

Policy on the Evaluation of Teaching, Research, and Service

Martin J. Mohlenkamp

April 18, 2019

Phase 2 of our current evaluation policy (included in section 2) describes how the TRS scores are determined, and thus is the core of our evaluation policy. It needs updating, at least to remove references to the quarter system etc. Now is then the time to consider other changes, large and small. Section 1 contains topics for discussion, which may lead to such changes. Section 3 contains our current (out-of-date, not-really-used) profiles.

1 Discussion Topics

I tried to separate questions and information from opinions, with mixed success.

1. The policy needs to handle Group I and Group II (possibly with new names) and regional faculty however they are eventually incorporated.

Martin: Possible language:

Faculty are divided according to the raise pools set by the college, and each set of faculty shall be evaluated separately. The chair and any other special cases shall be evaluated with the set in which they best fit. Scores of faculty in different sets are not comparable.

2. In current policy, the weight a faculty member has in a category is not taken into account when giving them a score. The main effect is high service scores for the chair and sub-chairs. We may have a wider variation in weights in the future, especially if regional faculty end up evaluated in the same set as Athens faculty. Should the weight be used during the rating process?

Martin: Possible language:

In comparing two faculty members in a category, their relative weights shall be taken into account. Thus a faculty member with a higher weight is expected to perform better, in quality or quantity, in order to receive the same score as a faculty member with a lower weight. The expectation need not be proportional to the weight.

3. The overall structure of the current policy is to give data to the committee members, and have them rank/rate “on a subjective basis”. The main alternative is a point system. Current practice uses the “subjective basis” for everything except teaching extras (such as independent studies). Policy and practice should be made to agree. What system should be used?

Martin: (a) My recollection is that the current practice of separating classroom performance from teaching extras was meant to reduce the effect of student evaluations on the overall teaching score.

- (b) Rotating or otherwise varying committee membership reduces the subjective effect by time-averaging over different viewpoints and values.

- (c) Some of the challenges of a point system are:
- i. Agreeing on the relative value of different activities. (How many conference talks equals a publication?)
 - ii. Agreeing when two instances of the same activity are the same. (Is a publication in a conference proceeding the same as in a journal?)
4. In current policy, each committee member is supposed to assign a score to each faculty member and then the mean scores are reported to the chair. This is subject to abuse, if a committee member gives a score of 10 out of favoritism or 0 out of malice. Recently we have used a system where discussion and compromise (averaging) takes place when ranking faculty members, and then the rank is mapped to a score. So far we have always achieved a consensus on the ranking and scores.

Martin: I find the ranking process meaningful, with pretty good agreement between evaluators, but the subsequent mapping to scores arbitrary. We could have the committee only report the ranking, with ties (and perhaps gaps) allowed. The chair would then map this to scores. Possible language:

The evaluation committee will attempt to come to a consensus on a ranking of all faculty members, with ties allowed. The vice-chair will report to the chair

- the ranking of the maximum set of faculty for whom consensus was achieved and
- a description of those faculty comparisons for which there was no consensus and the causes of disagreement.

For example, the vice-chair may report the consensus ranking $A > B = C > D > E$, that two committee members ranked $F = C$ while one committee member ranked $D > F > E$, and the reasons for disagreement regarding F . Other members of the evaluation committee have the option of including their own statements in cases when consensus was not achieved.

5. The scores that the evaluation committee send to the chair are advisory, so the chair could just ignore them. Should we limit the ability of the chair to change them? (The chair has a separate discretionary fund they can use.)

Martin: Here is possible language, continuing from the consensus ranking model above.

The chair sets the numerical scores, with the following constraints.

- For faculty members in the maximum consensus ranking, the scores must be consistent with this ranking.
- For faculty members for whom there was not consensus, the scores must be consistent with bounds imposed by the alternative rankings. In the example above, that means $C \geq F > E$.

6. Should scores be done “on a curve” by e.g. specifying the median score in policy? Current policy says “scores above 7 will be assigned to better than typical performances.”

