

UNIVERSITY REVIEW COMMITTEE
Tuesday, April 12, 2016
9 a.m., Hovey 102

MINUTES

Members present: Angela Bonnell, Diane Dean (via telephone), Joe Goodman, Christopher Horvath, Doris Houston, Andy Rummel, and Sam Catanzaro (non-voting)

Members not present: Rick Boser, Sheryl Jenkins, and David Rubin

Others present: Bruce Stoffel (recorder)

I. Call to order

Chairperson Doris Houston called the meeting to order at 9:05 a.m.

II. Discussion of ASPT suggestions and requests from Faculty Caucus

Update on ASPT document approval

Houston thanked committee members for their hard work on the ASPT document recommendations this academic year. Houston reported that URC has completed all work the Faculty Caucus (“the Caucus”) has asked the committee to complete by the end of the term. She noted that issues being discussed by URC working groups will not be considered by the Caucus until fall 2016. She also noted that discussion of the proposed new articles related to faculty discipline has been deferred by the Caucus until fall 2016.

Timeline for Appendix 1

Bruce Stoffel provided context for the timelines drafted for inclusion with Appendix 1 (ASPT calendar) per the suggestion of the Caucus. Stoffel explained that timelines (see attached) have been drafted for three of the six sections of the appendix: Calendar for Promotion and Tenure, Calendar for Performance Evaluation Review, and Calendar for Cumulative Post-Tenure Review. The other three calendars in Appendix 1 either set forth dates that may vary depending on an individual faculty member’s circumstance or identify deadlines not otherwise associated with a broader process. For those reasons, Stoffel said, timelines have not been developed for the other three sections of the appendix but can be.

Houston asked if ASPT document section or page numbers can be added to deadlines displayed in the timelines to direct the reader to more information about each deadline or action. Bonnell agreed with the suggestion. Sam Catanzaro responded that cross references can be added. He suggested citing section numbers rather than page numbers because pagination of the document has not yet been finalized.

Christopher Horvath moved to recommend the three timelines, as disseminated to URC prior to the meeting but with cross references added, to the Caucus. Joe Goodman seconded the motion. The motion carried on voice vote, all voting in the affirmative.

Revised passage. Appendix 2

Houston provided context for the request from the Caucus for URC to consider rewording the last sentence of the first paragraph of Appendix 2 (“Departments/schools must consider a demonstration of quality of accomplishment and a standard of excellence as they select guidelines and criteria for evaluation”). Houston said some Caucus members consider the sentence confusing because it seems to address multiple concepts. Catanzaro agreed, asking if the sentence is about establishing guidelines for setting forth criteria for faculty excellence or if the sentence is about guiding faculty members in making a case in their faculty evaluation papers for excellence in their work.

Catanzaro offered two options for the committee to consider: deleting the sentence without replacing it or endorsing the replacement sentence suggested by Caucus Chairperson Susan Kalter:

“Departments/schools must select guidelines and criteria for evaluation that guide faculty in how to demonstrate quality of accomplishments and achieve a standard of excellence.” Horvath said that, while he thinks the replacement sentence is better than the sentence that had been recommended by URC, the replacement sentence also conflates two concepts. Houston noted a possible third interpretation of the sentence: that it may refer to departments and schools demonstrating that evaluation standards are in line with high quality performance. Horvath suggested that the sentence be deleted to eliminate potential confusion resulting from its triple conflation. He added that the concepts addressed in the sentence are implicit elsewhere in the ASPT document. Goodman and Houston agreed.

Horvath moved that the last sentence of the first paragraph of Appendix 2 not be rewritten but instead be deleted. Goodman seconded the motion. The motion carried on voice vote, all voting in the affirmative. Houston reiterated the rationale for the committee recommendation to delete the sentence: that, by conflating concepts, the sentence may confuse readers and that the concepts the sentence is intended to convey are addressed elsewhere in the ASPT document multiple times.

III. Working group reports

Goodman asked if working groups are to frame their final reports as what URC recommends or what the working group recommends. Houston clarified that the reports should be framed as the URC recommendation to the Caucus.

Report from the working group on student evaluations (Christopher Horvath and Andy Rummel)

Horvath disseminated a written interim report from the working group on student evaluations. He first reviewed the Caucus charge to the working group: 1) whether the term “student reactions” should still be used or replaced with “student evaluations” or some other term, and 2) whether a requirement should be added to the ASPT document that multiple methods of teaching evaluation must be weighted equally in the faculty evaluation process. Houston noted that the Caucus also gave URC license to consider other issues related to student evaluations should URC want to do so. Horvath then summarized the interim working group recommendations. He explained that, in developing its recommendations, the working group has considered guidance published by the American Association of University Professors (AAUP), practices at benchmark institutions, and literature on the topic.

Horvath reported that AAUP has not issued specific guidelines regarding appropriate terminology for student feedback and that benchmark institutions use a variety of terms. Horvath said the working group recommends referring to student feedback as anything but “student evaluations,” because students do not evaluate their instructors but provide feedback regarding instruction in a specific course. Horvath said the working group recommends retaining the term “student reactions” or using the term “student responses.”

Horvath reported that he and Rummel found the second charge (whether multiple methods of teaching evaluation should be equally weighted) a bit more complicated. They surmise that the suggestion was made by the Caucus to prevent student reactions from being used as the predominant method to evaluate teaching performance. Horvath noted that the resources consulted by the working group agree that teaching evaluation should be an ongoing process involving multiple methods rather than based on information from a single point in time using a single method. He said that equally weighting all modes of evaluation could impede achieving an ongoing, comprehensive teaching evaluation process. Horvath also stated that student reactions to teaching are fraught with complex biases, which, he said, may suggest that departments should be allowed to consider various factors contributing to those biases on an individual basis.

Houston asked how the working group might recommend integrating such an approach to teaching evaluation in the ASPT document, which, Houston noted, would not need to be done until next academic year. Horvath referred to the following sentence in the interim working group report: “Instead, we would suggest language that encourages schools/departments to develop methods of teaching evaluation that

take into consideration multiple sources of data in ways appropriate to the particular ... faculty member, course load, pedagogy, course content, and discipline.” Horvath suggested rewording that sentence and adding it to the ASPT document. He offered to work on the wording and to help determine where in the ASPT document the wording should be added. Horvath suggested that generating some sort of model for teaching evaluation might help illustrate this approach to the Caucus but noted that models need to be tailored to individual disciplines. Generating a single model would be difficult for that reason, he said.

Houston noted that scores from student reactions to instruction continue to play a significant role in evaluating teaching performance in many departments and schools at the University. She asked committee members how they feel about that situation. Horvath responded that referring to student feedback as quantitative data is false. He cited lack of consistency across schools and departments with respect to methodology used to administer student feedback surveys and to interpret and apply results. Goodman agreed, also noting that rating scales differ across departments and colleges.

Horvath said he will revise the working group report and resubmit it to URC. There was discussion whether the revised report should refer to the term “student reactions” or to “student responses.” Bonnell suggested that, whatever the preferred term, it should be used consistently throughout the ASPT document. Horvath suggested using the term “student responses” to be consistent with the name of a common survey instrument used at the University, the Student Response Inventory. Committee members thanked Horvath and Rummel for their work on this issue, noting both the policy guidance and professional development implications resulting from their efforts.

Report from the working group on the evaluation process (Angela Bonnell, Rick Boser, and Sheryl Jenkins)

Bonnell reported that her group is working on recommendations regarding the time involved in compiling annual portfolios for performance evaluation. She said the group has consulted AAUP publications as well as policies at benchmark institutions. She said it seems clear to group members from their research that all faculty members at the University should be asked to prepare performance evaluation papers every year. She clarified that conducting performance evaluations annually is not just about salary incrementation. Conducting evaluations on an annual basis helps convey the importance of that feedback to the professional development of faculty members, she said.

Bonnell said the working group does not recommend stating explicitly in the ASPT document what departments and schools should be doing with respect to performance evaluations, rather it should be left to each department and school to decide. She added that there may be opportunities for sharing best practices regarding performance evaluation among units at the University, perhaps through the ASPT professional development sessions facilitated by Catanzaro.

Houston reported that several persons providing feedback about the proposed ASPT document changes have questioned the current policy of preparing performance evaluation portfolios in December and conducting evaluations in January, a period when faculty members are busy preparing spring courses. She asked Bonnell if her working group encountered guidance about the timing of performance evaluations while conducting its research. Bonnell responded that, while her group noted variations across institutions regarding the timing of evaluations, the group did not discuss the issue. She added that the group is willing to do so.

