

Faculty Senate Minutes #388

April 19, 2012

1:40 PM

Room L.61 NB

Present (38): Jana Arsovska, Andrea Balis, Elton Beckett, James Cauthen, Demi Cheng, Kathleen Collins, Lyell Davies, Virginia Diaz-Mendoza, James DiGiovanna, Mathieu Dufour, Janice Dunham, Terry Furst, Jay Paul Gates, Lior Gideon, Demis Glasford, Laura Greenberg, Maki Haberfeld, Veronica Hendrick, Charles Jennings, Shaobai Kan, Karen Kaplowitz, Kwando Kinshasa, Anru Lee, Yue Ma, Vincent Maiorino, Evan Mandery, Roger McDonald, Sara McDougall, Brian Montes, Catherine Mulder, David Munns, Richard Ocejo, Rick Richardson, Raul Rubio, Francis Sheehan, Staci Strobl, Denise Thompson, Pat Tovar

Absent (10): Michael Alperstein, Erica Burleigh, Jennifer Dysart, Beverly Frazier, Devin Harner, Richard Haw, Tim Horohoe, Richard Li, Mickey Melendez, Manouska Saint Gilles

Invited Guests: Professors Joshua Clegg, Yi Lu, Alexander Long, Keith Markus; Provost Jane Bowers; President Jeremy Travis

Agenda

1. Adoption of the agenda
2. Announcements & Reports
3. Adoption of Minutes #387 of the April 4, 2012, meeting
4. Elections
 - a. 3 faculty members to serve on the University Faculty Senate
 - b. 3 faculty members to serve on the Dean of Students Search Committee
 - c. 4 faculty members to fill two seats on the new Food Advisory Committee
 - d. 5 faculty members to serve on the Ad Hoc Committee on International Programs
5. Update on Pathways at John Jay and CUNY-wide
6. Review of the agenda of the April 25 College Council meeting
7. Update on the revision of the JJ student evaluation of the faculty instrument
8. Invited guests: President Jeremy Travis and Provost Jane Bowers

1. **Adoption of the agenda.** Senator Strobl withdrew her agenda item, which was a proposed resolution on Pathways; the amended agenda was approved.

2. Announcements & Reports

Senator Catherine Mulder said that as a member of the Committee on Scholarships she wants to report that the number of students who are applying for scholarships is down even though the money for scholarships is there. She urged faculty members to encourage students to apply for the scholarships and she also urged that faculty members take the letters of recommendation that they write for their students very seriously. She stated that the committee members do take those letters very seriously when handing out funds. Senator Andrea Balis recommended that examples of successful recommendation letters be distributed to the faculty, with identifying information removed, so that faculty members can have an idea of what to include in the letters of recommendation that they write for students.

President Karen Kaplowitz reported that the Academic Standards Committee has decided to withdraw its proposal for the establishment of an Academic Integrity Committee and that UCASC has agreed with that recommendation; she said this is a result of the important issues raised by the Senate when the proposal to create this committee was on the College Council agenda and she particularly thanked Senators James Cauthen, Roger McDonald, and Evan Mandery for their careful reading of the CUNY Policy on Academic Integrity and the proposal to create an Academic Integrity Committee at the College.

3. Adoption of Minutes #387 of the April 4, 2012, meeting. Having been amended by Senator Strobl and by Vice President Sheehan, the Minutes were approved.

4. Elections of members of various University and College bodies and committees

a. 3 faculty members to serve on the University Faculty Senate

Elected were: Professors Angela Crossman (Psychology); David Munns (History); and Thomas Kubic (Science) to serve three-year terms as representatives of the full-time faculty; Professor Ned Benton (Public Management) to serve as an alternate delegate representing the full-time faculty; and Adjunct Professor Richard Kempter (Psychology) to serve as an alternate delegate representing the adjunct faculty.

b. 3 faculty members to serve on the Dean of Students Search Committee

Elected were Professors Jama Adams (Africana Studies); Carmen Solis (SEEK/Graduate Programs); and Liza Yukins (English/ Women's Center).

c. 4 faculty members to fill two seats on the new Food Advisory Committee

Elected were Professors Jeanne-Marie Col (Public Management) and Raul Rubio (FLL).

d. 5 faculty members to serve on the Ad Hoc Committee on International Programs

Elected were Professors Silvia Dapia (FLL); Veronica Hendrick (English); Mangai Natarajan (Criminal Justice); Richard Ocejio (Sociology); and Patricia Tovar (Anthropology).

