

1. Introduction

The 2020-2021 Teaching Evaluation Committee set out to review the Departmental policies on teaching evaluation, as required in the Faculty and Professional Personnel Handbook (§44.6) “*The Teaching Evaluation Committee will review these procedures at least every three years and will review and report on student rating forms and departmental policies.*” The goals were to:

- confirm that each department was in compliance with having at least two methods of evaluating teaching,
- collect and summarize information about how each department evaluates teaching,
- make recommendations and suggestions about the evaluation of teaching for departmental policies

The Committee’s process began with a request for departmental policy on teaching evaluation and the creation of a policy review guide (Appendix A). Subcommittees were formed to review the policies by college. When the subcommittees reported their findings, the committee discussed the implications of the review (e.g., whether the department was in compliance, whether the policy might serve as a model) as well as any patterns within the college. Based on the completed policy reviews and the committee’s discussion, this report was drafted.

The committee reviewed documents from 43 departments. Most of the departmental policies clearly listed at least two methods to evaluate faculty teaching, with teaching observations (i.e., peer or chair review) and review of teaching materials being the most common second form. From the provided documents, it was unclear if the policies from some departments are in compliance with the Handbook’s (§43.2.2) requirement of having at least two methods of evaluating teaching. Additionally, from what was provided to the Committee, the policies of many departments did not clearly give the choice of the second method of evaluation to the individual instructor, as stated in §43.2.2. The Committee is concerned about the potential non-compliance to this important policy in the Handbook.

Nearly all of the documents about the evaluation of teaching described **summative assessment** processes. In other words, the evaluation’s goal was to assess the previous teaching (looking backward). **Formative assessment**, on the other hand, has a goal of improving future performance. When a teaching evaluation includes a formative function, it has the goal of improving the subsequent or future teaching practice. An evaluation of teaching which includes formative assessment would better align with the university’s commitment for “allowing faculty to develop to their full potential as teachers” (Handbook §43.1) to high-quality teaching.

The Committee did not have documents from every department when they completed their review. In some cases, the department told the Committee that the document was in the process of being approved. In other cases, there was no response to the Committee’s request. The lack of response may reflect no policy is needed or that the request was overlooked. A follow-up review should include the policies for these seven departments or obtain an understanding of whether the departments are justified in not having a policy on teaching effectiveness.

Section 2 of this report provides a narrative summary of the findings of compliance and other themes found in the provided documents, organized by college. Section 3 offers recommendations, informed by research, for departments about their policy on teaching evaluation, which is followed by suggestions to departments about potential improvements to their policy. Section 4 offers a summary list of the report. Section 5 describes issues which likely should be addressed by Faculty Council, namely, identifying a procedure for departments who appear to be out of compliance with the Faculty Handbook and clarifying that statistical summaries are based on the number of responses.

2. Summary by College

College of Architecture and Planning

Of the four departments in the college, only one department, Landscape Architecture, is clearly in compliance with the Handbook on evaluating teaching by listing peer-review as a second method of teaching evaluation. This peer-evaluation also offers formative assessment, diagnostic-focused open-ended comments for the instructor private use, which are not included in the instructor's file.

In Architecture, Urban Planning, and Construction Management/Interior Design, faculty are evaluated yearly seemingly only by student ratings of instruction. There is a Chair review that includes a review of course materials, but that seems to be only for faculty applying for tenure or promotion.

Further, the department policy for the department of Urban Planning and Architecture specifically mention a requirement for an average of 3.0 on the student completed course evaluations. If the average is below 3.0 average, the faculty member must submit a letter of explanation and a teaching portfolio. Since the letter and teaching portfolio are only required by some faculty in response to course evaluation ratings, this policy does not seem to be in compliance with the Handbook, based on the documents reviewed.

College of Communication, Information, and Media

Each department has two measures of teaching evaluation. Of note, the policy document for Journalism likely needs to be updated because it references pencil and paper forms for the students to complete. In the department of Telecommunication, faculty can choose the second measure (portfolio, observation). Currently, the department policies do not seem to include a formative feedback process.

College of Fine Arts

All three departments listed at least two methods, with the second method being peer-evaluation (observation). In this college, two departments also included the ability to attract students or student success as a possible third method for two departments, which may be a college specific method. The policies in this college seem to acknowledge the limitations of course evaluations.

