April 17, 2020 (with revisions of Aug. 25 and Sept.16, 2020)

2019-20 Ad Hoc Committee on Student Evaluations of Teaching (SET):

Report and Proposed Interim SET Form

Members: Mitra Sharafi (chair), Keith Findley, Linda Greene, Peggy Hacker, Kevin Kelly (ex officio), Ursula Weigold, Desmund Wu

A. Report:

1. Overview

In AY 2019-20, the Ad Hoc committee on Student Evaluation of Teaching (SET) was given the following charge:

> Implement the revised recommendations of the Ad Hoc Working Group on Course Evaluations[1] by proposing a new interim Student Evaluation of Teaching form to the APC and faculty within a year. Consult with outside experts and/or study and adapt existing validated forms to the extent feasible. Determine whether and how SET data should be used in nominating Law School faculty for university teaching awards.

We read and discussed research findings in this area[2] and the work of other UW departments (including History and Political Science) that had recently revised their SET form. We consulted with on-campus experts including Mo Bischof and Regina Lowery of the Student Learning Assessment unit (Office of the Provost), who discussed best practices with our committee. Our chair also discussed our interim SET form with the head of the Arts and Humanities Divisional Committee. E-mail correspondence with the head of the Social Sciences Divisional Committee in AY 2018-19 is included below as Appendix B. Our work was a continuation of recent efforts at the Law School, particularly the memo on teaching evaluation data by Gwendolyn Leachman (Feb. 6, 2018) and the Report of the Ad Hoc Working Group on Course Evaluations, chaired by David Schwartz (March 25, 2019). The long-term aim of all of these efforts has been to make our SET form better reflect teaching effectiveness and student learning while reducing bias in student responses.

The Law School's existing SET form consists of five questions, rated on a five-point Likert scale (from poor to excellent), with space for comments. The proposed interim SET form consists of ten questions, seven of which (Q3-10) are rated on a six-category scale (from strongly disagree to strongly agree, plus not applicable), with space for comments. Both the existing and proposed interim SET forms appear below (see B-C). The committee

---

[1] The revised recommendations were approved on May 3, 2019.
[2] A summary of the research findings is available here: http://bit.ly/evalsmap

has been guided by current research findings and best practices in revising the SET form. A

growing body of research finds that there is difference and bias in scores and comments referring to instructors who are female or of color (for example, see Kristina M. W. Mitchell & Jonathan Martin, *Gender Bias in Student Evaluations*, 51 PS: Political Science & Politics 648–652 (2018); Lisa L. Martin, *Gender, Teaching Evaluations, and Professional Success in Political Science*, 49 PS: Political Science & Politics 313–319 (2016); Meera E. Deo, *A Better Tenure Battle: Fighting Bias in Teaching Evaluations*, 31 Colum. J. Gender & L. 7–43 (2015)). Gwendolyn Leachman's 2018 memo analyzes UW Law School evaluations specifically. Our aim has been to minimize the use of highly subjective questions (which leave more space for bias) and questions that students may not really have the knowledge base to answer. We have tried to replace these with questions that measure less subjective features of an instructor's teaching. In some instances, we have included questions that assess the student's own personal experience with the class, which students *are* capable of assessing even if it is subjective. We have aimed to ask about only one thing per question, providing greater clarity when interpreting student responses. We have put greater emphasis on the student's own role in learning. And we have tried to avoid questions that have a "popularity contest" quality to them.

2.  Changes:

We have made a couple of major changes to the Law School's existing SET form. We have cut the question about the professor's knowledge of the subject matter (original Q1). Colleagues have noted that students are not in a position to judge because they are new to the subject matter themselves and do not have access to the professor's full cache of knowledge on the subject. Original Q1 measured confidence levels more than actual knowledge. Some colleagues also felt that this question operated against instructors who had not been teaching the course for many years. We have also cut the original global question (original Q5 on overall quality). One might assume that original Q5 would be a summary or compilation of original questions 1-4. However, colleagues reported that Q5 was often not the average of Q1-4. Instead, it seemed to be its own independent question, and one that could take on a "popularity contest" quality that we seek to avoid. It is worth noting that the Law School no longer calculates a faculty mean because of the statistical problems with such a calculation.[3]

We have also revised some questions. The original Q2 (preparation and organization) was poorly designed because it asked about two different things, making the results unclear. Did a poor score for Q2 mean that the instructor was poorly prepared, poorly organized, or both? There was also concern from some committee members that

