



Student Course Evaluation Improvement Project

PRESENTED BY Thomas Cohen, Chair

REVIEW DATES SEC – April 8, 2019 | SENATE – April 24, 2019

VOTING METHOD In a single vote

RELEVANT POLICY/DOCUMENT N/A

NECESSARY APPROVALS Senate, President

ISSUE

In January 2017, the Associate Provost of Learning Initiatives and Executive Director of the Teaching & Learning Transformation Center (TLTC) submitted a proposal to the Senate Executive Committee (SEC) recommending improvements to the CourseEvalUM system. The proposal noted that the University's method of evaluating courses had not changed significantly since an electronic system was first introduced in 2008 and called for a comprehensive review of the existing approach to evaluating courses and instructors. In February 2017, the SEC voted to charge the Academic Procedures & Standards (APAS) Committee with reviewing past Senate action on course evaluations, reviewing scholarship on course assessments and practices at Big 10 and peer institutions, consulting with a range of stakeholders across campus, and recommending whether changes should be made to the current system.

RECOMMENDATIONS

The APAS Committee makes a series of recommendations regarding changes to the University's CourseEvalUM system as shown in the attached report.

COMMITTEE WORK

In spring 2017, the APAS Committee met with the proposers to discuss their concerns with course evaluations and researched past Senate action on course evaluations. It also met with representatives of the Office of Institutional Research, Planning, and Assessment (IRPA), which oversees implementation of course evaluations, and the Course Evaluation Advisory Group, which advises IRPA on the development of the system and its uses on campus. A subcommittee with members from the APAS Committee and members from the Course Evaluation Advisory Group was formed to consider the charge in-depth and make recommendations to the APAS Committee. The subcommittee met approximately twenty times between July 2017 and January 2019, and consulted with a range of subject-matter experts and stakeholders.

The subcommittee developed recommendations based on its review and submitted its report to the full APAS Committee in February 2019. APAS shared its preliminary directions with the stakeholders the subcommittee previously consulted, as well as with the Senate at its March 2019

meeting. It also consulted with the Office of General Counsel. The committee developed a final set of recommendations based on the feedback it gathered. After due consideration, the APAS Committee voted to approve its recommendations on the course evaluation system at its meeting on March 29, 2019.

ALTERNATIVES

The University could continue using the current CourseEvalUM system.

RISKS

There are no known risks.

FINANCIAL IMPLICATIONS

There may be some cost associated with implementation of the recommendations. Specifically, there could be costs associated with incorporating survey results into tools used by students when registering for classes depending on how the recommendation is implemented.



Student Course Evaluation Improvement Project

2018-2019 Committee

Thomas Cohen (Chair)
Deanna Barath (Graduate Student)
Progyan Basu (Faculty)
Lauren Brown (Undergraduate Student)
William Cohen (Ex-Officio Provost's Rep)
Linda Coleman (Faculty)
Adrian Cornelius (Ex-Officio University Registrar)
Jeffrey Franke (Ex-Officio Graduate School Rep)
Lee Friedman (Faculty)
Shannon Gundy (Ex-Officio Rep for Director of Undergraduate Admissions)
Agisilaos Iliadis (Faculty)
Lisa Kiely (Ex-Officio Undergraduate Studies Rep)

Byung-Eun Kim (Faculty)
Roberto Korzeniewicz (Faculty)
Marilee Lindemann (Faculty)
Celina McDonald (Faculty)
Benjamin Parrish (Undergraduate Student)
Julian Savelski (Exempt Staff)
Michael Sparrow (Exempt Staff)
David Straney (Faculty)
Elizabeth Warner (Faculty)

Date of Submission

April 2019

BACKGROUND

In January 2017, the Associate Provost of Learning Initiatives and Executive Director of the Teaching & Learning Transformation Center (TLTC) submitted a proposal to the Senate Executive Committee (SEC) recommending improvements to the CourseEvalUM system. The proposal noted that the University's method of evaluating courses had not changed significantly since an electronic system was first introduced in 2008, and pointed out several areas where the current CourseEvalUM system could be improved. It called for a comprehensive review of the existing approach to evaluating courses and instructors, one that would be informed by recent scholarly literature and intended to revise the course evaluation items used. In February 2017, the SEC voted to charge the Academic Procedures & Standards (APAS) Committee with reviewing past Senate action on course evaluations, reviewing scholarship on course assessments and practices at Big 10 and peer institutions, consulting with a range of stakeholders across campus, and recommending whether changes should be made to the current system (Appendix 9).