5. Update on Pathways at John Jay and CUNY-wide [Attachment A]

President Kaplowitz distributed a letter which she had sent to the Board of Trustees on April 17 [Attachment A]. She plans to make the letter public after it has been received by the Trustees at their business addresses. There was a discussion about a letter that President Travis sent to the faculty about Pathways in which he stated that he and Provost Bowers support Pathways and in which he asked that any protests against Pathways not take place in the governance bodies of the College. A Senator questioned the permissibility of the person who will chair the body, the College Council, at which Pathways is scheduled to be voted on taking a public position about the matter before it is even debated; he said that under Robert's Rules, as he reads it, President Travis should step down as chair of the College Council for that part of the meeting as should Provost Bowers, who would normally chair the meeting should President Travis step down as chair.

President Kaplowitz explained that for those who oppose Pathways but are not sure whether to vote no or to abstain, it is important to know that an abstention carries the same weight and impact as a no vote because for a motion to pass, 16 yes votes (an absolute majority) are required at UCASC and 35 yes votes (an absolute majority) are required at the College Council; whether the other votes are no votes or abstentions does not affect the outcome. A Senator wondered if Pathways fails at UCASC the following day, might the college-option proposal be brought by an administrator before the College Council under the heading of new business. President Kaplowitz said that it could be but that she strenuously doubts that it would be.

President Kaplowitz noted that the students plan to vote in support of Pathways, as do the administrators, and, therefore, so will the HEOs who work for and report to the administrators. She noted that 60% of the seats on the College Council are held by faculty members and, therefore, Pathways cannot be approved by the College Council if they faculty do not vote for it. She suggested that it would be good to know whether the academic departments had taken a position on Pathways, whether department representatives had been instructed as to how to vote, and, if not, if how College Council member plan to vote. The following were the reports:

Senator James DiGiovanna: the Philosophy Dept has decided to either abstain or vote no.
 Senator Roger McDonald: the Political Science unanimously decided to vote no.
 Senator Andrea Balis: Interdisciplinary Studies Program is abstaining.
 Senator Mathieu Dufour: Economics: everyone voted no except for one abstention.
 Senator Janice Dunham: the Library decided to oppose by abstaining.
 Senator Laura Greenberg: Art & Music is abstaining.
 Senator Demi Cheng: the Science vote is no.
 Senator Charles Jennings: Security, Fire and Emergency Dept does not have a formal position but he will vote no.
 Senator Kwando Kinshasa: Africana Studies votes no.
 Senator Lyell Davies: Communication and Theater Arts votes no.
 Senator Raul Rubio: Foreign Languages and Literature has voted yes.
 Senator Anru Lee and Senator Terry Furst: Anthropology does not have a formal position but both will vote no.
 Senator Demis Glasford: Psychology has not discussed this but he will ascertain his department's position.
 Senator Richard Ocejo: Sociology has no formal vote or statement but he has been instructed to vote no.
 Senator Sara McDougall: History votes no.
 Senator Evan Mandery: he will vote no on behalf of the Criminal Justice Department.
 Senator Veronica Hendrick: the English Department has no formal position as of now but the consensus is a vote of no.
 Senator Catherine Mulder: she will vote no as an at-large member.
 Senator Maki Haberfeld: Law, Police Science and CJA will decide before tomorrow.
 President Karen Kaplowitz: she will vote no as an at-large member.
 Vice President Francis Sheehan: he will vote no as an at-large member.

6. Review of the agenda of the April 25 College Council meeting. Noted.

7. Update on the revision of the JJ student evaluation of the faculty instrument: Invited guests: Members of the Student Evaluation of the Faculty Committee: Professors Alex Long (Chair), Joshua Clegg, Keith Markus, and Yi Lu [Attachment B]

The committee discussed the draft of their recommendations [Attachment B] and gave an explanation for some of the revisions to the current instrument. Senate members gave input and the committee members made note of their suggestions. President Kaplowitz asked that after the instrument is revised, that the committee come back to share the draft guidelines it is developing.

The Committee members explained that their document is road map of the new, revised student evaluation of faculty they are developing. They expressed their hope that the Senate will endorse their recommendations. President Kaplowitz requested that the Committee summarize the recommendations and provide this summary in writing so the Senate can review them and vote at the next Senate meeting. This was agreed to.