---

[3] On why the calculation of an average across Q1-4 and comparison of this average to a faculty mean are statistically unsound, see Philip B. Stark & Richard Freishtat, *An Evaluation of Course Evaluations*, ScienceOpen.com, (2014), http://bit.ly/freishtat.

instructors who encourage discussion and decide to extend a group conversation in class (for instance, catching up the next day) were penalized by original Q2. Relatedly, in some instances, students may not be well positioned to assess the professor's organization or preparation, when both are quite present but the class session is fluid by design. Students may not realize that what feels like free-flowing and impromptu coverage of material is actually well planned and thought through by the professor. We have instead introduced new Q9 (the organization of the course helped me learn the material), which asks about organization in a broad way that refers to the entire course, creating space for overflow discussions. The question is also designed to focus on what students *are* well positioned to assess—their own understanding of the organization of the course. The committee believes that instructor preparation is built into other questions on the new SET form and therefore does not require its own separate question. We have replaced original Q3 (professor's personal receptiveness to students) with questions that ask about less subjective qualities, which students are capable of assessing, as in Q5 (instructor encouraged student participation), Q6 (instructor's answers to student questions), Q7 (availability for consultation outside of class), and Q10 (class environment). Finally, many colleagues were dissatisfied with original Q4 (success in getting you to think in greater depth about topics discussed in class): it seemed to imply that good teaching meant making students think about the course materials outside of class. Colleagues found this random and not deserving of its own question. Many of the new questions (Q3-10) get at students' "digestion" of course ideas, but by asking about features that we considered more essential. We have added new questions specifically about readings (new Q4); consultation outside of class, including office hours (new Q7); and class environment (new Q10). We have also added new questions about student preparation (new Q1) and interest level in the course subject (new Q2).

3. Other issues:

The following issues have emerged about the presentation of SET data in the tenure and university award contexts:

(a) Tenure and Promotions Committee:

What approach should the Tenure and Promotions Committee take to numerical scores in annual pre-tenure reports and ultimately, in tenure files sent to the Divisional Committee? There are two sub-issues here:

i. Our current interim procedure has been that instructors are given their written SET comments, but not their SET scores unless they request them. For pre-tenure candidates, this means that an instructor may not receive their scores at the end of every semester. However, these scores will ultimately be sent to the Divisional Committee in the candidate's tenure file. The question is therefore: in

annual pre-tenure reports (and ultimately the tenure file), should the scores be included or described in any way, or omitted? If a pre-tenure candidate wants to avoid knowing their scores, putting them in the annual report will mean that they *will* see their scores. We are currently in a grey zone on this question, and the Tenure Committee seeks clarification.

ii. There may be a tendency to de-emphasize SET scores in relation to SET comments in annual pre-tenure reports. Both kinds of SET data may exhibit bias. But if comments are privileged over scores, the selection of comments to be highlighted in the pre-tenure annual report (or tenure file letter) becomes even more important than it used to be. In other words, it would be possible to create a skewed impression of a candidate's teaching by cherry-picking the comments. The 2019-20 Tenure Committee's favored approach is this: the author of the annual report (or tenure file letter) should summarize patterns in the SET comments without quoting any particular comment.

(b) Divisional Committees:

How would the Divisional Committees (Social Sciences, Arts and Humanities) regard a tenure file that contained no scores at all? This is not something we are proposing at the moment, but the importance of numbers may be a key question for longer-term SET planning at the Law School. The Social Sciences Divisional Committee seems receptive to eliminating all scores from tenure files (see Appendix B). At a meeting with the Arts and Humanities Divisional Committee chair in spring 2020 (Marcelo Pellegrini), our committee chair found him to be potentially open to this idea. However, Prof. Pellegrini emphasized that he could only speak for himself, not the whole Divisional Committee. He also noted that it is unlikely that all of the Divisional Committees will produce a statement reflecting a unified approach among them. He emphasized that once the Law School has finalized its position, the future committee chair should attend a meeting of each relevant Divisional Committee (ideally the first or second meeting of the year, in the fall) to explain the Law School's new approach. In a potential future tenure file without scores, it would be essential for the chair's letter to provide context for the change and to fully explain our new standards of teaching assessment.