College of Health

Five departments in this college clearly had two methods of evaluation of teaching in their documents. After course evaluations, the second methods included, peer- or chair evaluation (4), teaching portfolio (2), and invitation of survey/questionnaire/pre-post testing (1).

One department, Nursing, did not have a clear teaching evaluation policy with two measures of teaching evaluation, based on the documents provided. The provided documents did identify a minimum threshold of 2.5 average student rating as part of the minimum acceptable performance. No documents were provided by Military Science.

College of Sciences and Humanities

In this college, department policies include a second (and even a third) method of the evaluation of teaching. Observation of teaching (seven departments) and a review of teaching materials (two departments) were the most common second method of evaluating teaching. [It should be noted that the Committee was given the peer-review policy from the department of Philosophy and Religious Studies, which the Committee assumed supplemented the university required student-completed course evaluations.] In five departments, the faculty member has the choice of either teaching observation or review of materials. Although the teaching evaluation policy from the Department of Computer Science gives the faculty member a choice of chair- or peer-evaluation, the chair may not approve of the faculty member choices for peer-evaluation. The Department of Computer Science annual review document

identified a 4.0 threshold on “a majority of the instruction related questions” of the student completed course evaluations in order to be considered meritorious teaching.

The policies of multiple departments (e.g., Psychological Science, Anthropology, and Sociology) suggested that faculty members will receive formative feedback.

The committee did not review the policies from Physics and Astronomy, Political Science, or Women’s Studies.

Honor’s College

The Committee was given a document for non-tenure track faculty. The peer-evaluation policy is described in the Appendix, but, unfortunately, the Appendix was not included. Assuming that the peer-evaluation process supplements the student completed course evaluations, Honors College would be in compliance with the Faculty Handbook on this issue, but the Committee did not have enough information to complete our review.

Miller College of Business

Of the seven departments in the Miller College of Business, three departments, Economics, Management, and Finance and Insurance, seem to have two methods to evaluate faculty teaching in compliance with the Faculty Handbook. The departments of Finance and Insurance and Management both recommended portfolios and observation. Based on the documents the Committee reviewed, it appears that the department policy from Marketing and Information Systems and Operations Management did not have a clear second method of teaching evaluation (such as peer/chair observation, portfolio). Finally, the department policy for Accounting and for Applied Business Studies were not reviewed by the Committee.

Teachers College

Of the six departments in the college, five departments, Early Childhood, Youth, and Family Studies, Educational Leadership, Educational Psychology, Elementary Education, and Special Education, clearly had at least two methods of evaluation. From the documents provided, it was not clear that the policy from Educational Studies included a second method of teaching evaluation. The provided document listed procedures for the evaluation of teaching for the department which was solely about student completed course evaluations, and then the document provided the policy for the College, which states that each department will collect additional data. No department-specific second method was identified.

In this college, the most common second method of teaching evaluation was teaching observation, with four departments offering that method. Importantly, the policy from the department of Early Childhood, Youth and Family Studies states that “the goals and potential concerns of the faculty member and the Department Chairperson” should drive (not only inform) the evaluation of teaching.

The evaluation of teaching in the department of Educational Leadership included thresholds of 3.5 or 3.0 on the rating forms in the points system for merit.

Intensive English Institute (IEI)

The teaching of the faculty in the IEI is evaluated with observations (for each course taught), teaching portfolio review, and student responses on the course evaluation forms. The course evaluation form results are aggregated by faculty.

3. Recommendations for Departments

Departmental policies on the evaluation of teaching must meet the obligations outlined in the Faculty Handbook. Additionally, these documents should reflect the University's values outlined in the Inclusive Excellence Plan. Based on the review of the departmental policies, the committee offers the following recommendations:

1. Ensure that department policy is in compliance with the Faculty Handbook §43.2.2, which states that the evaluation of teaching must include additional methods beyond student course evaluations. The Handbook (§43.2.1) states that the department procedures should be approved by its college. This Committee's review, however, found that some policies may be out of compliance.
2. The Handbook (§43.2.2) states that the peer review of teaching, chairperson review of teaching, and peer and chairperson review of a teaching portfolio are all "available to the individual faculty member." Few departments (e.g., Anthropology, Communication Studies, Journalism, History) clearly gave the faculty member choice on the second method of teaching evaluation. And, even fewer listed a possible third method of teaching evaluation. The documents submitted by several departments (e.g., Mathematical Sciences) appear to be out of compliance with the Handbook because the faculty choice of the second method of teaching evaluation is limited or not necessarily honored.
3. Have a consistent form to guide chair- peer-evaluation/observations. Such a document is required by the Handbook for chairperson review (§44.5.1) and peer-review (§44.5.2), but few departments provided these documents.
4. Use the student completed course evaluations information appropriately. The Handbook (§43.2) requires that the "evaluation of teaching for personnel decisions must be fair and systematic." Factors such as mode of delivery, level, and subject influence student responses in course evaluations. Moreover, research on student completed course evaluations have shown that course evaluations are subject to bias. Boring et al. (2016) concluded that "student evaluations of teaching (mostly) do not measure teaching effectiveness." Heffernan (2021) summarized the research findings as "...highest scores are awarded in subjects...taught by white English first language speaking, able-bodied, male academics who are neither too young nor too old... , and who the students believe are heterosexual." In other words, certain faculty have an advantage in these scores based on demographic information. The process appears to be methodologically objective, but the raw scores are influenced by bias and prejudice (Heffernan 2021). Further, the bias is large enough for more effective faculty to receive lower student ratings of instruction (Boring et al., 2016). Accordingly, setting an arbitrary threshold on the expected average (or median) scores inadvertently may give an advantage to particular faculty, based on demographic information regardless of teaching effectiveness. In fact, Block (2021) which supports the continued use of student evaluations, advocated for "dismissing numerical averages." Given the known factors which influence rating in course evaluation (e.g., mode of delivery, demographics of the student, demographics of the instructor), which obviously differ between faculty members, the use of numerical averages in promotion and/or merit documents seems to be in conflict with the Inclusive Excellence Plan goal "of creating an atmosphere that will lead to the retention of diverse faculty."

The Faculty Handbook (§44.3.1) requires that statistical summary of the data should only be done when there has been at least ten responses. The department policies which have a threshold do not seem to clearly limit the calculation to courses with sufficient responses, which likely makes the policy in conflict with §44.3.1. Further, the difference between a class mean of 3.4 vs. 3.5 is extremely unlikely to represent a meaningful difference, so setting an arbitrary

minimum threshold will be difficult to defend if appealed. Instead, if a department does want to make distinctions among class means, the distribution of the student ratings of instruction should be taken into account as well as factors which may influence the student ratings of instruction (e.g., instructor, subject matter, and student characteristics). Considering the distribution of class means as well as mitigating factors will help identify substantial differences more clearly than an arbitrary threshold. The threshold statement could then be replaced by something like this: “Class means that are substantially below the department average should be examined for context and may be judged unsatisfactory.” If the department policy wants to more exactly define “substantially below,” a policy could be described where class means that are two standard deviations below the department mean are carefully reviewed, and those that are three standard deviations below may be deemed unsatisfactory.

After the review of the policies across the university, the Committee offers the following suggestions to support transparency in the evaluation of teaching (1, 5, 6) and to support faculty members professional development of pedagogical skills (2, 3, 4).

1. Create a specific **evaluation of teaching document** where the options and procedures are described. Having a teaching evaluation policy makes this important process more transparent. (Early Childhood, Youth and Family Studies teaching evaluation policy might be a good example). According to the Handbook (§43.2), the primary purpose of the evaluation of teaching is to help create “the best possible educational experience and in allowing faculty to develop to their full potential as teachers.” Most department policies of the evaluation of teaching are embedded in promotion and tenure or salary documents. Having a departmental policy on the evaluation of teaching would also make it easier for the department to confirm that their practice is in compliance with the Handbook. Although the evaluation of teaching is part of personnel decisions, those documents have a wider purpose and focus, which subsequently encourage a summative perspective on the evaluation of teaching.
2. Include an aspect of **formative assessment**, supporting the building of the instructor’s pedagogical skills. As stated in the Handbook (§43.1) the university recognizes that a goal of the evaluation of teaching is “allowing faculty to develop to their full potential as teachers.” This option is currently part of the department policy in multiple departments, (e.g., Anthropology, Landscape Architecture, Psychological Sciences, Theater and Dance).
3. Allowing faculty to contextualize or describe their teaching goals in a reflective statement (Block, 2021) as suggested in the Handbook (§43.2.3).
4. Permit instructors to exclude the course evaluation for at least one course from their teaching evaluation. This option recognizes that unanticipated hiccups sometimes emerge when implementing teaching innovations (e.g., immersive learning projects) or teaching strategies (e.g., active learning). The flexibility allows some risk-taking by faculty, rather than continuing with a safe teaching practice. This option is offered to faculty in some departments (e.g., Communication Studies, Educational Leadership).
5. Identifying the relative weight course evaluation will have in the evaluation of teaching. For instance, the policy from the Department of Special Education states that the observation by the Chairperson should be at least equal to the weight given to the student-completed course evaluation.
6. Drafting guidance for students about the “effective critique, bias and the purpose of evaluations” (Block, 2021)

4. University Summary

There were seven departments whose policies were not reviewed at this time, which require follow-up:

- College of Health
 - Military Science
- College of Sciences and Humanities
 - Physics and Astronomy
 - Political Science
 - Women's Studies
- Honor's College
- Miller College of Business
 - Accounting
 - Applied Business Studies

Based on the documents provided to the Committee, seven departments seem to be out of compliance with the Handbook (§43.2.2) requirement of having at least two methods of teaching evaluation:

- College of Architecture and Planning
 - Architecture
 - Urban Planning
 - Construction Management/Interior Design
- College of Health
 - Nursing
- Miller College of Business
 - Marketing
 - Information Systems and Operations Management
- Teachers College
 - Educational Studies

Based on the provided documents, few departments seem to be in full compliance with (§43.2.2), which gives the choice of the second method for the evaluation of teaching to individual faculty members. Accordingly the Committee suggests that EVERY department reviews their policy and procedures regarding teaching evaluation for compliance, specifically to §43.1, §43.2.2, §44.5, and §44.3.

Recommendation:

Ensure that department policy is in compliance with the Handbook, which is required at least once every three years (§43.1). The Handbook states that the evaluation of teaching must include:

- additional methods beyond student course evaluations (§43.1).
- faculty choice of the second method of evaluation (§43.2.2).
- have a form for guiding peer and chair reviews (§44.5).
- remove arbitrary thresholds on the numerical averages (Block, 2021) to ensure compliance with best methodological and statistical considerations (Handbook §44.3).

Suggestions:

- Have a Teaching Evaluation document which outlines the options and procedures for faculty.
- Include an aspect of formative assessment to support the instructor's pedagogical skills.
- Allow faculty to describe or contextualize their teaching.
- Identify the relative weight of course evaluations in the evaluation of teaching.
- Require fewer than all course evaluations in the evaluation of effective teaching.
- Draft guidance for students about effective critique, bias, and purpose of course evaluations.

5. Recommendations for Consideration by Faculty Council

The Faculty Council should revise the Handbook to:

- Outline the process when the Committee's review suggests that a department's policy might be out of compliance. Perhaps the Committee's report triggers
 - a Dean review within the College
 - a response from the Dean with a summary of the steps taken to get the departmental policies within compliance..
- Amend §44.3.1 to "Statistical summary of information from classes with fewer than ten (10) **responses** should be avoided."

References

Block, S. (2021). Yes, student reviews of classroom teaching have value. *Inside Higher Education*. https://www.insidehighered.com/views/2021/03/17/colleges-can-use-student-teaching-evaluations-improve-classroom-experience-opinion?utm_source=Inside+Higher+Ed&utm_campaign=fdfa395c5b-DNU_2021_COPY_02&utm_medium=email&utm_term=0_1fcbc04421-fdfa395c5b-197382789&mc_cid=fdfa395c5b&mc_eid=58db3950ed

Boring, A., Ottoboni, K., Stark, P. B. (2016). Student evaluations of teaching (mostly) do not measure teaching effectiveness. Retrieved from Science Open Research. doi:10.14293/S2199-1006.1.SOR-EDU.AETBZC.v1

Heffernan, T. (2021). Sexism, racism, prejudice, and bias: a literature review and synthesis of research surrounding student evaluation of courses and teaching. *Assessment and Evaluation in Higher Education*. Doi: 10.1080/02602938.2021.1888075