8. Invited guests: President Jeremy Travis and Provost Bowers

President Travis recalled that our own Gen Ed reform at John Jay had been undertaken by the faculty and took four to five years to complete. He stated that he wished that CUNY would have taken as much time with their CUNY-wide revision of Gen Ed for the University. He said that the College came together as a community during its revision of Gen Ed and expressed his hope that Pathways will not cause a divide in the College. He stated that although he understands what he calls the faculty's "moment of resistance," the Board of Trustees is not going to change its mind regarding Pathways and he notes that the Bylaws are clear as to the Board's authority to set guidelines about the curriculum. He said that he does not want the College to lose its forward momentum that we had had with our own Gen Ed revision as well as the sense of unity the College had regarding our own process and curricular design.

President Travis also said that he wants to clear up a rumor that he has heard that he and Provost Bowers are opposed to Pathways. He stated that both he and the Provost support Pathways and that they both are in favor of doing whatever needs to be done to meet the requirements of Pathways. He also stated that he is not a fan of how we got to where we are and wishes that the University had done things differently. But what he does not want to lose is the momentum that the College had a year ago regarding Gen Ed. He said that if that was to all get slowed down it would be an unfortunate moment in the life of the College.

Provost Bowers said that she feels the same way as the President. She stated that in 2007 she sent an email to the faculty in which she started the College's Gen Ed reform when she was Undergraduate Dean. Provost Bowers said that she and the provosts throughout CUNY want more than 12 credits for the senior college option; on the other hand, she said that voting against Pathways is a vote against the 12-credit senior college option. She added that Pathways follows the model of our General Ed revision to some extent.

Vice President Sheehan said that he is not opposed to Pathways if Pathways is defined as a mechanism for enabling students to transfer more easily from one college to another. But that is not what the opposition is about; rather, what is being imposed on John Jay is the actual curriculum. He then stated that if the faculty vote yes for the 12-credit senior college option for John Jay, the Chancellor will automatically characterize that as John Jay's faculty support for Pathways, which the faculty does not support.

President Kaplowitz agreed, saying that, indeed, the Chancellor has done that by stating that there are three colleges that have voted for Pathways because they approved a college option.

Provost Bowers said that she does not see a way out for John Jay or for any of the CUNY colleges that are in the same situation. She said that her concern is if we delay, since Pathways is inevitable, our College would be left behind.

Senator Evan Mandery asked President Travis if he would be willing to support the College Council's simultaneous vote of protest against Pathways while voting for the college option. President Travis replied that he and Provost Bowers do support Pathways; he said they do not support a simultaneous vote of protest by the College Council, however much he understands the desire there may be for such a vote.

Senator Mandery said that in his eleven years at the College, he has never seen such an uprising in the John Jay faculty, such united opposition to anything. He said that he believes that if the faculty were dissatisfied with the process but thought a good result would be achieved, the faculty would vote for the College option; he added that he sees no satisfactory way to resolve this situation.

President Travis stated that the core Pathways problem in his view is the process the University took to get to the Pathway's result. President Kaplowitz stated that the substance of Pathways is entirely objectionable, including the inadequate number of credits involved and the learning outcomes, which are terrible. She added that the latest directive from Vice Chancellor Logue shows how fatuous and cynical the process and the resulting curriculum are: Vice Chancellor Logue is saying that we still must check off at least three learning outcomes for every Pathways course that we submit to 80th Street for approval but we now do not have to list them on our course syllabi.

President Kaplowitz said that she understand the difficult position both the President and Provost are in. But the faculty must have control of the curriculum and if the faculty does not vote its opposition to Pathways then the faculty is ceding responsibility for the curriculum to the administrators at 80th Street.

The meeting was adjourned at 4:05 PM.

Submitted by

Virginia Diaz-Mendoza
Recording Secretary

ATTACHMENT A

April 17, 2012

Board of Trustees
The City University of New York

Dear Members of the CUNY Board of Trustees,

I write to you as the faculty member on the CUNY Board Committee on Academic Policy, Program and Research (CAPPR). Two weeks ago, on April 2, during the most recent CAPPR meeting and during informal discussions afterward with several of the Committee's members, I came to the realization that not all the Trustees may know about the many votes of opposition to Pathways that have taken place across CUNY during the past three and a half months since January 2012, none solicited by the University Faculty Senate. Copies of these statements and other relevant statements are enclosed for your information.

During and after that same meeting I also came to the realization that not all the Trustees may know about a petition in opposition to Pathways which more than 4,700 non-student members of the CUNY community have signed onto; this petition was emailed to each of us at our CUNY email address, from which those who signed onto the petition did so. You will find my name among the signatories. The url of the petition is
http://action.aft.org/c/521/p/dia/action3/common/public/?action_KEY=3633

Please do let me know if you would like to discuss this or if you have questions. In addition to being the CAPPR faculty member, I am also the President of the Faculty Senate of John Jay College of Criminal Justice, one of the elected faculty bodies that issued yet another statement in opposition to Pathways just two weeks ago and which is among the documents enclosed.