(c) University Teaching Awards:

i. How should SET data be used in nominating Law School faculty for university teaching awards? We recommend that the Law School

nomination file include teaching evaluation scores (although not faculty mean) and quotations from SET forms. Unlike the tenure process, the nomination of faculty for these awards is not evaluative; it is meant to highlight praise for faculty teaching. As a result, we are not as concerned with potential cherry-picking and distortion of comments in this context as in the tenure process. Also, unlike the tenure dossier, an awards nomination does not include full SET forms. Including quotations will therefore provide information that will not otherwise be available to the awards committee. We recommend providing scores (including potentially the conversion of the new Q3-10 into numerical form) because excellent teachers who may not be terribly innovative in their teaching methods may lose out to candidates in other units otherwise. (University teaching award committees put a high premium on innovative teaching methods.)

ii. A larger question for the future is whether the Law School should consider expanding the process of nominating instructors for university teaching awards. Currently, the Law School follows the lead of students and recent alumni, who vote for Law School teaching awards. If our aim is to avoid "popularity contests," does this existing process not mean that the university award nomination process is simply amplifying the effects of the earlier student vote? Future SET committees may want to consider alternative ways of broadening the pool.

We propose to discuss these important issues at an upcoming faculty meeting.

4. <u>Role of SET data in relation to other kinds of information about teaching:</u>

Following the faculty approval of the revised recommendations of the Ad Hoc Working Group on Course Evaluations on May 3, 2019, the Law School's tenure rules were amended. The revised rules are as follows:

<u>UW Law School Rules for Tenure Decisions, and Pre-Tenure Review of Untenured Faculty (version of May 6, 2019):</u>

<u>Part Two: Rules for Tenure Promotions</u>

2.1 Criteria for Granting Tenure:

(b) Elaboration on Tenure Standards

(1) Teaching. Teaching quality can be judged on a variety of measures, including peer evaluations, candidate self-assessment,

and evidence of effort, activity and innovation observed by members of the law school community or provided by the candidate. Student evaluations based on the Student Evaluation of Teaching ("SET") form in use at the Law School as of 2018 shall not be used comparatively to those of other instructors, or to Law-School-wide numerical means. SET forms may be consulted, though as evidence of teaching quality, SET data shall not be given conclusive or predominant weight, nor given more weight than any other type of evidence, and due regard shall be given to the possible influence of race and gender bias in such forms.

This revision is consistent with the Fall 2017 memo by the UW Social Sciences Divisional Committee Chair, announcing an initiative to "broaden the criteria by which we assess excellence in teaching" for tenure reviews (see Appendix A).

In the longer term, the Law School should generate and draw upon a wider body of information to assess teaching. This will have the effect of de-emphasizing SET data. Alongside SET data, here are the other kinds of teaching-related information that the Law School may consider:

(a)     Peer visit reports: We already do this, of course. If these reports are to play a larger role in the assessment of teaching, it is more important than ever that there be no gaps in a pre-tenure candidate's file (i.e., no semesters for which no peer report was submitted). It will also be worthwhile to create a standard peer visit process to promote greater consistency across evaluators. Future SET committees should provide guidance, including the development of best practices and a standardized form. See Appendix C for a model from Letters & Sciences. Pedagogy literature and other institutions' websites may also provide useful examples. Another idea could be to video record the class for the instructor's own review, or so that the instructor and assessor could review parts of it together.

(b)     Teaching dossier: Instructors could compile a teaching dossier that includes syllabi, self-evaluations (see below), a teaching philosophy statement (as in the tenure file), and evidence of continuing pedagogical education (see below), for instance. Tenure candidates already do something like this. The Law School could consider requiring something similar post-tenure, too.

(c)     Self-evaluation: Instructors could write a short assessment (1-2pp.) of their own courses and teaching at the end of each semester, describing how they changed their courses and teaching techniques in relation to previous semesters, their assessment on the effectiveness of those changes, and perhaps how they plan to change their teaching and courses for the future.

6

(d)     Evidence of "effort, activity and innovation…provided by the candidate" (quoting the revised rules): instructors may attend on-campus teaching workshops (like the Discussion Project, Active Teaching Labs, Blend@UW, or a variety of one-time teaching events on offer) or perhaps some kind of pedagogical training off campus or online. If the Law School decides to require this kind of continuing education, it should compile a list of resources so that instructors can learn about such opportunities easily. A list of resources online on inclusive and effective teaching could include titles like Meera E. Deo, *Unequal Profession: Race and Gender in Legal Academia*; Gabriella Gutiérrez y Muhs, *Presumed Incompetent: The Intersections of Race and Class for Women in Academia*; and Margaret Price, *Mad at School: Rhetorics of Mental Disability and Academic Life.* It could also include links to campus resources like Teaching and Learning Office of the Provost and external events like the AALS Workshop for New Law School Teachers.