Martin: I think we have had grade inflation (all scores move up) and upward grade homogenization (compression at the top). This may help keep the peace, but prevents the top performers from getting much reward. In 2018, there were only 6 scores below 7 awarded in any category to Group I faculty.

7. Should we update and use the profiles in section 3? They would help calibrate year-to-year scores and could be an alternative to setting a median.

8. It would look bad if a faculty member who is accused of serious misconduct receives high ratings. However, the evaluation committee is not equipped to judge serious misconduct. What should we do?

Martin: I suggest specifically saying we do not judge misconduct. Possible language:

Serious misconduct, such as sexual harassment of a student, falsification of research results, or theft, is not considered by the evaluation committee. Any sanctions for such behavior must be handled separately by the appropriate authority.

9. Current policy has the following appeals process:

A faculty member can request the Evaluation Committee to reconsider his/her rating by writing to the committee within one week of the evaluation being placed in the faculty member's mailbox. The request should also indicate why he/she believes the rating should be revised.

How should appeals be handled?

Martin: (a) The committee only knows the rating it sent to the chair, which might not be the rating the faculty member received. Only the chair can change the awarded rating. Therefore, I think appeals should go to the chair, who could then request further information or re-evaluation from the committee.

(b) Since the faculty member does not know the comparisons made with other faculty or other considerations in setting their score, it is hard for them know if it was unfair.

10. How should leaves and other special cases be handled?

(a) New faculty. Current policy says "ad hoc"

(b) Faculty fellowship leaves. The presumption is that research continues but teaching and service do not.

i. One semester.

ii. Full year.

(c) Medical or other leave.

i. Short term.

ii. One semester.

iii. Full year.

11. Should we set the evaluation year(s) in policy? Current practice is to use the academic year for teaching and service and the calendar year for research.

Martin: A research year starting October 1 would be easier administratively, since the faculty, committee, and chair work could be done in the Fall semester.

12. What activities should be moved to a different category?

(a) There was a suggestion to move writing/grading PhD comprehensive exams from service to teaching. Tutorial comprehensive exams could move too.

(b) Dissertation students are counted in research in the year they complete. Although it does not say this in the policy, this has been interpreted to mean they should not count in teaching in their final year. We could count them only in teaching throughout, only in research throughout, in teaching throughout and additionally in research in the year they finish, etc.

- (c) Current policy allows some activities to be counted either as research or as service (organizing a conference or session, editorships, reviews, refereeing, ...). This makes comparisons between people harder and sometimes people double-report. Should we specify some of these as research and some as teaching?

Martin: I would put refereeing and reviews in research, and editorships and organizing in service. In 2018, 11 people reported refereeing as research and 8 as service, including 5 who double-reported them.

- (d) Currently, being a member of a thesis/dissertation committee is under service, and is supposed to count only in the year of completion. Should this be in teaching? Should they count every year?
- (e) Should there be a separate category for teaching publications? Should they count for three years or one?
- (f) How should grants with a teaching component be handled?
13. What changes should be made in the research information gathered?

Martin: Policy says that each faculty member's portfolio should contain their preprints and reprints. When I was on the evaluation committee years ago I learned that the reprints I submitted went into a box in the office that did not make it to the committee, and that most faculty members submitted nothing. Recent practice has been to ask for a sample, which the committee members can/do read. Policy and practice should be made to agree in both what is collected and what is looked at by the committee.

14. What changes should be made in the teaching information gathered?

- (a) Of the things the current policy says to gather, those that might relate to classroom performance are:

first-day handouts, schedules, enrollment figures, grades given out, course evaluation material, summaries of complaints coming to the Chair and/or to the Dean's office, any memos or other material that come to the office about teaching-related activity of the person, and any other material deemed relevant.

For most people that means syllabi, grade distributions, and student evaluations. The past couple of years we have also collected a sample. There are known issues with using student evaluations, but without them we are left with only syllabi, grade distributions, and perhaps a sample. What else can/should we collect/use?

