially of female faculty.

Recommendation 5.1 Further investigate two related issues: a) whether the merit evaluation process has features that generate gender-based anomalies; b) how to use the flexibility of the merit evaluation process, in conjunction with other policies, to make UW a more family friendly employer. [Responsibility: Provost and FAUW President]

Conclusion

No performance review process is free of problems. No single process could do all the good things that might be accomplished with performance reviews, since these things are sometimes in tension with one another. Bearing this in mind, it is the view of the working group that the existing process has served the University fairly well. Nevertheless, we think that the changes recommended in this report, along with a review of relevant sections of the Memorandum of Agreement and relevant policies by the Faculty Relations Committee to ensure their consistency with the changes, will result in a much improved process. It will be more transparent and better understood, more equitable, and provide more useful feedback to faculty members, especially tenure-track faculty who need it most. Implementing the recommendations of this report will enhance morale, reward, motivate and reinforce productive activity of faculty at the University of Waterloo.