Sincerely,

Karen Kaplowitz

Karen Kaplowitz, Ph.D.
Member, Board of Trustees Committee on Academic Policy, Program and Research
kkaplowitz@jjay.cuny.edu
John Jay: 212-237-8724

cc. Chancellor Matthew Goldstein
Executive Vice Chancellor Alexandra Logue
Vice Chancellor Jay Hershenson

ATTACHMENT B

DRAFT

Blueprint and Specifications for Revised Student Evaluation of Faculty Form
27 February 2012 (Version 6)

Student Evaluation of Faculty Committee

Alex Long, Chair

Joshua Clegg

Yi Lu

Keith A. Markus

Navila Abbas

Emiliya Abramova

Blueprint and Specifications for Revised Student Evaluation of Faculty Form

This document presents a blueprint and specifications for the revised student evaluation of faculty instrument. The first section summarizes the background and context. The next section presents the blueprint. The final section presents the item and content specifications. The term *item* refers to both rating scales and prompts for written responses. This document is intended to serve three main purposes: (1) It specifies how the items for the new form should be written. (2) It provides a means of obtaining feedback and, ideally, reaching a degree of consensus on the design of the new form before the items are written. (3) It documents the form design and content specifications, providing a source of information for evaluating the validity of the evaluation form in the future.

Background and Context

During the Fall 2010 semester, the Student Evaluation of Faculty (SEOF) committee began work on revising the current SEOF instrument, which was introduced in Fall 1999. This was done in response to a charge from the College Council. Shortly thereafter, the Faculty Senate reaffirmed four basic purposes served by the SEOF process: (1) providing information about expectations for new faculty, (2) providing feedback from students to faculty, (3) providing information for personnel decisions involving adjunct faculty, and (4) providing information for personnel decisions involving full-time faculty¹. As a way of informing the revision process, the SEOF committee conducted a survey of SEOF stakeholders and developed recommendations on the basis of an analysis of survey data. For the purposes of the survey, the SEOF committee identified students, junior faculty, senior faculty, and Personnel & Budget committee members as primary stakeholder groups. Most of the concerns raised by survey respondents involved the SEOF process rather than the questions on the form itself². The present blueprint for a revised form can thus only address a narrow subset of the concerns raised by stakeholders.

Broad Design Issues Shaping the Blueprint

This section covers the following broad design issues: (a) compatibility with paper and online delivery, (b) uniformity across traditional and online courses, (c) specificity of prompts, (d), dimensionality of ratings, (e) inclusion of demographic items, (f) student learning, and (g) rating scales.

Compatibility with Paper and Online Delivery

At present, traditional courses are evaluated using a paper form completed in class. A separate committee developed a procedure for using the same questions (with one minor modification for applicability) in an online format for online courses. For reasons described in the next section, the committee considers it essential that the same questions be used for both

¹ See Faculty Senate resolution for exact wording.

² These are summarized in a separate report on the survey results that will be made available through the College intranet ("Inside John Jay").

types of courses. As a result, the blueprint is written to be compatible with both paper and online delivery of the evaluations.

Uniformity of Evaluation Across Traditional and Online Courses

The item specifications presented in the section of the same title stipulate that all items should be worded to apply equally well to both traditional and online courses. Maintaining this equivalence has practical advantages with respect to maintaining a single SEOF process and handling intermediate cases such as hybrid courses. However, the stipulation is also motivated by a more fundamental principle. The third and fourth purposes of the SEOF process involve personnel decisions for faculty members. The use of SEOF results in personnel decisions constitutes high-stakes evaluation. In contrast, the second purpose – providing feedback – constitutes low-stakes evaluation. Although there are many opportunities to tailor assessment to the unique aspects of online learning for purposes of low-stakes assessment (Middle States Commission on Higher Education, 2011), assessing different faculty using a different SEOF instrument compromises the use of the ratings for personnel decisions. It is the view of the committee that under no circumstances should the low-stakes use of the SEOF ratings for outcomes assessment be allowed to compromise the high-stakes use as a form of personnel evaluation.

A fundamental principle of personnel evaluation states that an evaluation tool must be fair to all employees in the sense that it measures the same thing for each of them and is as free as possible from potential sources of bias (AERA, APA & NCME, 1999; Joint Committee on Standards for Educational Evaluation, 2009; Society for Industrial and Organizational Psychology, 2003). The use of a different SEOF form for different types of courses introduces potential bias by holding faculty to different sets of standards in different courses. This source of bias is exacerbated by the fact that teaching assignments vary across faculty and the fact that teaching assignments may vary across various groups protected by equal opportunity employment concerns. Consequently, to minimize the potential for bias in the SEOF process, the committee has produced a design for the form that will apply equally to all types of courses.

Specificity of Prompts

A second concern that arose in the survey results involves the specificity of the prompts for the written responses. One suggestion was to have separate written-response prompts for each item. Implementation of that suggestion within the context of paper-and-pencil administration, however, would require changes to the basic layout of the form that would likely dilute the quality and specificity of open-ended responses as well as increase production and scanning costs. As a result, the present blueprint addresses this concern (and others) by clustering rating items into content domains and providing separate prompts for each domain.

Dimensionality of the Ratings

One concern raised in the Spring 1999 validation report of the current SEOF form is that respondents do not seem to distinguish the various rating dimensions and this makes the form less effective with respect to providing useful feedback to faculty (Markus, 1999). A similar theme arose in the survey results. Clustering items into content domains offers a strategy to address this concern. The hope is that such clustering will make more salient to students the fact that they are rating different dimensions of the faculty members. The result will ideally be

greater statistical separation between responses in various domains, and thus more valid ratings for specific dimensions of faculty performance. The survey also elicited some suggestions for broadening or refining the content of the SEOF form. These suggestions are reflected in the design presented below.

Demographic Items

The original version of the current form contained student demographic items. These were used in the validation analyses reported in Spring 1999 in order to provide empirical assessments of various possible sources of bias in the ratings. Technically, a measurement tool is considered invariant across groups if the relationship between the scores and what they measure remains the same across groups. Demographic items provide group variables for such analyses. When the form was approved by the College Council, it was agreed that these items would be dropped after the first administration because a senior administrator felt that the form should not "evaluate students" but only "evaluate faculty" and that the demographic items constituted evaluation of students. Student members of the SEOF committee have strongly advised that these demographic items be returned to the form in order to facilitate continued monitoring of the effective functioning of the SEOF evaluations. The demographic items are included in the blueprint presented in the demographics subsection of the content specifications. This includes an expected grade question in part because the issue of the relationship between grades and ratings came up in the survey responses.

Student Learning

If one were to design an instrument for the purpose of measuring student learning, self-report would be about the worst possible strategy for obtaining valid and reliable assessments. Nonetheless, student learning is a central outcome of good teaching and this makes the inclusion of learning in the SEOF ratings attractive. However, student learning is a complex multidimensional construct. One aspect of student learning is determined by the curriculum. How much students learn is in part constrained by the course objectives, the course level, and the function of the course in the broader curriculum. Moreover, learning comes in different kinds including various forms of knowledge acquisition, cognitive restructuring and elaboration, skill development and professional socialization. The relative proportions of these various types of learning vary among different courses. These differences among courses produce potential sources of bias that could result in an uneven playing field working in favor of professors who offer certain kinds of courses and against others. Further there is no standard metric for amounts of learning that would facilitate obtaining ratings on a common metric across courses. Independent of the various factors that impact judgments of amount of learning, however, student satisfaction with opportunities to learn offers a construct more conducive to establishing comparable ratings across courses and sections. Moreover, the primary strength of student evaluations of faculty in comparison with other assessments is that they give voice to students' perspectives with respect to the quality of faculty and the courses that they offer. In light of these considerations, a section on student satisfaction offers an attractive option for augmenting the content of the SEOF ratings.

Rating Scales

Finally, it was recognized at the time of the previous revision that the scale response options are not optimal, and indeed constitute a significant weakness of the form. The problem

with agree/disagree Likert scales is that they do not fix a common frame of reference with respect to how much agreement or disagreement aligns with each numeric point on the scale. The use of verbal anchors for each response helps, but remains ambiguous. This introduces a possible source of error variance into the ratings.

The technically ideal solution would be to conduct a critical incidents study to produce behavioral anchors (Flanagan, 1954). However, the committee considered and dismissed this approach early on for lack of time and resources to conduct the required study. Although in many ways optimal, critical incidents is a very labor intensive process.

A second option would be to make the ratings norm referenced, asking students to compare the professor to other professors. The committee rejected this option for several reasons. First, students vary in their experience with different professors, depending upon how many courses they have taken, their division (undergraduate versus Masters), their majors, and a variety of other factors. As a result, the yardstick by which ratings are made would vary depending upon the students a professor teaches. This is especially acute for professors who teach freshmen courses because such students may have only had a handful of college courses when they complete the form. Second, it would be necessary to specify in the instructions the desired comparison group (all professors, all professors at this college, all professors within a fixed number of terms, all professors for courses in the major). The optimal choice of comparison group is not obvious. All choices have potential problems, and it is not clear that all students have access to any one comparison group (e.g., freshmen, transfer students). Third, the norm referenced approach is similar to grading on a curve: It presumes that a fixed proportion of faculty members should receive negative evaluations. There is no basis in fact to support such a presupposition. As such, in a sense, norm referenced ratings would be invalid a priori, simply on the basis of these presuppositions.

A third option is to construct a criterion referenced rating system that does not require the investment in time and resources that critical incidents do. This seems to offer the best compromise between the quality of the resulting scales and the resources required to produce them. Most such methods were developed for knowledge tests with correct and incorrect answers. A modification of Angoff's (1971) method for setting cut scores offered a good fit to the present purposes. The modified method can be outlined as follows: (a) A range of behavioral anchors are sorted into a linear ordering from low to high. (b) Each member of the committee identifies the behavioral anchors that reflect cut points between unacceptable teaching, below average teaching, above average teaching, and exemplary teaching. (c) The resulting ordinal values are averaged across committee members. (d) The resulting averages are used to determine the behavioral anchors included on the final rating scale (e.g., as 2, 4, and 6 on a 7 point scale).

The SEOF committee considered all of these (and other) basic form construction issues and the result of these deliberations can be seen in the blueprint for the revised SEOF form, described below.

Form Blueprint

A test blueprint summarizes the overall distribution of content across items on a measurement instrument (Crocker & Algina, 2008). The blueprint below contains three columns.

The first column contains the four broad content domains and also a category of demographic items. The second column contains the number of rating items for each domain. The third column indicates our plan to include one open-ended prompt for each domain. The content domains are discussed in more detail in the content specification section.

Domain	Rating Items	Prompts for Written Comments
Direct Instructional Activities	4-5 items	1 item
Course Design and Implementation	4-5 items	1 item
Interaction with Students	4-5 items	1 item
Student Satisfaction	4-5 items	1 item
Overall Assessment	3 items	1 item
Student Demographics	5 items	None

The above blueprint results in a form containing between 27 and 31 scaled items, 5 open-ended items, and 5 demographic items. By contrast, the current form contains 15 scaled items, one open-ended item, and no demographic items. The current form was shortened to its current length to address an administrator's concern that the SEOF process not take too much class time. It is the consensus of the committee that the time currently allotted suffices to complete a form of the proposed length. If the timeline before adoption permits, completion time can be tested empirically on a sample of volunteers.

If necessary, the additional items can be accommodated by using legal-sized scan forms and/or by arranging the form in landscape instead of portrait orientation. We propose that the open-ended items should appear immediately after the corresponding set of scaled items on both sides of the form. However, if scanning a two-sided form is cost prohibitive, the current arrangement of scaled items on one side and open-ended items on the other side could be retained as an alternative. Should this second alternative be necessary, the front side should contain sentences after each section referring the student to the open-ended item on the back.

Item Content Specifications

Test specifications delineate both the content and format of the items on a test (Schmeiser & Welch, 2006). This section presents content specifications first, following the above blueprint. Format specifications are presented separately for rating items and written response prompts. The rating items reflect three broad content domains and a domain of overall evaluations. Some possible topics, organized by domain, are listed below. Asterisks (*) indicate topics derived from items on the current form and daggers (†) indicate topics suggested in survey responses.

We wish to emphasize that the topics listed are not sample items. Multiple items could be written for one topic. Moreover, writing items will require careful attention to clarity and ambiguity beyond the brief topic labels listed below.

Direct Instructional Activities

This domain includes all activities through which the faculty member presents course material. This applies equally to activities such as lecturing in a traditional classroom and textual or video presentation in an online course.

- * Lesson organization.
- * Presentation of course material.
- * Use of instructional time.
- * Clarification of difficult points.
- * Thorough knowledge of the subject³.
- * Encouragement of reasoning.
- * Handling of classroom questions and comments.

Course Design and Course Mechanics

This domain includes activities involved in designing a course and completing activities associated with a course that extend beyond direct instructional activities. Course design involves choices about course content, course materials and student assessments. Course mechanics involve non-instructional activities such as classroom management, grading, and responsiveness to communications outside of class.

- * Classroom management.
- * Fair grading.
- * Timely grading.
- † Timely response to email and voice mail.
- † Appropriateness of reading.
- † Use of technology (if applicable).