CURRENT PRACTICE

In 2002, the Senate considered a proposal relating to teaching evaluations, which led to the formation of a Joint Provost/Senate Task Force on Course Evaluations and Teaching in spring 2003. The task force presented a preliminary report and recommendations in February 2004; the Senate subsequently passed a resolution recommending that "there be a University-wide requirement for student evaluations in all undergraduate and graduate courses." The task force submitted its final report in April 2005, identifying four purposes for a new course evaluation system:

- **Formative Evaluation:** To provide diagnostic feedback to faculty for the improvement of teaching.
- **Summative Evaluation:** To provide one measure of teaching effectiveness for use in the Appointment, Promotion, and Tenure (APT) and post-tenure review processes and in annual productivity reviews.

- Informative Evaluation: To provide information to students for their use in the selection of courses and instructors.
- Outcome Evaluation: For the purposes of documenting student learning.

The task force also recommended steps to enhance the institution's ability to assess and improve curriculum and instruction. The Senate voted to approve the Recommendations for the Implementation of Web-based Student Course Evaluations (Senate Document #02-03-39) in December 2005. The Provost subsequently appointed an implementation committee, which presented items for a new University-wide course evaluation system to the Senate as an informational item in April 2006. A complete overview of past Senate action on course evaluations can be found in Appendix 1. Today, practices and priorities associated with the CourseEvalUM system are informed by the Course Evaluation Advisory Group, a body composed of representatives from each College/School and various administrative units.

There are currently sixteen CourseEvalUM survey items about instruction that are asked of students in all courses, and four that are asked about teaching assistants (Appendix 2). Fifteen of the items about instruction are forced-choice items, and one is open-ended; three of the teaching assistant items are forced-choice, and one is open-ended. Items focus on either the course or the instructor. In cases where there are multiple instructors, the instructor-specific items are asked for each individual. In addition, Colleges/Schools and departments have the ability to add additional items to the evaluations; those that do so are included in Appendix 3. In some cases, the number of these additional items is significant and more than doubles the length of the instrument. Additional items are most often used to gather insights on courses and instructors and align with the original purposes behind course evaluations; in some instances, however, they are used to gather data for accreditation purposes, and may have little or nothing to do with the course being evaluated.

Results from different items on the evaluations are available to different groups. Responses to eight of the items (Administrator Items) are only visible to instructors and authorized campus administrators, and are intended for use in evaluating and improving instructor performance. The Administrator Items also include a single open-ended item. Given that responses to the Administrator Items can inform personnel decisions, they are kept confidential and only made available to the instructor and relevant administrators. In order to ensure that students benefit from the system, eight additional items are included (Student Items) that are only visible to students and instructors. This separation is known as the "firewall." Results from these items are primarily intended to help students select courses. There is some overlap in the themes addressed by the Administrator and Student Items.

Course evaluations are administered near the end of each term, and conclude before the start of the final exam period. For standard, fifteen-week courses, the system opens two weeks before the last day of classes. Results are not available until after final grades have been submitted. Reports to instructors and administrators include the score distribution, average, and standard deviation for each item. Additionally, comparative averages by College/School, department, and course level are reported. An "overall score" summarizes the average of all five Likert-scale Administrator Items.

Presently, results dating back to 2007 are available to currently registered students. In 2014, however, the University adopted a new vended platform to conduct evaluations. The ability for students to view results gathered after 2014 was not implemented until fall 2018, meaning students have only recently been able to access results from courses offered in the last several years. The current platform cannot show student grade distributions, which were previously available. Results

for courses with five or fewer students are not made available to students, and students can only view results for a particular course/section if the response rate exceeds 70%. Over the past three years, University response rates in fall and spring semesters have ranged from about 55% to 60%, thereby making results from many courses inaccessible to students. Additional information on how CourseEvalUM results are used may be found in the Committee Findings section.