Note: One way to combine (c) and (d) would be for the next SET committee to develop a short questionnaire or template with instructions such as: *Please list three pieces of constructive criticism you have received on your teaching and what you have done to respond to each. Include any professional development opportunities you have engaged in to improve your teaching (may include events at the Law School), noting what you have learned and how you plan to incorporate it into your teaching.*

(e)     Focus group or facilitated discussion by an assessor: For pre-tenure candidates, a member of the Tenure Committee or other "assessor" could visit the instructor's class for one or part of class session in order to host a focus group or whole-class discussion of the instructor's teaching. The instructor would be absent for this session. The UW Teaching Academy may offer guidelines for this kind of assessment and/or training for assessors.[4] This model provides direct feedback from students, unlike the other suggestions noted here. It could also eliminate the more extreme and even abusive anonymous comments that students sometimes submit on SET forms. There are a few issues with this concept. First, the group dynamics of the students involved could produce a snowball effect with skewed results. Second, this model is labor-intensive and would require training for assessors. Standardized questions would have to be developed, and the feedback might be recorded and coded. As a result, the Law School could consider having only one or two such assessments per TT instructor pre-

[4] https://teachingacademy.wisc.edu/

tenure, at most. Two would allow observations about change over time (for instance, one at the 1/3-mark to tenure, and the second at the 2/3-mark).

The larger point is that it is worth thinking broadly and creatively about alternatives or additions to SET data.

B.  Existing SET Form (2001 (or earlier)-2020):

The Law School's existing SET form consists of five questions, rated on a 5-point Likert scale:[5]

1.  The professor's knowledge of the subject matter of the course.
2.  The professor's preparation and organization of the class, including organization of the entire course and preparation for each class.
3.  The professor's personal receptiveness to students, including receptiveness to consultation outside of class.
4.  The professor's success in getting you to think in greater depth about the topics discussed in the class.
5.  The overall quality of the professor's teaching in this course.

Students are also asked to write comments, with the following instructions:

Please write any observations, especially what you liked or disliked about the course or teacher, in order to encourage effective practices or to help identify weaknesses and possibilities of improvement. Thoughtful comments by many students often provide helpful information which 'circle the number' responses cannot. (There is no limitation to the answer length.)

C.  Interim SET Form (proposed in spring 2020; approved at faculty meeting of Sept.4, 2020):

We propose the following SET form:

Student evaluations of teaching play an important role for faculty. Your opinions influence the review of instructors that takes place every year and can help instructors improve their courses. The University of Wisconsin Law School recognizes that student evaluations of teaching may be influenced by students' unconscious and unintentional biases, including but not limited to bias relating to the instructor's race and gender. Women and instructors of color are systematically rated lower in their teaching evaluations than white men, even when there are no actual differences in the instruction or in what students have learned.

As you fill out the course evaluation please keep this in mind and make an effort to resist stereotypes about professors. Focus on your opinions about the content of

---

[5] Key: 5 (excellent), 4 (very good), 3 (good), 2 (fair), 1 (poor)

the course (the assignments, the readings, the in-class material) and not unrelated matters (such as the instructor's appearance).[6], [7]

You must answer questions 1-10 before submitting this form. Comments are optional.[1]

Student role: Please indicate the option that best applies:

1. Student preparation: On average, I spent 0-2 hours/2-4 hours/4-6 hours/more than 6 hours preparing for this course each week.
2. Having taken this course, my interest level in this subject matter is less/the same/more than before.

Instructor's teaching: Please indicate the option that best applies: strongly disagree/disagree/neither agree nor disagree/agree/strongly agree/not applicable

3. The instructor helped me understand the course subject.
4. The assigned readings helped me understand the course subject.
5. The instructor encouraged student participation, as appropriate to the class format.
6. The instructor's answers to student questions aided my learning, whether in class or in subsequent follow-up.
7. The instructor was available for consultation outside of class.
8. The course was intellectually challenging.
9. The organization of the course helped me learn the material.
10. The class environment felt like a welcoming place to express my ideas.