Student Relations

This domain includes items assessing how well the instructor maintains helpful and professional relationships with students in various contexts related to the course. As with the other domains, these items will be worded in terms of observable behaviors.

- * Demonstrates enthusiasm about teaching.
- * Respectful treatment of students.
- * Encourages different points of view about subject matter.
- * Motivates students.
- † Professor demeanor toward students.
- † Responds constructively/helpfully to requests for help outside of class.

Student Satisfaction with Course

This domain involves student ratings of student satisfaction with various aspects of the course provided by the instructor. The goal of this section is to allow students to rate more

³ Survey feedback suggested that this aspect of instructor performance is not something that students can accurately evaluate. If this feedback is correct, then this would not be an appropriate topic for a rating item. This topic is also potentially problematic in that it might discourage faculty members from acknowledging that they do not know the answer to a question asked in class (in a traditional course). At the same time, the original concern behind the question is clear enough and an alternative way to address that aspect of faculty teaching may be possible.

holistic aspects of the instructor's course delivery that extend beyond the teaching behaviors rated in previous sections. It is possible for two faculty members to engage in similar behaviors but with sufficient differences that they produce different learning opportunities for students. This section is intended to address that possibility⁴.

Satisfaction with selection of course content.

Satisfaction with opportunities to develop critical analysis skills.

Satisfaction with course design.

Satisfaction with course execution.

Overall Evaluation

This domain involves overall evaluation of the professor. This is not a scale summed over sub-topics. The student directly provides an overall evaluation.

* Instructor overall effectiveness.

† Student would take the professor again.

† Recommend course to other students.

Demographics

Whereas rating items are numbered with Arabic numerals, demographic items are lettered with Roman letters. Demographic items come last on the form. The instructions will inform students that these items will be used for statistical purposes only. The basic goal for this section is to make it possible to test the assumption that the rating form works similarly for different students and different courses. The objective is that, to the extent possible, course assignments should not favor some professors and disadvantage others. The specific demographic items have been selected to address common hypotheses from the research literature and common concerns about course evaluation ratings. However, these are not set in stone. The specific demographics could be rotated or modified on an ad hoc basis to respond to specific concerns as they arise.

Class Standing & Division

GPA

Expected grade

Course load

Reason for taking course

Employment: Hours worked per week

Item Format Specifications

This section describes the formatting of ratings item stems and response options as well as the formatting of prompts for written responses.

Rating Items

⁴ The committee has considered measuring the frequency with which satisfactory course elements are delivered as a means of providing a criterion referenced rating scale. Another option is to ask directly about satisfaction, but this option does not allow for criterion referenced scaling.

The general instructions on the current form read "Instructions: Fill in the oval that most accurately represents your view about the statement. If you are unsure of your evaluation or if the question is not applicable, then leave the question's response blank." The new form will retain these instructions.

Each rating item has its own unique ordinal item number and begins on a new line. Each item consists of a stem and a set of standardized response options.

Item stems are simple declarative statements. No conjunctions are used to avoid ambiguity. Each item stem describes a specific activity of the professor in terms of a single attribute. Statements are in present tense. Statements express a consistent and persistent property of the professor's actions. Statements should use descriptive action verbs: These are verbs that describe rather than interpret and refer to an observable action rather than a mental state (Semin & Fiedler, 1988). For example, "The professor arrives at class on time" uses a descriptive action verb. "The professor encourages promptness by example" is an interpretive action verb and "The professor takes promptness seriously" is a state verb. The latter two depend upon students' interpretations of the professor's intentions and mental states respectively. The descriptive action verb does not. Similarly, "The professor treats students respectfully" is a descriptive action verb whereas "The professor shows respect to students" and "The professor respects students" are not.

Each item must make a statement that the student is able to evaluate based upon observations derived from his or her participation in the course.

Each item must apply to all courses so that every professor evaluated is expected to do what the item describes, irrespective of the course.

Item wording should avoid making the gender of the instructor salient through the use of gendered pronouns.

Item responses utilize a 9 point Likert scale. In all cases, responses are numbered sequentially from left to right with lower numbers representing less favorable evaluations. The current form uses a 5 point scale from Strongly Disagree to Strongly Agree with no verbal anchors for the middle three points. This format has the disadvantage that different students can interpret the responses differently from one another. To address this issue, the new form will use a criterion referenced scale system. Ideally, this will take the form of item-specific behavioral anchors as described above. Alternatively, items asking about specific teaching behaviors could ask whether the instructor performs these behaviors (1) Never, (2) Less than 40% of the time, (3) 40% of the time, (4) 50% of the time, (5) 60% of the time, (6) 70% of the time, (7) 80% of the time, (8) 90% of the time, (9) Always. This response scale is deliberately skewed to reduce the likelihood of cases bunching up at the top of the scale due to a ceiling effect. This will maximize the informativeness of the resulting ratings.