COMMITTEE WORK

In spring 2017, the APAS Committee met with the proposers to discuss their concerns with course evaluations and researched past Senate action on course evaluations. It also met with representatives of the Office of Institutional Research, Planning, and Assessment (IRPA), which oversees implementation of course evaluations, and the Course Evaluation Advisory Group. A subcommittee with members from both the APAS Committee and the Course Evaluation Advisory Group was formed to consider the charge in-depth and make recommendations to the APAS Committee. Subcommittee members included:

Phil Evers (faculty, APAS Chair from 2016-2018)

Susan Hendricks (faculty, past APAS member)

Marilee Lindemann (faculty, APAS member)

Michael Passarella George (staff, Assistant Director for Decision Support in IRPA/Course Evaluation Advisory Group member)

Doug Roberts (faculty, past APAS member)

Joseph Sullivan (faculty, Course Evaluation Advisory Group member)

Kaci Thompson (staff, Course Evaluation Advisory Group member)

The subcommittee met approximately twenty times between July 2017 and January 2019, and consulted with a range of subject-matter experts and stakeholders. The subcommittee:

- Reviewed past Senate action establishing the purposes of course evaluations;
- Reviewed research on course evaluations at UMD conducted by IRPA, including research on what our current items measure, bias in course evaluations, and how results are used by students in the course selection process;
- Met with experts in the field, including Dr. Sandra Loughlin, an educational psychologist who directed the Office of Transformational Learning in the Robert H. Smith School of Business; and Dr. Alice Donlan, an educational psychologist and Director of Research for the TLTC;
- Reviewed a survey of current literature on student assessments of teaching;
- Met with the Associate Provost for Faculty Affairs and the Council of Associate Deans for Faculty Affairs (CADFA) to discuss the use of course evaluations in making teaching assignments and improving instructor effectiveness;
- Reviewed the TLTC's new Mid-Semester Evaluation of College Teaching (MSECT) pilot;
- Consulted with both undergraduate and graduate students at two meetings of the Senate Student Affairs Committee, as well as a focus group of students;

- Evaluated possible replacement items through cognitive interviews with students conducted by Loughlin and Donlan; and
- Reviewed practices at Big 10 and peer institutions.

The subcommittee developed recommendations based on its review and submitted its report to the full APAS Committee in February 2019. In reviewing the recommendations, APAS considered whether or not the practice of conducting University-wide assessments of teaching effectiveness should be discontinued altogether. Ultimately, the committee determined that collecting University-wide data on student experiences serves a useful purpose that should be continued and improved. Its recommendations are premised on a belief that the University should take this opportunity to reimagine the instrument in light of current best practices; narrow its purpose; and develop a tool that provides more and better information to administrators, instructors, and students.

APAS shared its preliminary directions with the stakeholders the subcommittee previously consulted, as well as with the Senate at its March 2019 meeting. It also consulted with the Office of General Counsel, and sent a survey to a select group of department chairs to gather feedback on its proposed recommendations for making teaching assignments and evaluating instructor performance. The committee considered the subcommittee's recommendations and the feedback it gathered in order to develop a final set of recommendations. After due consideration, the APAS Committee voted to approve its recommendations on the course evaluation system at its meeting on March 29, 2019.

PEER INSTITUTION AND SCHOLARLY RESEARCH

Big 10 and Peer Institutions

The committee reviewed information provided by the proposer on course evaluation practices at Big 10 and peer institutions (Appendix 4). Most peer institutions have a set of campus-wide questions that are asked, and many allow colleges, departments, and sometimes instructors to include additional questions. Several articles published over the course of the committee's work indicate that institutions have identified concerns and are conducting reviews similar to the ones called for in APAS' charge (see Flaherty, 2018; Supiano, 2018; Doerer, 2019). The committee found that while institutions have adopted a range of approaches, it is difficult to identify any single instrument or set of best practices the University might adopt. Well-designed instruments are developed for particular contexts and to answer specific questions. In consultation with campus experts, the committee determined that the University should first identify what it wants to assess regarding courses and instructors, and then design or adapt an instrument that targets those elements as precisely as possible.

Scholarly and Institutional Research

The committee reviewed recent literature relevant to student evaluations of teaching, including studies addressing bias in teaching evaluations. While documenting bias can be difficult, the literature suggests that broad or vague items, and items that allow open-ended comments in particular, are more susceptible to bias (Felton et al., 2008; Lindahl and Unger, 2010; Porter, 2011). Studies also indicate that student learning is not highly correlated with student perceptions of teaching (Uttl et al., 2017). The committee found that much of the literature on student evaluations is particular to the instrument being studied, and does not necessarily yield broadly applicable insights (Linse, 2017). The committee's work was also informed by a memo from Dr. Sandra Loughlin reviewing literature on student evaluations of teaching (Appendix 5). The memo noted that such evaluations often ask students about things for which they are not the best source of data.