Comments: Please elaborate on any of the above or on other matters, including: what you found most and least helpful to your learning, and any changes you would suggest for the future.

## Appendix A:

### 2017 Memo from the Social Sciences Divisional Committee

---

[6] A recent experimental study found that students given this text before filling out SET forms ranked female instructors higher than students in a control group. David A. M. Peterson et al., *Mitigating Gender Bias in Student Evaluations of Teaching*, 14(5): e0216241 PLOS ONE (2019). The authors of the study designed the prompt to overcome the influence of bias by making students aware of the possibility of bias, attempting to motivate them to suppress its effects, and providing cues about what other considerations to use when answering the questions about the instructor. *Id.* at 3.

[7] Some members of our committee were concerned about the risk of amplifying bias among some students, who might take these opening paragraphs as a provocation. Given the research noted in the preceding note, we are proposing this text for our interim SET form. The next committee that develops the new permanent SET form may want to re-visit this text.

---

[1] There was debate among faculty in 2020 over whether to make comments optional. Specifically, there were differing views over whether students should be able to click "submit" without adding any text to the comments box. Future committees may want to re-visit this question.

To: Department Chairs, Directors and Deans

From: Maryellen MacDonald, Chair, Social Sciences Divisional Committee

Re: Tenure Reviews during the 2017-2018 Academic Year

I'm writing to alert you to changes in, and recommendations from, our Social Sciences Divisional committee, which are useful to know even if you do not have any tenure cases coming up this year in your department/school.

1. The major change this year is a revision to the role that student teaching evaluations play in establishing excellence in teaching. The Social Sciences Divisional Committee still requires these materials, but it aims to broaden the criteria by which we assess excellence in teaching, and in the process reduce the weight that the students' teaching feedback have in our deliberations. There are several reasons for this shift in emphases. First, it has long been known that students' ratings have a patina of quantitative rigor while in fact being extremely noisy measures, both across and within departments. They may be influenced by many factors that don't necessarily reflect teaching excellence, including the difficulty of the course, the quality of the classroom, time of day, etc. It is difficult to compare a candidate's ratings to those of other faculty in the same department, because other faculty typically teach a different course, of different size, level, and topic. Some departments are stacked with outstanding teachers and others less so, meaning that over-reliance on comparisons across faculty are not straightforward. Thus, while these measures might be informative about change within one instructor over time, we will not overly scrutinize absolute numerical values or comparisons across faculty in a department.

   More recently, there have arisen new data that further argue against over-interpretation of student course feedback, namely that large studies of student evaluations show systematic biases in those evaluations in ways that may disadvantage women and minorities. Some examples of evidence include:

   This article by Lisa Martin, UW Political Science professor and Associate Dean, https://faculty.polisci.wisc.edu/llmartin3/wp-content/uploads/2013/11/Gender-and-teaching.pdf

   This blog post summarizing a number of studies in the social sciences, including studies of online classes in which the instructor's actual gender was crossed with the students' belief about their instructor's gender, http://blogs.lse.ac.uk/impactofsocialsciences/2016/02/04/student-evaluations-of-teaching-gender-bias/

2. This reduction on weight of student feedback is balanced by additional

consideration of other evidence of excellence in teaching. These include:

    a.  A clear statement in the chair's letter how the candidate's teaching fits in the teaching mission of the department or unit.

    b.  The candidate's teaching statement, which should include not only a description of the candidate's approach but also relevant evidence or examples of attention to teaching excellence, such as teaching workshops or classes that the candidate attended or led, successful teaching methods implemented, changes in response to feedback, and other evidence as relevant.

    c.  Peer teaching observations, which are required for probationary faculty every year from their department or school, starting in the candidate's second year. This detailed observation conducted by a faculty member, based on both course materials and observation of teaching, should be documented for inclusion in the tenure dossier. An example of a form for peer evaluation of teaching is available on the divisional committee website, if it is useful to your department/school. The peer teaching evaluation requirement was passed several years ago, yet the committee has been assessing dossiers where this has not been followed. It is extremely useful for probationary faculty to receive feedback on their teaching so that they can make changes and improve over time. Such evaluation and feedback [make] it more likely that a probationary faculty member will have and demonstrate excellence in teaching when evaluated by divisional committee. We also will attend carefully to these evaluations.