Horvath said that, from a pragmatic perspective, he would rather compile his portfolio for evaluation during winter break than in May. Catanzaro pointed out that if portfolios were due in May, DFSCs would be working off contract (i.e., after the May 15 end date for most faculty contracts). Diane Dean noted that moving the evaluation portfolio due date to May could complicate decisions regarding salary incrementation and appeals. If such matters are to be decided before the fall semester, the evaluation process has to start early, she added. Catanzaro said that changing the timeline could help separate the promotion and tenure application and decision processes so they are treated as the distinct processes they are, but he is unsure whether the benefits of such a change would outweigh the costs. Horvath agreed, noting that the two processes are so closely related.

Report from the working group on salary incrementation (Joe Goodman and David Rubin)

Goodman reported that the primary concern of his working group is the dollar amount of the salary increment associated with promotion from assistant professor to associate professor and the amount associated with promotion from associate professor to full professor. He said that, while all peer institutions researched by the working group define salary increments by fixed amounts rather than by percentages, those amounts are higher than increments granted by Illinois State. Goodman said that, despite these differences, the working group prefers to be measured in its recommendations given the continuing budget crisis in the state. He asked for guidance from the committee in framing working group recommendations.

Catanzaro agreed that the issue is sensitive given the current public dialogue about the cost of higher education. Recalling that the last increase in salary increments was likely approved in the early to middle 2000s, Catanzaro suggested that the working group consult ASPT documents in the Provost's office to determine how often, when, and in what amounts changes have been to the salary increments by the University. That information might help make a case for discussion of an increase despite the financial uncertainty facing the state.

IV. Approval of minutes

Because the meeting was nearing the established time for adjournment, Houston suggested that minutes of prior meetings be reviewed and approved via consent agenda. Meeting attendees concurred. Minutes of the following meetings will be reviewed and approved in this manner: February 3, 2016; February 16, 2016; March 2, 2016; and March 30, 2016.

V. Other business

There was none.

VI. Adjournment

Horvath moved, Bonnell seconded that the meeting adjourn. The motion carried on voice vote, all voting in the affirmative. The meeting adjourned at 10:04 a.m.

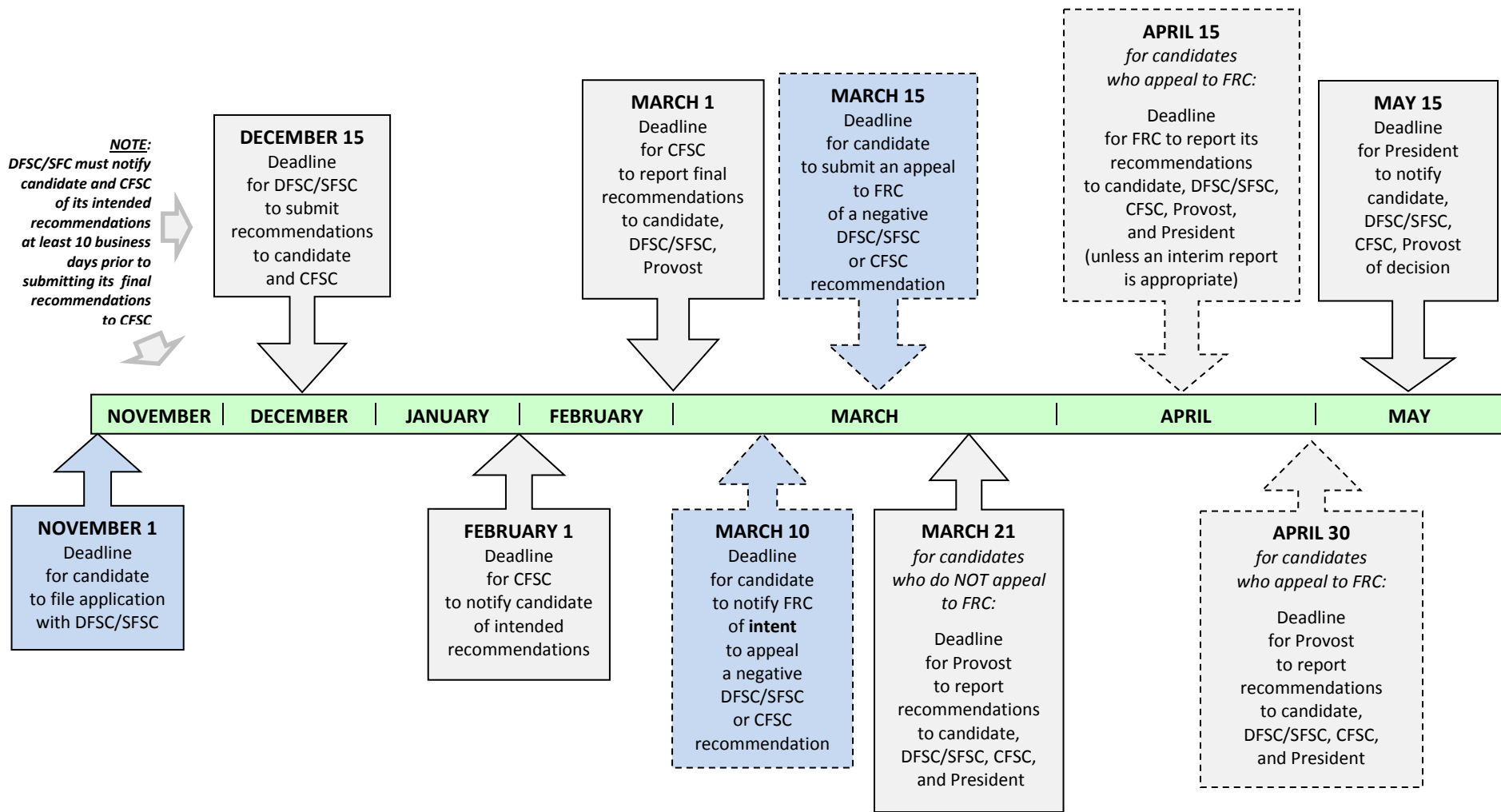
Respectfully submitted,
Christopher Horvath, Acting Secretary
Bruce Stoffel, Recorder