The committee also reviewed several studies IRPA has conducted on the current course evaluation system that investigate usability, reliability and validity of the items; the relationship between response rates and instructor scores; and whether an instructor's race/ethnicity/citizenship and gender explain differences in ratings. IRPA also conducted phone interviews to identify practices associated with higher response rates, finding that students are more likely to complete evaluations when instructors emphasize that they value the feedback and are interested in improving a course. These studies are summarized in Appendix 6.

COMMITTEE FINDINGS

Administrator & Instructor Use of Course Evaluation Results

In the course of its review, the committee consulted various administrators, including the Associate Provost for Faculty Affairs and the Senior Vice President and Provost. It also attended several meetings of CADFA and the Undergraduate Academic Programs Committee (UGAP). Once it had identified preliminary recommendations, the committee distributed them, along with a survey, to a select group of chairs from both large and small departments, as well as the dean of a non-departmentalized College.

The committee learned that course evaluation results are used by administrators and instructors to assess performance in tenure and promotion cases. They are also used to identify potential problems in an instructor's teaching and indicate where additional intervention may be necessary. Practices vary, sometimes significantly, when it comes to the way results are used in appointment and promotion processes, though the committee identified the following generalities.

- **Tenured/tenure-track (T/TT) faculty:** The role that course evaluation results play in the APT process varies by College/School, though the University has been shifting to more holistic evaluations of faculty teaching that involve peer evaluations, student mentoring, curriculum/course development, a teaching portfolio, and other instructional activities. The Office of Faculty Affairs provides a template that can be used when compiling and interpreting the numeric results of evaluations; responses to open-ended items are typically included at the discretion of the individual faculty member. In general, however, the committee found that course evaluation results play a limited and decreasing role in the APT process, particularly given teaching is only one aspect on which a T/TT faculty member is assessed.
- **Professional track (PTK) faculty:** The role course evaluations play in the Appointment, Evaluation, and Promotion (AEP) process for PTK faculty is more significant than it is in APT. There is no consistent standard for peer evaluation process for PTK faculty, and course evaluations are, in some cases, the sole or most significant factor used in making renewal or promotion decisions, particularly for purely instructional faculty.

Most administrators are aware of the results' limitations and contextualize them with other sources of data on instructor performance. This caution is not universal, however, and evaluation results are sometimes used in inappropriate ways. These include averaging all of an instructor's results into a single number for comparisons with peers or to give teaching awards, and comparing instructor averages to the College/School- and department-averages for courses of a similar level.

Instructors and administrators often use the open-ended comments to contextualize and nuance the numeric data; as one administrator put it: "the numbers tell you there is a problem and the comments tell you what the problem is." Some instructors expressed significant concerns with the

comments, however, noting that they can be biased and even abusive. Some instructors also indicated that they ignore the comments entirely. Both IRPA and the Office of Faculty Affairs have received complaints from instructors about the open-ended comments. To help address these concerns, IRPA has added the following language before the open-ended items on the evaluations:

In order to help instructors and administrators best use your feedback to improve teaching and learning at the university, please be thoughtful and constructive when writing comments. Inappropriate or offensive comments do not reflect the civil engagement we value as an institution, and they are generally not effective in stimulating improvements to instruction. Should any comments take the form of actual threats, they will be forwarded, with the student's identifying information, to campus police for threat assessment.

While it has received requests from instructors to remove specific comments, IRPA is not equipped to evaluate the nature of student comments or make decisions regarding which comments should be removed. In addition, the current system does not offer a way to easily delete comments. The committee acknowledged that the use of open-ended comments can be abused, but determined that their value was significant enough that they should be retained. The committee recommended including two open-ended items that use specific prompts related to positive aspects and areas for improvement. The committee hopes that this will yield more actionable responses and may reduce the number of biased comments.

Student Use of Course Evaluation Results

The committee met twice with the Student Affairs Committee and hosted a small focus group of students from different disciplines. From these sessions, the subcommittee gained insights into how students use course evaluation results and other resources to select courses. Students reported using a range of resources including CourseEvalUM results, third-party websites (among them ratemyprofessor.com, ourumd.com, and planetterp.com), and word of mouth. Student responses to a 2016 Campus Assessment Working Group Snapshot indicated 43% of students considered CourseEvalUM a “major factor” when choosing courses (Appendix 7). Students expressed uncertainty as to whether the results were for instructors, administrators, or other students, and did not always understand which items referred to the course and which to the instructor. The distinction between Administrator and Student Items was also unfamiliar.

When asked what would make a course evaluation system more useful and improve completion rates, students asked for access to the open-ended comments, grade distributions, and a “star system” for providing a simple snapshot of student satisfaction with a course or instructor. Students reported that incentivizing participation by assigning extra credit and devoting class time to completing evaluations are both effective. Knowledge that their responses would make a difference in how a course was taught in the future is also a motivating factor, which is supported by IRPA’s phone interview project looking at response rates (Appendix 6).

The committee considered ways to increase the value of the system to students. In addition to recommending that students be given access to all of the numeric results, the committee discussed ways to increase response rates to ensure that the threshold of 70% needed to release results to students is more consistently met. Its recommendations include encouraging instructors to emphasize the value they place in student feedback and set aside class time to complete evaluations. Responding to anecdotal feedback from both students and instructors that an excessive number of items decreases response rates, the committee also considered the length of the instrument. It determined that the number of University-wide items should remain the same, and recommended that the number of College/School/department items be limited to five.

TLTC Initiatives on Evaluating Teaching Effectiveness

Many of the same issues the subcommittee was considering have been considered by the TLTC. The committee learned that the TLTC has been developing a [Mid-Semester Evaluation of College Teaching \(MSECT\)](#), which was piloted in 2017 and 2018 (Donlan, 2019). MSECT is a survey administered through Qualtrics and modeled on the [Fearless Teaching Framework](#) (Donlan et al., 2019). It grew out of repeated requests by instructors for a way to evaluate and improve their teaching during the semester. Per the purposes identified for the current course evaluation system, instructors were interested in *formative feedback* that could be used immediately, rather than the next time a course was taught. The tool is intended primarily to help instructors improve their teaching; while instructors could incorporate results into a teaching portfolio, results would not be visible to administrators as with current CourseEvalUM results. Preliminary analysis from pilot data provides evidence that the measure is a valid and reliable assessment of teaching effectiveness. The committee was impressed by the initial results of the pilot, and its recommendations encourage further development and adoption of MSECT and other mechanisms to gather mid-semester feedback on teaching.

Firewall Between Student and Administrator Items

When the University developed its current course evaluation system, the results were treated as “evaluations” of instructors, both rhetorically and in decisions over who should have access to the results. As described above, this approach resulted in two sets of items visible to different audiences with a firewall between them. In the course of its work, and after consulting extensively with the Office of General Counsel, the committee determined that this division is no longer necessary, as results are not “evaluations” used to assess instructor performance.

Student responses are opinions regarding their experience in a particular course. Students are not subject matter experts who can speak to the organization of a course's content, nor are they trained in pedagogy and able to accurately assess an instructor's teaching. Further, there are no standards, rubrics, or specific criteria for students to apply when completing evaluations. Given this, students are not equipped to directly “evaluate” an instructor's performance. Their perspectives can be used by administrators to gain insights into instructors’ teaching effectiveness, though in such cases it is the administrator conducting the evaluation, the results/conclusions of which are the actual evaluation to be considered part of the personnel record.

Additionally, all information used in personnel evaluations need not be confidential. There are multiple other sources of information that are not confidential, including numbers and names of publications, syllabi, the number and value of grants, etc. The committee decided that allowing students, instructors, and administrators identical access to numeric results would not invalidate their use in certain personnel processes, and would increase the amount of information available to both administrators and students. It determined, however, that responses to open-ended items should remain confidential and visible only to instructors and administrators as they could contain personally identifiable information and are unaggregated, unit-level data.

Given that results are not performance evaluations, the committee determined that the “course evaluation system” should be renamed to better communicate that it gathers students’ perceptions and experiences about a course or instructor, a distinction that current terminology may blur.

Limitations of the Current Course Evaluation System

Based on the reviews of relevant literature and consultation with campus experts addressed above, the committee identified significant concerns with the CourseEvalUM items and their ostensible purpose. The items invite students to speak to themes that they are not in a position to credibly

address. For example, one item asks students to indicate whether “the standards the instructor set for the course were Too Low/Appropriate/Too High,” which is better assessed by other experts in the discipline. In addition, many items are not sufficiently specific to yield information that is actionable by an instructor. This lack of specificity is also concerning as vague or imprecise items are more open to bias. The items are also unable to adequately address all four of the system’s original purposes:

- **Formative Evaluation:** While results can help instructors improve their teaching, they cannot be used to inform mid-semester adjustments. As addressed above, instructors now have access to more dynamic and timely mechanisms to gather feedback throughout the semester. Further, many of the current items address vague or subjective criteria and do not ask about specific classroom practices associated with effective teaching, making it difficult for instructors to directly address possible deficiencies.
- **Summative Evaluation:** The results remain a potentially valuable measure of teaching effectiveness for use in the APT, AEP, and post-tenure review processes. Yet results are open to misuse, given they yield potentially biased information and tend to measure a single factor (general satisfaction with an instructor). In light of these shortcomings, their use by some as the primary or even sole measure of teaching effectiveness is particularly problematic.
- **Informative Evaluation:** The results remain a valuable resource for students making course selections, though increased completion rates increase information available to students.
- **Outcome Evaluation:** The current system is ill-suited to measuring student learning, which is better addressed by learning outcomes assessments and other mechanisms. Studies have also shown that student learning is not highly correlated with students’ perceptions of learning, which are often informed more by other factors (how much the student enjoys the topic, whether the course was required, etc.).

The committee debated at length the purpose of the instrument, eventually determining that it should be redesigned to focus primarily on summative and informative feedback. Surveys should be summative to the instructor and to administrators and serve as one measure of teaching effectiveness to use in evaluating and improving teaching practices. Surveys should also be informative to students, in order to assist them in selecting courses and instructors. The committee determined that the current items are not able to adequately or efficiently meet these goals, and decided to recommend that they be replaced.

FRAMEWORK FOR REPLACEMENT SURVEY ITEMS

The committee decided that the number of items should remain the same as in the current instrument. It considered new items that fall into three conceptual categories:

1. Those designed to provide summative feedback for use in evaluating and improving teaching;
2. Those designed to inform student course decisions; and
3. Those intended to assess teaching assistants.

The committee determined that items in the first category should either assess baseline teaching practices that should be met or identify the utilization of best practices of teaching effectiveness. In discussing the relative balance between these two purposes, the committee considered focusing survey items solely on baseline or core teaching practices that should generally be expected of

every instructor, where consistently low scores can serve as a red flag and inspire discussions between instructors and administrators. It decided that such an approach would only exacerbate concerns that course evaluations are designed to emphasize deficiencies in teaching. It might also suggest that the University's standard for instruction is mere adequacy. Therefore, the committee determined that items addressing baseline practices should be accompanied in approximately equal numbers by items addressing best practices, which presents an opportunity to both identify practices that the University values and encourage adoption of those practices.

In light of the scholarly research discussed above, the committee determined that the work of developing and testing sound replacement items is sufficiently complex that it should be entrusted to those well versed in the scholarship. The committee decided to focus its efforts on identifying constructs that address specific teaching practices and recommend that subject-matter experts be tasked with developing the specific wording associated with each item based on those constructs. Examples of possible wording of survey items associated with most of the constructs can be found in Appendix 8. With the exception of the items intended to inform student course selection, the constructs the committee included in its recommendations are all supported by a large body of literature supporting their connection to learning. The student course selection constructs are based on requests made by students. Following their development, specific survey items would be tested and piloted by the Course Evaluation Advisory Group, IRPA, and subject-matter experts. The final survey items would then be shared with the Senate for its feedback before being implemented. IRPA would also ensure that items are presented in a logical order on the survey, rather than presenting them in the conceptual categories that informed the development of the constructs.

Stakeholder Feedback

The committee circulated its draft recommendations with a range of stakeholders including the Course Evaluation Advisory Group, the Office of Faculty Affairs, administrators, and students. A few stakeholders noted that replacing the current items would disrupt the ability to measure teaching improvement over time. Faculty going up for tenure in the next several years, for example, would have to modify their promotion materials to account for the sudden shift. After consulting with the Office of Faculty Affairs, the committee determined that the APT process could accommodate the change. The AEP process relies more on CourseEvalUM results in some cases, which led the committee to recommend that the University explore ways to provide more holistic reviews of instructional faculty. Adding past results from CourseEvalUM and data from the new survey items to the data warehouse would also facilitate more sophisticated analyses that could smooth the transition, which led to another of the committee's recommendations.

Some stakeholders raised concerns about replacing the current items and about the availability of grade distributions. The committee considered potentially retaining some of the current items for several years to bridge the transition. Stakeholder feedback identified two items in particular as valuable: "I learned a lot from this course" and "Overall, this instructor was an effective teacher." The committee determined not to retain these two items. As noted above, student learning is not highly correlated with students' perceptions of learning, and "effective teacher" is an ambiguous concept subject to significant interpretation. However, several other current items closely align with the proposed constructs, which will facilitate some comparisons between new survey data and CourseEvalUM results. Ultimately, the committee decided that a clean break with the majority of the current items was in the best interests of instructors and administrators. The committee also considered the importance of making grade distributions available to students. Students consistently request them and IRPA annually receives and complies with FOIA requests for grade distributions from third-party websites. Given that students are able to access the information regardless, the committee decided to recommend that results once again be provided directly to students.

RECOMMENDATIONS

I. Guiding Principles & Parameters

1. The University should rename the current “CourseEvalUM” system to better communicate that it gathers students’ perceptions and experiences about a course or instructor and does not serve as an evaluation.
2. The University should replace the existing CourseEvalUM items with new survey items that follow these principles:
 - a. The number of University-level items should be approximately the same as in the current survey instrument.
 - b. The survey items should focus on measuring progress relative to baseline teaching practices and on utilization of best practices of teaching effectiveness.
 - c. The survey should include only those items that students can reliably answer and should focus on those items where students are the best or only source of data.
 - d. The majority of survey items should be designed so that responses can inform actionable changes by the instructor.
 - e. The survey items should be written using language that makes clear what is being asked of students.
 - f. Students should understand who will be reviewing their responses, in order to inform their thinking as they are filling out the survey.
 - g. The survey items should focus on asking students to speak to their own student experience, rather than asking for general feedback or input based on other students’ experiences, unless there is a compelling rationale to do otherwise.
 - h. The survey items should be relevant for in-person, blended, and online courses.
 - i. The survey items should clearly indicate whether they relate to the instructor or the course.
 - j. The survey items should be positively worded so that a high score on an item is positive and a low score indicates that adjustments in practices may be needed.
3. The survey item development process should involve a pilot or other mechanism for testing and refining the new items.
4. The Course Evaluation Advisory Group and Office of Institutional Research, Planning, & Assessment (IRPA) should provide an informational report to the Senate on new survey items to gather feedback before implementation.
5. The University should consider ways to ensure that survey results are not utilized as the sole basis for giving teaching awards or for assessing progress towards accreditation standards.

II. Recommendations on Constructs & Items

1. The Course Evaluation Advisory Group and IRPA should work with subject-matter experts to develop new survey items that align with the below constructs that assess teaching effectiveness, inform student registration decisions, provide opportunities for open feedback, and allow for feedback on teaching assistants.

Constructs that Assess Baseline and Best Practices in Teaching Effectiveness

- Timely feedback
- Clear assignment expectations
- Clear grading expectations
- Focus on course content in class sessions
- Value of required texts
- Climate
- Instructor support
- Quality feedback
- Scaffolding
- Cognitive engagement and/or rigor
- Alignment of instruction to assessment

Constructs that Inform Student Registration Decisions

- Course satisfaction
- Instructor satisfaction
- Time invested
- Major/non-major

Constructs for Open-Ended Feedback

- Positive Aspects
- Areas for Improvement

Constructs Related to Teaching Assistants

- Climate
- Timely feedback
- Effective use of class time
- Open-ended item on positive aspects
- Open-ended item on areas for improvement

2. The Course Evaluation Advisory Group, in consultation with the Teaching and Learning Transformation Center (TLTC), should develop a bank of additional items—based on baseline and best practices of teaching effectiveness and literature in the field—that Colleges/Schools and units may include in addition to the University-level items.
3. The University should limit Colleges/Schools and units to a maximum of five additional survey items, which should be developed in consultation with the TLTC.
4. The Course Evaluation Advisory Group should carefully consider the order in which items are presented to students on the survey and whether they should correspond to the order of responses provided in reports available to administrators and instructors.
5. The Course Evaluation Advisory Group should ensure that survey items are clearly identified as applying to either the instructor or to the course.

III. Recommendations on Implementation and Usage of Survey Results

1. The University should encourage instructors to gather mid-semester feedback on their teaching, using tools such as Qualtrics and resources provided by the TLTC.
2. The University should encourage instructors to set aside time in class for students to complete surveys and to explain to students the value and impact of survey responses on teaching practices.

3. The University should make numeric data from survey results available to instructors, administrators, and students. Responses to open-ended items should remain accessible to instructors and administrators only, not students.
4. The University should consider ways to incorporate survey results in information available to students during the course selection process.
5. The Course Evaluation Advisory Group should prioritize efforts to add existing CourseEvalUM data and future survey results to the data warehouse.
6. IRPA should discontinue the practice of including department-wide and College-wide averages across all courses of a given level in survey results.
7. The University should again make course grade distributions available to students.
8. The University should not release survey results from courses with fewer than 5 students and should continue the practice of not releasing results to students if the response rate for a given course is less than 70%.
9. The University should consider how best to ensure that survey results are not utilized as the sole basis for personnel determinations of PTK faculty.
10. The Provost's Office should develop guidance on best practices for utilizing statistical analysis of data from survey results in the Appointment, Promotion, and Tenure (APT) and Appointment, Evaluation, and Promotion (AEP) processes.

APPENDICES

- Appendix 1 — Past Senate Action on Course Evaluations
- Appendix 2 — Current CourseEvalUM Items
- Appendix 3 — Number of Evaluation Items by Unit
- Appendix 4 — Big 10 and Peer Institution Research on Course Evaluations
- Appendix 5 — Memo from Dr. Sandra Loughlin (November 1, 2017)
- Appendix 6 — IRPA Studies on UMD's Course Evaluation System
- Appendix 7 — CAWG Snapshot of Student Experiences
- Appendix 8 — Sample Item Wording for New Constructs
- Appendix 9 — Charge from the Senate Executive Committee

REFERENCES

- Doerer, Kristen. 2019. "Colleges Are Getting Smarter About Student Evaluations. Here's How." *The Chronicle of Higher Education*. Retrieved from (<https://www.chronicle.com/article/Colleges-Are-Getting-Smarter/245457>).
- Donlan, A. E., & Byrne, V. L. (2019). "Gathering formative feedback on college teaching: Evidence-based mid-semester evaluation in higher education classrooms." Poster presented at the biennial meeting of the Society for Research on Child Development, Baltimore, MD.
- Donlan, A. E., Loughlin, S. M., & Byrne, V. L. (2019). "The Fearless Teaching Framework: A model to synthesize foundational education research for university instructors." *To Improve the Academy*, 38(1), 33-49.
- Felton, J., Koper, P. T., Mitchell, J., & Stinson, M. 2008. "Attractiveness, easiness and other issues: Student evaluations of professors on Ratemyprofessors.com." *Assessment & Evaluation in Higher Education*, 33(1), 45-61.
- Flaherty, Colleen. 2018. "Teaching Eval Shake-Up." Inside Higher Ed. Retrieved from (<https://www.insidehighered.com/news/2018/05/22/most-institutions-say-they-value-teaching-how-they-assess-it-tells-different-story>).
- Lindahl., M. W., Unger, M. L. (2010). "Cruelty in Student Teaching Evaluations." *College Teaching*, 58, 71-76.
- Linse, A. R. (2017). "Interpreting and using student ratings data: Guidance for faculty serving as administrators and on evaluation committees." *Studies in Educational Evaluation*, 54, 94-106.
- Supiano, Beckie. 2018. "A University Overhauled Its Course Evaluation to Get Better Feedback. Here's What Changed." *The Chronicle of Higher Education*. Retrieved from (<https://www.chronicle.com/article/A-University-Overhauled-Its/243803>).
- Uttl., B., White, C. A., & Gonzalez, D. W. (2017). "Meta-analysis of faculty's teaching effectiveness: Student evaluation of teaching ratings and student learning are not related." *Studies in Educational Evaluation*, 54, 22-42.