The current form relies only on the general instructions to indicate to students that they should skip items that they cannot evaluate or that do not apply. Space allowing, the new form will include a response option for "Insufficient opportunity to observe" to make this option more salient.

Written Response Items

Each prompt invites the student to provide open-ended comments on one of the four content domains. The prompts are constructed to be as parallel as possible. As noted in the blueprint section above, if possible each open-ended item will appear directly after the corresponding scaled items.

The item prompts must be specific enough that responses will be both concrete and comparable. For example, a prompt like "Additional comments" could elicit responses that are vague and overly general as well as completely unrelated to the domain being evaluated. In contrast, prompts like "Please use the space below to comment on how well your instructor organized and administered the course" or "Please use the space below to comment on, or provide additional information about, your responses to the previous 5 questions" are more likely to produce domain-specific responses.

The item prompts must be general enough that student responses can include concerns or observations not directly addressed in the Likert-style items. For example, prompts like "Please use the space below to rate the timeliness of your instructor's feedback" or "Do you feel like you were fairly graded?" would likely constrain student responses and so defeat one of the primary purposes of open-ended items – namely, to provide an opportunity for respondents to comment on elements of instructor performance not accounted for in the design of the forced-choice items.

Item Response Characteristics

It is undesirable to construct items with means too close to either extreme because such items fail to make distinctions between professors who fall at or beyond the extreme in question. As a result, it is highly desirable to generate items with means that fall in the middle third of the scale. Historically, SEOF forms used at the College have suffered from ceiling effects with too many ratings clustered at the top of the scale. Behavioral anchors may help address this.

Items should generate consistent ratings across multiple observers. Some variation is expected as a result of differences between students. However, items with smaller standard deviations across students within the same section are desirable. Ideally, these standard deviations should fall below 1 so that 50% of the students rate professors within the same 2-point range (e.g., between 3 and 5 on a 7 point scale).

Items should effectively distinguish different sections from one another. As a result, large standard deviations in mean section ratings across sections are desirable. Items with too little variation may require replacement or revision.

Items should cluster together by content area. Correlations between items should be higher within content clusters, consistent with a single dimension within clusters, but clearly not unidimensional across clusters. All items should share variance with other items in their cluster (e.g., high factor loadings, high item discrimination parameters).

Constructed response items should produce responses that are related to the question (no more than 10% off topic), that are specific and useful (e.g., 50% offer actionable feedback on a specific aspect of teaching or course design), and that produce a sufficiently high response rate (e.g., at least 30% of students make a response).

References

- American Educational Research Association (AERA), American Psychological Association (APA) and National Council on Measurement in Education (NCME) (1999). Standards for educational and psychological testing. Washington, DC: AERA.
- Angoff, W. H. (1971). Norms, scales, and equivalent scores. In R. L. Thorndike (Ed.), *Educational Measurement* (2nd ed., pp. 508–600). Washington, D.C.: American Council on Education.
- Crocker, L. & Algina, J. (2008). *Introduction to classical and modern test theory*. Mason, OH: Cengage.
- Flanagan, J. C. (1954). The critical incident technique. *Psychological Bulletin*, 54, 327-358.
- Joint Committee on Standards for Educational Evaluation (2009). The personnel evaluation standards: How to assess systems for evaluating educators. Thousand Oaks, CA: Corwin.
- Markus, K. A. (1999). Validation report for the student evaluation of faculty: Spring 1999. <http://inside.jjay.cuny.edu/docs/Student%20Evaluation%20of%20Faculty%20Committee.pdf>
- Middle States Commission on Higher Education (2011). Distance education programs: Interregional guidelines for the evaluation of distance education (online learning). Philadelphia, PA: Author.
- Semin, G. R. & Fiedler, K. (1988). The cognitive functions of linguistic categories in describing persons: Social cognition and language. *Journal of Personality and Social Psychology*, 54, 558-568.
- Schmeiser, C. B. & Welch, C. J. (2006). Test development. In R. L. Brennan (Ed.), *Educational measurement* (4th ed., 307-353). Westport, CT: Praeger.
- Society for Industrial and Organizational Psychology (2003). Principles for the validation and use of personnel selection procedures (4th ed.). Bowling Green, OH: Author.