Attachments:

Draft Appendix 1 timelines

URC Working Group on Student Evaluations: Chris Horvath (CAS/Philosophy), Andy Rummel (CFA/Music), Date Submitted:
April 11, 2016

Timeline: Promotion and Tenure



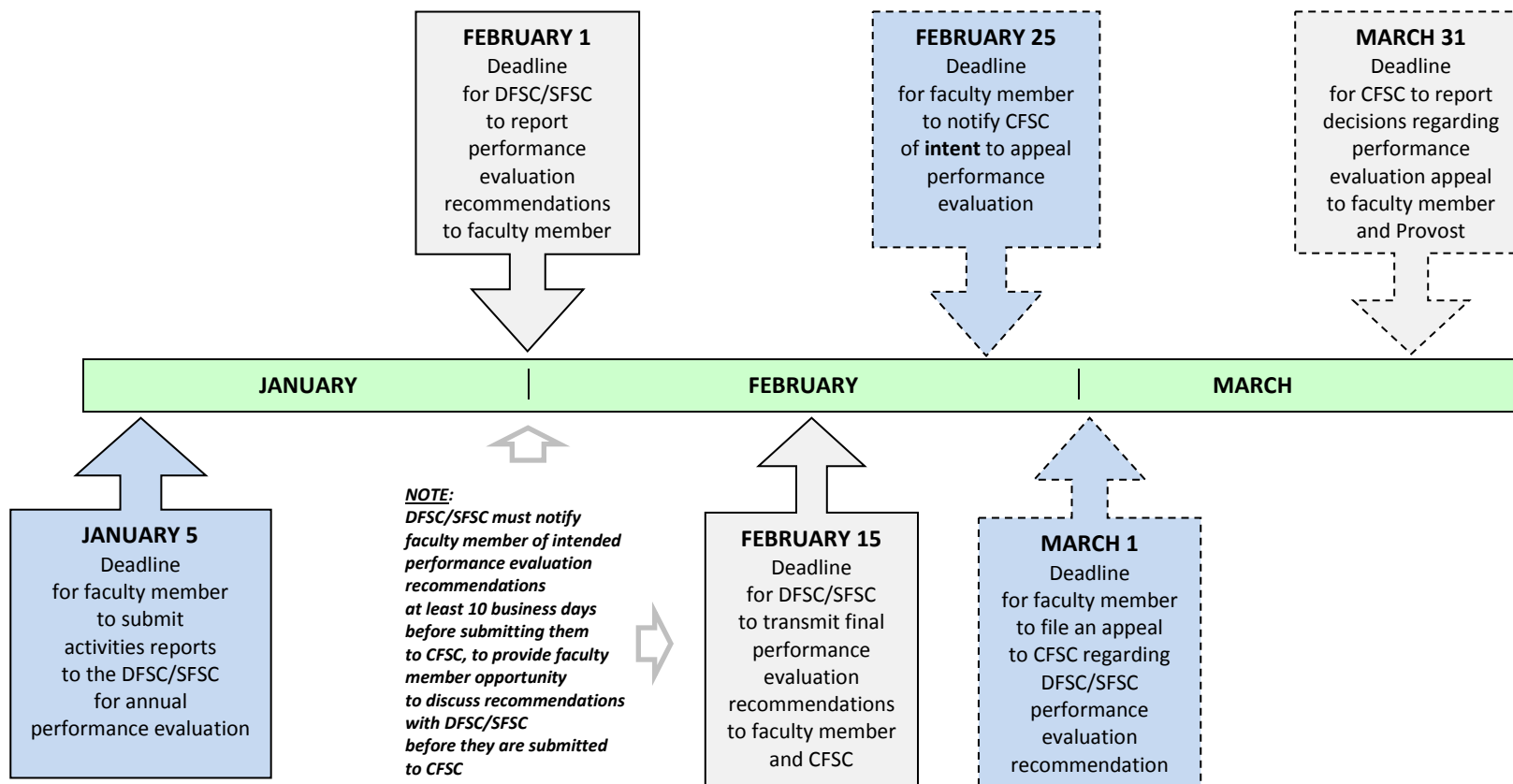
KEY

DFSC: Department Faculty Status Committee
 SFSC: School Faculty Status Committee
 CFSC: College Faculty Status Committee
 FRC: Faculty Review Committee

BLUE denotes an action initiated by the candidate
 GRAY denotes an action initiated by a committee or an administrator
 Dashed outline denotes an action related to an appeal by the candidate

This timeline is provided as a planning tool.
 For a complete description of the promotion and tenure process, including policies and procedures for appealing promotion and tenure decisions, please refer to Articles VIII, IX, and XIII of this document.