Martin: i. My recollection is that the current practice of separating classroom performance from teaching extras was meant to reduce the effect of student evaluations on the overall teaching score.

ii. The sample exams collected in 2018 were not at all useful. The optional extra sample was slightly useful.

iii. Some universities wishing to de-emphasize student evaluations are doing peer evaluations. Such evaluations need to be done systematically and with trained evaluators in order to provide useful information. I think it would be great if we did this. The feedback from these peer evaluations could also improve our teaching.

- (b) Should we specifically exclude sources of information like ratemyprofessor?

Martin: Possible language:

The evaluation committee will not consider rumors, unofficial student comments on websites, or similar materials.

15. What changes should be made in the service information gathered?

(a) Current policy says

...reports from departmental committee chairs (for committee members) and reports from the Department Chair for committee chairs, the graduate chair and vice-chair.

I also collected forms from the chair on committee chairs, course coordinators, and the advisory board. (Graduate chair and vice chair appear on both committee chairs and advisory board, but did not get individual evaluations.)

Martin: The reports from committee chairs sometimes say someone did a great job but rarely (never?) say anything bad. We have no way to compare quality or quantity between committees.

16. What information gathered in the evaluation process should be made public? We could have an annual report with publications, grants, talks, etc.

17. What information gathered in the evaluation process should be made available to the faculty? Should faculty members be able to see each other's student evaluations, syllabi, samples, submitted forms, etc.?

18. What information gathered in the evaluation process should be shared with the Promotion and Tenure committee? Whatever we decide should be specified in (evaluation or P&T) policy and done consistently.

19. Should any information gathered by the P&T committee be used in evaluation?

Martin: If we had a teaching observation letter about one faculty member and no letter about another, it would not help when comparing them.

2 Current Policy, Phase 2

Procedure for Determining Faculty Raises in the Department of Mathematics

....

Phase 2: Evaluation of research, teaching and service.

Ratings for research, teaching and service will vary from 0 to 10. A satisfactory performance will be assigned a score in the range 6 to 7. A score below 6 reflects a sub par performance and scores above 7 will be assigned to better than typical performances. These ratings will be based on information from a variety of sources including the Chair, departmental committee chairs, student evaluations and, on the appropriate departmental forms, information provided by the faculty member. The faculty member's overall rating will be a weighted mean of the evaluations in each of the three areas.

2.1 The Evaluation of Research.

Research and scholarly activity will be evaluated based on publication during the year and the preceding two years and scholarly activity during the year. The faculty member's research portfolio will contain

- 1 Reprints of published articles and books.
- 2 Preprints of unpublished articles and books.
- 3 A summary of research and scholarly activity provided by the faculty member.

The portfolio should include activity related to publication for the preceding three years. Using the departmental form, the faculty member should report events related to the publication of an advanced monograph, advanced texts or refereed papers. The creation of a preprint, the acceptance of a paper and publication of a paper are regarded as distinct events; the (approximate) date of each event should be given in the faculty member's report.

The following are activities that will be reported on an annual basis:

- 1 External research grants (new and continuing).
- 2 Internal research grants received. These are grants that are reviewed by external referees (i.e. OURC grants, Baker Fund grants).
- 3 Editing a monograph or proceedings of a conference, revisions of previously published monographs.
- 4 Invited talks. This category includes colloquium talks, AMS Special Session talks, seminars presented at other universities, and conference talks.
- 5 Contributed papers presented in person.
- 6 Submitting a research grant proposal (internal or external).
- 7 Attending a research workshop or conference.
- 8 Contributed paper, not presented in person.

- 9 Having a Ph.D. student complete a dissertation.
- 10 Speaking at or attending an ongoing in-house seminar.

The following activities can be reported either as scholarly activity or service (but not both):

- 1 Organizing a conference (reported the year of the conference.)
- 2 Editor of a Journal.
- 3 Organizing an AMS Special Session (reported in the year of the session).
- 4 Book and article reviews.
- 5 Referee for journal, outside examiner of a dissertation in mathematics or reviewer for a tenure application.

Each member of the Evaluation Committee will use the above information to rate a faculty member. The Evaluation Committee will read each person's folder and, on a subjective basis, each committee member will assign the person a rating x , $0 \leq x \leq 10$. The Evaluation Committee will report the mean of these ratings to the Chair as a suggested research rating.