Some other reminders from changes to the guidelines made 1-2 years ago:

"The chair's letter should note if the candidate's probationary period ("tenure clock") was extended. However, the letter should not describe the specific circumstances for those extensions unless they were for relevant professional reasons rather than for personal reasons, e.g. parental leave or sick leave."

Last year, the Social Sciences Divisional Committee has added an option for demonstrating excellence through integration of research, teaching, and service. I encourage you to read these guidelines, as some departments may wish to put language into new faculty offers, when relevant, about position expectations that are consistent with integration of research, teaching, and service. Chairs should also discuss the integrated tenure case option with relevant current probationary faculty who are not yet eligible for tenure consideration, because the decision to pursue an integrated tenure case ideally should be made and documented well before the year in which the tenure dossier is prepared.

Highlights and tips:

- The department chair's cover letter is the single most important document in the

dossier. Please take the time to write a concise yet comprehensive review according to the guidelines. More than half the cases that are denied by the Social Sciences Divisional Committee have weak chair's letters that did not provide adequate context about the nature of the discipline, the candidate, and/or articulate reasons for support for the candidate's promotion. More specifically, we expect a tenure file to explain how a particular mixture of activities relates to the department's mission, the terms of the candidate's appointment and the communication of expectations and support provided to the candidate.

- Outside reviewers should be outstanding scholars in the candidate's field. Importantly, at least five reviewers should be at arm's length from the candidate. The 2016-17 year saw several cases plagued by concerns over the degree to which letter writers were arm's length. In reviewing the tenure documents, the committee didn't see need for changes to the guidelines in response to these events, but we recommend especially careful attention to two elements of the guidelines that can have variable interpretation. These are a) the exclusion of the candidate's mentors, where occasionally letter writers use the term "mentor" to mean something other than the candidate's academic advisor (such as someone who runs a mentoring session at a conference for a group of young people in the field, including the candidate). In this case, the chair's letter can clarify that this use of "mentor" does not constitute a departure from arm's length. And b) the exclusion of letter writers who have a "personal interest" in the candidate's success or their work. Examples that may rise to the level of personal interest include situations where the letter writer has invited the candidate to contribute an article or chapter to something the letter writer is editing, even though there is no collaborative research between the letter writer and candidate. The committee is serious about seeing at least five unambiguously arm's length letters. Two tips:

  o Use the letter request template, which gives the letter writer our committee's full definition of arm's length, and explicitly ask the letter writer to indicate their arm's length status, clarifying any relationships for the committee.

  o Ask for more than the five required letters, both because some fields are seeing more requests declined, and also because departments can be surprised to find, when the letter comes back, that the writer is not in fact arm's length.

- Research collaboration should be documented indicating the effort among team members.

- All documents relevant to the submission of tenure dossiers to the Social Sciences Divisional Committee, including the updated, most recent version of the committee's tenure guidelines and template letters are available at: https://secfac.wisc.edu/tenure/social-sciences-divisional-committee/tenure-

documents/

- A number of exemplary dossiers from past successful tenure cases are available for review. To arrange time to look over these dossiers, please contact Divisional Committees Coordinator Michaela Aust (contact information below).

Finally, we anticipate a high volume of cases this year. As such, submitting cases earlier rather than later both the committee and your department. If there are too many cases for a given meeting, some cases are deferred to the next meeting. Our past experience suggests that cases submitted in fall and early winter are much less likely to be deferred because of case overload than cases that arrive later in the spring.

The committee seeks to work with departments/schools to ensure fair and judicious reviews of candidates. Please contact me (mcmacdond@wisc.edu) or Divisional Committees Coordinator Michaela Aust (Michaela.aust@wisc.edu or 263-5741) with any questions, comments, or suggestions.

<u>Appendix B:</u>

<u>Correspondence with Social Science Divisional Committee in AY 2018-19</u>

From: Susannah Tahk <<u>susannah.tahk@wisc.edu</u>>
Sent: Friday, April 5, 2019 2:11 PM
To: Michaela Aust <<u>michaela.aust@wisc.edu</u>>
Subject: Course evaluation numerical scores


Dear Ms. Aust,

This is Susannah Tahk, associate dean for research and faculty development at the Law School, writing with a question regarding divisional committees' practices…

My question concerns providing numerical scores on course evaluations to divisional committees. Based on widespread concern about reliability, validity and bias, a Law School committee has proposed eliminating the use of numerical scores on course evaluations. Instead, the evaluations would only ask open-ended questions that call for written responses. However, concern was raised that divisional committees (in our case, usually Social Sciences but occasionally Arts and Humanities) in fact require units to provide numerical evaluation scores for tenure candidates. I was wondering if the Law School could get some guidance as to this possible requirement (before we eliminate the form that collects it!).