Timeline: Performance Evaluation Review



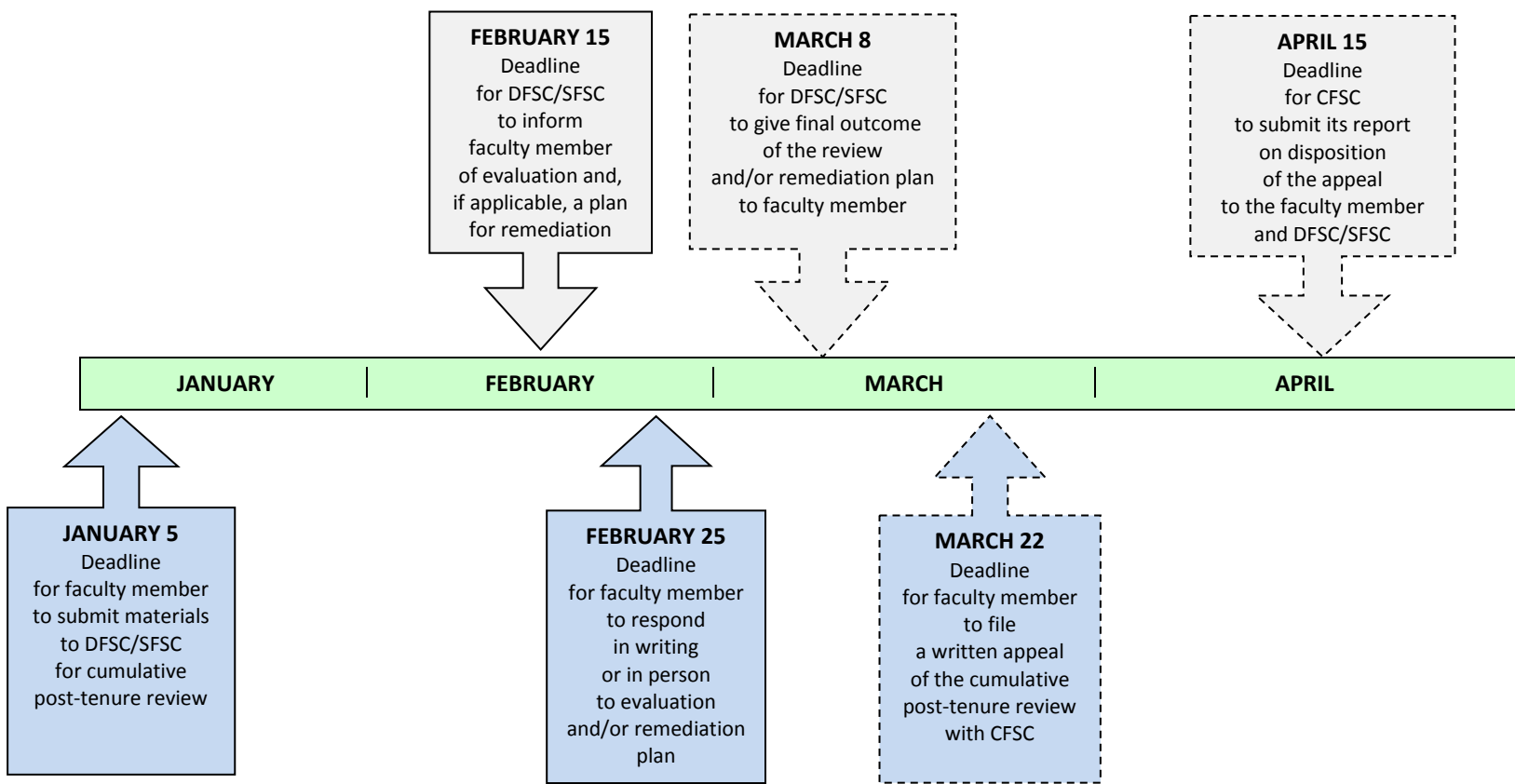
KEY

DFSC: Department Faculty Status Committee
 SFSC: School Faculty Status Committee
 CFSC: College Faculty Status Committee

BLUE denotes an action initiated by the faculty member
 GRAY denotes an action initiated by a committee
 Dashed outline denotes an action related to an appeal by the faculty member

This timeline is provided as a planning tool.
 For a complete description of the promotion and tenure process, including policies and procedures for appealing promotion and tenure decisions, please refer to Articles VIII, IX, and XIII of this document.

Timeline: Cumulative Post-Tenure Review



KEY

DFSC: Department Faculty Status Committee
 SFSC: School Faculty Status Committee
 CFSC: College Faculty Status Committee

BLUE denotes an action initiated by the faculty member
 GRAY denotes an action initiated by a committee
 Dashed outline denotes an action related to an appeal by the faculty member

This timeline is provided as a planning tool.
 For a complete description of the promotion and tenure process, including policies and procedures for appealing promotion and tenure decisions, please refer to Articles VIII, IX, and XIII of this document.

URC Working Group on Student Evaluations:

Chris Horvath (CAS/Philosophy)

Andy Rummel (CFA/Music)

Task:

The subcommittee was asked to review Article VII.B.2 (pg 57) and provide guidance to the URC regarding the following suggestions/requests from the Faculty Caucus (11/4/15):-

- (i) *Should the term "student reactions" still be used or should the phrase be replaced with "student evaluations" or some other term?*
- (ii) *Consider adding a requirement that multiple methods of teaching evaluation be weighted equally.*

The Faculty Caucus requested that the subcommittee consider both AAUP Guidelines and recent research on the use of student input in the evaluation of faculty teaching.

Review:

The subcommittee reviewed material available on-line in order to reach its recommendations. In addition to AAUP material and recent research on student evaluations, we chose to examine the practices of "Benchmark Institutions" (list attached) in order to determine "best practices" with respect to the use of student input in faculty evaluations.

The subcommittee addressed the following questions in their review.

1. What are the AAUP guidelines with respect to the use of student course evaluations in the evaluation of faculty teaching?
2. How do our "Benchmark Institutions" administer student course evaluations and how are those evaluations used in the evaluation of faculty teaching? Are other forms of teaching evaluation required for faculty evaluation? If they are required, are different modes of evaluation given equal weight?
3. What are the most recent research findings on the reliability of student evaluations as a measure of faculty performance/learning outcomes assessment? Is there evidence of systematic bias in student course evaluations with respect to female faculty, faculty of color, LGBTQ faculty, ESL faculty?

Findings and Recommendations:

*With respect to request/suggestion (i), the subcommittee recommends **retaining** the less-formal term "student reaction".*

Justifications:

- There is a great deal of heterogeneity across departments and colleges in both the instruments used to generate student feedback and in the methodology used to administer those instruments.

- Some instruments are clearly designed to elicit comments on the instructor's performance (e.g. "Was the instructor regularly late or absent from class?" "Did the instructor return graded material in a timely manner?") and others are designed to elicit feedback on the course itself (e.g. "Were the reading assignments interesting and relevant?" "Was the course well organized?")
- Some instruments use primarily open questions and others use a numerical scale. (Some departments use 5 as a positive response and other departments use 5 as a negative response.)
- Some faculty self-administer their "evaluations" with little guidance or oversight while other departments have elaborate procedures for administering and collecting evaluations.
- The AAUP has no specific guidelines regarding this issue.
- Our Benchmark Institutions take a variety of approaches. Most use the terms "course evaluation" or "instructor evaluation".
- A review of the relevant literature suggests that "evaluation" is a misnomer. The data gathered on the typical student response instruments do not provide reliable information about the quality of instructor's performance in the classroom or about the instructor's success in achieving desired learning outcomes.
- There is ample evidence of inherent bias in many student "evaluations" with respect to race, gender, sex, and sexuality. Cis-gender, white male faculty may benefit from a race and gender based "assumption of competence". Female, non-white, and non cis-gender faculty suffer the effects of the opposite assumption.

This disparity coupled with the documented problems with bias inherent in the student evaluation process lead us to suggest that the student feedback should not be considered "evaluative" in any formal sense. Rather, student "course evaluations" should be treated as an opportunity for students to provide feedback regarding their experience with a particular instructor in a particular course. Whatever we call these student "reactions" should reflect this reality.

With respect to request/suggestion (ii), the status quo seems to privilege student course evaluation. We believe the intent of this suggestion is to increase the relative importance of modes of teaching evaluation other than student evaluation. The subcommittee endorses this basic idea. However, simply requiring that all sources of data regarding teaching performance be treated equally seems to miss the real target. A review of the relevant literature and "Best Practices" suggests that the evaluation of teaching should be a holistic and on-going process not limited to a single source of data or a single day at the end of the semester. We believe requiring all schools/departments to treat all sources of input equally (i.e. treating student course evaluations with equal weight to peer review of a comprehensive teaching portfolio or peer observation in the classroom) would, in fact, impede the development of comprehensive and on-going methods of teaching evaluation. We **do not recommend** the suggested change. Instead, we would suggest language that

encourages schools/departments to develop methods of teaching evaluation that take into consideration multiple sources of input over an extended period of time and weight the various sources of data in ways appropriate to the particular the faculty member, course load, pedagogy, course content, and discipline.

Justifications:

- The AAUP has no specific guidelines regarding this issue. However, “a recent AAUP survey finds declining response rates on student reviews of professors, too many colleges that do little beyond student reviews, and concerns about bias against women, minorities and adjuncts. But association panel wants to improve system, not end it.” (Inside Higher Ed June 10, 2015)
- IDEA is a non-profit organization doing research to improve higher education. Several schools and departments on campus use “student response inventories” developed by IDEA (e.g. the College of Fine Arts) According to IDEA, “Student ratings of instruction (SRI) should be supplemented with peer review and ongoing faculty development. We were pleased to read that 69 percent of respondents see the need for student feedback about their teaching. We also agree that institutions should end the practice of allowing SRI to serve as the only or primary indicator of teaching effectiveness. IDEA has long recommended that they count no more than 30 percent to 50 percent of the overall teaching evaluation.” (IDEA June 22, 2015)
- Our Benchmark Institutions take a variety of approaches to faculty teaching evaluation. All include some form of student input. Most require additional sources of data, most often peer review of teaching material and less often classroom observation. The relative weight given to different sources of data regarding teaching performance varies significantly both between and within institutions. Most commonly, college and university level policy requires multiple sources of input on teaching performance while decisions about specific kinds of assessment required and the relative weighting of are made at the department level.
- The variety in policies and procedures at the department level within Benchmark Institutions reflects the differences in course content and pedagogy within different disciplines. These differences should be respected.
- A review of the literature reveals a persistent problem of gender and race bias in student course evaluations. This bias is most often revealed in a complex interaction of student gender, instructor gender, and course content. (e.g. Basow, 1998 and Laube, 2007.) For example, a consistent gender bias is found against female faculty who introduce (appropriately) feminist content into non-gender studies courses, though a similar negative response does not apply to male faculty who do the same thing.
- The same bias response has been demonstrated with respect to race and race-focused course content.

Selected Bibliography:

Andersen, K., & Miller, E. D. (1997). Gender and student evaluations of teaching. *Political Science & Politics*, 30, 216-219.

Explores the potentially damaging effects of gender bias in student evaluations of teaching, specifically with regard to student expectations. Reviews a number of laboratory and "real life" studies and summarizes their conclusions. Notes the different and conflicting expectations of students and recommends a broader approach to teacher evaluations.

Arreola, R. A. (2000). Developing a comprehensive faculty evaluation system: A handbook for college faculty and administrators on designing and operating a comprehensive faculty evaluation system (2nd ed.). Bolton, MA: Anker Publishing Company, Inc.

This handbook provides a practical model for developing and using a comprehensive faculty evaluating system that responds to the specific needs, concerns, and characteristics of the faculty and administration of an individual academic unit. It outlines an eight-step procedure that focuses on the determination of: (1) the faculty role model; (2) faculty role model parameter values; (3) roles in the faculty role model; (4) role component weights; (5) appropriate sources of information; (6) information source weights; (7) how information should be gathered; and (8) appropriate forms and protocols. It also examines the selection and development of forms for the student evaluation of faculty, providing samples of student rating form items is included. An appendix contains a sample faculty evaluation manual.

Basow, S. A. (1998). Student evaluations: Gender bias and teaching styles. In L. H. Collins, Chrisler, J.C., & Quina, K. (Eds.), Career strategies for women in academe: Arming Athena. (pp. 135-156). Thousand Oaks, CA: Sage.

Using a quantitative approach, Basow argues that the overall effect of gender on student evaluations is small, accounting for about 3% of variance. However, there may be significant interaction effects between gender and other context variables that may cumulatively disadvantage female faculty.

Cashin, W. E. (1995). Student ratings of teaching: The research revisited. IDEA paper No. 32. This paper attempts to summarize the conclusions of the major reviews of the literature on student ratings of teaching. It is an update of a paper by the same name published as IDEA Paper No. 20 from the Center for Faculty Evaluation and Development in 1988. Viewing student ratings as data rather than evaluations may help to put them in proper perspective. Studies have considered the multidimensionality of student ratings and their reliability and validity. They have been compared to student learning outcomes, the self-ratings of the instructor, and the ratings of others, and possible sources of bias have been studied. There are probably more studies of student ratings than of all the other data used to evaluate college teaching combined, and there are certainly enough studies to allow some conclusions. In general, student ratings tend to be statistically reliable, valid, and relatively free from bias and need for control. Nevertheless, they are only one source of data about teaching and must be used with multiple sources of data to make judgments about all the components of teaching.

Laube, H., Massoni, K., Sprague, J., & Ferber, A. L. (2007). The impact of gender on the evaluation of teaching: What we know and what we can do. *NWSA Journal*, 19(3), 87-104.

Merritt, Deborah J. (2008). Bias, the Brain, and Student Evaluations of Teaching. *St. John's Law Review* 82, 235-287.

Miller, J., & Chamberlin, M. (2000). Women are teachers, men are professors: A study of student perceptions. *Teaching Sociology*, 28(4), 283-298.

Sociology students' perceptions of their instructors' educational attainment levels are examined empirically. The authors find gender disparities: students misattribute in an upward direction the level of education actually attained by male graduate student instructors, while they misattribute in a downward direction the level of formal education attained by women, even when the female faculty member is a full professor. The misattributions are linked to the imputed statuses "teacher" for women and "professor" for men, regardless of the actual positions held or the credentials earned by faculty members and graduate student instructors. The authors suggest that a process of marginalization explains the empirical findings - a process that is attributed by others, but chosen by the self, regardless of the social and economic costs incurred.

Miller, Claire Cain. "Is the Professor Bossy or Brilliant? Much Depends on Gender." *New York Times* 6 Feb 2015.

Ratings Agency. (2016). "Students Judge Their Teachers. Often Unfairly." *Economist* 23 Jan 2016.

Simpson, R. D. (1995). Uses and misuses of student evaluations of teaching effectiveness. *Innovative Higher Education*, 20(1), 3-5.

While student evaluations of teaching performance can provide useful feedback on faculty, particularly on dimensions of course delivery, there are serious limitations. Bias and distrust are often overlooked in interpreting student ratings. An inappropriate use is in rank-ordering faculty in a department. Student evaluation data must be integrated with other sources of information on teaching quality.

Travis Russ, Cheri Simonds & Stephen Hunt. (2002). Coming Out in the Classroom ... An Occupational Hazard?: The Influence of Sexual Orientation on Teacher Credibility and Perceived Student Learning. *Communication Education* 51(3), 311-324.

Wachtel, H. K. (1998). Student evaluation of college teaching effectiveness: A brief review. *Assessment & Evaluation on Higher Education*, 23, 191-212.

This paper presents a brief review of the existing research on student written evaluations of the teaching performance of college and university instructors. First, a short historical background is given.