2.2 The Evaluation of Teaching.

The procedure for the teaching assessment is as follows. There will be a "teaching folder" compiled for each person. This will contain all material that the office has about the courses taught: first-day handouts, schedules, enrollment figures, grades given out, course evaluation material, summaries of complaints coming to the Chair and/or to the Dean's office, any memos or other material that come to the office about teaching-related activity of the person, and any other material deemed relevant. In addition the faculty member will provide a summary of teaching related activity. A departmental form will be used to report on the number of tutorial students, Ph.D. students, course development work, and so forth. There is also a place for the person to report any other activity relevant to the evaluation of teaching.

There are a variety of activities that can positively affect one's teaching evaluation. These include, but are not limited to, such activities as:

- 1 Taking on tutorial or dissertation students.
- 2 Offering approved independent studies.
- 3 Developing new courses or modifying existing ones.
- 4 Attending teaching related conferences or workshops.
- 5 Being well regarded by students.

There are also a variety of activities that can negatively affect one's teaching evaluation. These include, but are not limited to, such activities as:

- 1 Failing to pass out first day handouts.
- 2 Failing to keep office hours.
- 3 Failing to meet classes every time and for the full period.
- 4 Failing to make effective use of class time.

- 5 Failing to hand out evaluations on approved department forms.
- 6 Failing to follow college policy with regard to final examinations.
- 7 Failing to cover the syllabus of the course.

and, in general, for failing to do the routine things required of all faculty by the Faculty Handbook and/or departmental policy.

Both the quality and quantity of such activities will be considered in the evaluation.

Each member of the Evaluation Committee will use the above information to rate a faculty member. The Evaluation Committee will read each person's folder and, on a subjective basis, each committee member will assign the person a rating x , $0 \leq x \leq 10$. The Evaluation Committee will report the mean of these ratings to the Chair as a suggested teaching rating.

2.3 Evaluation of Service.

The service portfolio will consist of information provided by the faculty member, reports from departmental committee chairs (for committee members) and reports from the Department Chair for committee chairs, the graduate chair and vice-chair.

The following is a list (though not an exhaustive list) of activities that can be considered as "Service."

- 1 Proctoring a Ph.D. comprehensive examination.
- 2 Supervising a multi-section course.
- 3 Writing and grading a tutorial comprehensive.
- 4 Substituting for an absent faculty member.
- 5 Writing and grading a Ph.D. comprehensive examination.
- 6 Being a member of a thesis/dissertation committee. (Will be counted in the year of completion of the degree.)
- 7 Member of an ad hoc committee (e.g., member of a College Promotion and Tenure Committee.)
- 8 Advising students.
- 9 Being an active member of a department committee.
- 10 Being a member of a college/university committee.
- 11 Service to the community.
- 12 Service to the profession.

The following are among the things that will impact negatively on the service evaluation.

- 1 Refusing appropriate service requests.
- 2 Accepting an assignment and then not carrying it out satisfactorily.
- 3 Failure to attend meetings of assigned committees.
- 4 Failure to attend department meetings.

Each member of the Evaluation Committee will use the above information to rate a faculty member. The Evaluation Committee will read each person's folder and, on a subjective basis, each committee member will assign the person a rating x , $0 \leq x \leq 10$. The Evaluation Committee will report the mean of these ratings to the Chair as a suggested service rating.

2.4 Exceptions.

There are some exceptions to the above procedure. The Evaluation Committee will evaluate untenured faculty in their first two years and faculty on leave on an *ad hoc* basis. Faculty taking two or more quarters of sabbatical leave will:

- 1 For the academic year of their sabbatical leave, receive the mean of their teaching and service ratings from their three previous evaluations.
- 2 For the academic year of their sabbatical leave, have their weights set at the mean values of their weights for the previous three evaluations.
- 3 Have their research ratings determined as described in section 2.1.

2.5 Incomplete portfolios.

The Evaluation Committee will evaluate all faculty members on all aspects of their duties. If the faculty member does not present any information about his/her activities to the Committee, then he/she will be evaluated on what is provided by the Chair. This will include information given to the Chair by other official sources. The faculty member will be evaluated as if there is no other activity.

2.6 Appeals.

A faculty member can request the Evaluation Committee to reconsider his/her rating by writing to the committee within one week of the evaluation being placed in the faculty member's mailbox. The request should also indicate why he/she believes the rating should be revised.

3 Current Profiles

Profiles in Teaching, Research and Service

This section contains some examples of research, teaching and service and the ratings they would receive. These will establish guidelines for the Budget Committee and to assist faculty members in evaluating their own performance. Note should be taken that these examples are not the exclusive ways to attain these ratings; rather, they should be taken as representative members of equivalence classes.

Profiles in Research

1. Attends and speaks in an in-house seminar, attends a conference. Rating: 1 - 2
2. No publication record for last 3 years. Attends an in-house seminar without giving talks. Attends a research conference (does not present a paper). Does minimal reviewing (1-2 papers). Rating: 2 - 3
3. During the past three years had an acceptance or publication of an article. During the year attended and spoke at a research workshop or conference and attended an in-house seminar. Rating: 4 - 5
4. Has one paper published (accepted) in one of the good journals (in the last three years). Active participant in an internal seminar (1-2 talks). In the past year, has given an invited talk and presented a paper at a conference (Contributed talk). Rating: 6 - 7
5. During the past three years published one paper and another one accepted; one or two invited talks, referees and speaks in an internal seminar. Rating: 6 - 7
6. The professor has published 3 joint papers in the last three years. He/She has 2 joint papers accepted but not published. He/She gave 2 invited talks at other universities in the last year. He/She refereed 5 papers and wrote 2 reviews. Rating: 7 - 8
7. Publication activity is consistent with publishing one average paper per year in average journals (i.e., in a three year time period, three papers have appeared, one is accepted but not appeared and one is under review), attends a conference, attends and speaks in an in-house seminar. Rating: 7 - 8
8. Publication activity is consistent with publishing almost one average paper per year, attends a conference, referees papers and reviews books or articles. Rating: 7- 8
9. Publication activity is consistent with publishing about 2 average papers per year in average journals (during last 3 years). Attends and speaks in an in-house seminar. During the year, presents one contributed paper at a conference and attends a research workshop/conference (different from the conference listed above). Does above average reviewing/refereeing (7-8 papers/year). Rating: 8-9.5

10. Publication activity is consistent with publishing 2.2 average papers per year. Gave one invited talk and attended a conference during the year. Rating: 9-10

11. Publication activity is consistent with publishing 1.67 average papers per year and has an external grant. During the year, the professor gave an invited one-hour address at a conference in his/her field, 2 contributed talks and one of his/her Ph.D. students finished his/her degree. Rating: 9-10

Profiles in Teaching

1. Although an adequate lecturer, the professor regularly fails to cover the required syllabus in multisection courses (i.e. 250B, 163, 263). The professor does not distribute a first day of class handout and spontaneously cancels classes. Rating: 0 - 3

2. Student regard the professor as poor: student evaluations indicate that the professor creates a poor learning environment in the classroom (i.e. questions are discouraged, lectures are disorganized, etc.). The professor performs the minimum duties. Rating: 3 - 4

3. Students regard the professor as adequate. Fails to complete syllabus for multi-section courses on a regular basis. In addition, the professor has taken on some additional teaching duties (i.e. a tutorial or dissertation student or some independent studies). Rating: 4 - 6

4. The professor fulfills his/her minimum duties, teaches adequately and is generally unavailable outside of office hours or class time. Rating: 5 - 6

5. Students regard the professor as adequate: student evaluations from the Calculus sequence reflect that the instructor is conscientious, generally prepared for class but is not an especially inspiring instructor. The faculty member takes on no other teaching related activities, but is generally available to students and colleagues outside of regular office hours. Rating: 6 - 7

6. Students regard the professor highly: the professor delivers clear lectures, uses class time effectively, is generally available outside of class. Rating: 8

7. Students regard the professor highly. In addition, the professor has developed a new course or experimental course or substantially revised the content of an existing course OR taken on some additional teaching duties (i.e. a tutorial or dissertation student or some independent studies). Rating: 8 - 9

8. Students regard the professor very highly. In addition, the professor has taken on some additional teaching duties (i.e. a tutorial or dissertation student or some independent studies). The professor has also developed a new course or experimental course or substantially revised the content of an existing course. Rating: 9 - 10

9. Students regard the professor very highly. As part of supervising a multiple section course, the professor meets with the TAs teaching the course. The professor offers regular 'help' sessions in addition to regular classes. The professor also regularly attends teaching workshops and helps

keep the Department abreast of developments in undergraduate and graduate education.
Rating: 9 - 10

Profiles in Service.

The profiles mention 'major' committees, 'minor' committees and 'routine activities'. In 'normal' years, the following committees and activities would be categorized as follows:

Major committees include: Graduate, Undergraduate, Promotion and Tenure, Evaluation, Assessment, Staffing, College Curriculum, College Staffing, University Curriculum Council subcommittees, OURC review panel.

Minor committees include: Library, Visitors, Outreach, Grievance, Nominating, Publicity, Joint EECS/Math, Building Committee.

Routine activities include: Member of an ad hoc committee such as Dean's P&T committee, college representative on a masters or Ph.D. committee, departmental task force, or comprehensive exam committee. Also, proctoring comps, advising, etc.

The committee recognizes that in some years, due to special circumstances, major committees may become minor committees and vice versa. The committee also recognizes that chairing a committee can be substantially more work than being a member of a committee.

LOW:

1. Although a member of a departmental committee, the person is inactive, in the sense that he/she rarely attends committee meetings, and when attending doesn't contribute much. Rarely attends departmental meetings. Rating: 0 - 1
2. A member of a Department committee and attends regularly, but does not contribute very much. Takes on assignments but does not do them and refuses some assignments altogether. Advises fewer than 6 students/quarter for UC, A&S or the department. Rating: 1 - 2.5
3. An active member of a minor departmental committee. Advises 6 to 10 students per quarter and proctors a comprehensive exam. Rating: 3.5 - 4.5
4. An active participant of two minor committees, regularly attends faculty meetings, some advising and two or three routine tasks. Rating: 5 - 6

MEDIUM

5. An active member of a major committee and a minor committee. Performs three or four routine activities. Regularly attends faculty meetings. Rating: 6 - 7
6. Active member of one major committee and two other committees (departmental or college). Performs three or four routine tasks for the department. Regularly attends faculty meetings. Rating: 7 - 8
7. Active member of one major committee and chairs an active minor committee. Performs three or four routine tasks for the department. Regularly attends faculty meetings. Rating: 7 - 8

HIGH

8. Is an effective agent in developing a new (or revitalizing an old) service to the Department (i.e. improving library facilities; modernizing a computer lab). Serves on a minor committee. Performs 5 or 6 routine services at any level. (e.g., sits on Dean's P&T committee, outside reader on a few masters or Ph.D. thesis committees, writes a comprehensive exam, helps out with a math competition, makes the arrangements for a Kennedy/Frontier of Science lecturer) Rating: 8 - 9.5
9. Is the chair of a major departmental committee. Completes a major service assignment (e.g., coordinates a program review, supervises the creation of a math contest, sets up and maintains a computer lab.) Organizes a Special Session at an AMS meeting and referees more than 5 papers. Does four or five routine services. Rating: 8 - 10
10. Member of two major departmental committees and of a college committee that requires a significant amount of work. Advises 10 students/quarter, writes and proctors a comprehensive exam, referees five papers, and one or two other routine activities. Rating: 9 - 9.5
11. Serves on two active departmental committees (or on one such along with another time intensive task such as the scheduling officer), participates on several ad hoc committees or performs a number of routine services for the department, serves on a significant university

committee (such as Faculty Senate, or University Promotion and Tenure) and reviews or referees some articles.

Rating: 9 -10

12.Active participant on a 4-5 committees (departmental, college, or university), chairing two of them. does 3 or 4 routine tasks for the Department such as covering classes, having 10 or more advisees/quarter or refereeing some articles. Rating: 9.5

- 10