If you would like to discuss, I'm also available via phone or in person.

Thank you very much for your assistance.


Sincerely,

Susannah Tahk

Associate Dean for Research & Faculty Development


From: Michaela Aust
Sent: Monday, April 8, 2019 1:28 PM
To: Susannah Tahk
Cc: Kristine Kwekkeboom
Subject: RE: Course evaluation numerical scores

Dear Ms. Tahk,

It was great to see you at the event as well. Your question comes timely as we are currently revising the tenure guidelines.

As you might have seen in past tenure guideline changes, the divisional committee is well aware of bias in student evaluations and takes that into consideration when evaluating tenure cases. The committee welcomes all additional information on student feedback/the candidate's teaching. All divisional committees also meet with a WISELI at the beginning of the academic year, which serves as an orientation for new members and a reminder for continuing members.

As for the total elimination of numerical scores, I would like to defer to Kris Kwekkeboom, current chair of the Social Sciences Divisional Committee (copied on this email).

With best regards,

Michaela

From: Kristine Kwekkeboom <kwekkeboom@wisc.edu>
Sent: Monday, April 08, 2019 2:11 PM
To: Michaela Aust <michaela.aust@wisc.edu>; Susannah Tahk <susannah.tahk@wisc.edu>
Subject: Re: Course evaluation numerical scores

Hi Susannah,

Thanks for your interesting and important question. My read of the tenure guidelines is that the divisional committees would like to see some kind of student evaluation of a candidate's teaching effectiveness. The current language is flexible, indicating that "systematic surveys of student opinion" are just one type of information that is *appropriate* for the teaching case. And that numeric ratings are *an example* of data that can be used to document the case. As long as the law school collects some sort of information (quantitative or qualitative) from students, and can describe what or how that information demonstrates a standard of "excellence" in teaching within your school, the divisional committee members should have the necessary information to make an evaluation.

We have a meeting on Wednesday. I'll run your question by the full membership and let you know if there is disagreement on what I've suggested, here.

Kris

From: Kristine Kwekkeboom <kwekkeboom@wisc.edu>

Subject: Re: Course evaluation numerical scores

Date: April 11, 2019 at 10:42:07 AM CDT

To: Susannah Tahk <susannah.tahk@wisc.edu>

Hi Susannah,

I just wanted to let you know that I brought your question to the Divisional Committee yesterday. Members agreed that some student evaluation of teaching quality is necessary, but it doesn't have to be numeric. Cases can be evaluated with qualitative data only.

One of the members did raise the issue that qualitative data are also subject to bias, and may not resolve your original concerns. For example, we sometimes see letters from external reviewers that describe a female candidate's work and productivity differently than a male candidate's work/productivity. Just food for thought!

Kris

<u>Appendix C:</u>

<u>Sample Evaluation of Teaching form (UW Letters & Sciences)</u>

EVALUATION OF TEACHING

| Name of Instructor | |
|---|---|
| Date | |
| Course number and name | |
| Name of evaluator | |

1. Please write a brief non-evaluative description of the class you visited.


2. Please write an evaluation of this class session. Your evaluation should include responses to the following questions:

    a.     How well was the material organized?

    b.     How clearly was it presented?

    c.     Did the instructor encourage critical thinking?

d.      Was time well utilized?

e.      Did the instructor communicate enthusiasm and interest in the subject?

f.      If discussion took place, how well did the instructor moderate it?

g.      Did the instructor have a thorough knowledge of the material?

h.      Was the material presented up-to-date?

3.  Please write an evaluation of the syllabus and any other written material (e.g., exams) that you considered. Your evaluation should include responses to the following questions:

a.      Does the syllabus set out clear learning objectives for the course?

b.      Is the syllabus well organized and well conceptualized?

c.      Does the syllabus make clear the basis for grading?

d.      Is the instructor covering the major areas that should be covered in this course?

4.  Please provide constructive criticism that the instructor can use to improve the course. Issues to consider include:

a.      Delivery methods

b.      Student interactions

c.      Types of material presented and distributed

d.      Grading and evaluation methods

General Comments: