

White Paper on Evaluating Teaching Effectiveness

By

Academic and Professional Status Committee of VCU Faculty Senate

Kristine Artello, Chair

Masoud Manjili, Co-Chair

Committee Members:

Kyungeh An, Mary Boyes, Dave Dixon, Deborah DiazGranados, Kelly Lockeman,
Jeff Lodge, & Steven Price

With Special Thanks to Enoch Hale, Director,
Center for Teaching & Learning Excellence at VCU

Abstract

VCU faculty are largely unclear about how, when, and according to what criteria their teaching is evaluated annually. There is no shared definition of that what constitutes substantive due process for teaching evaluation (e.g. the shared expectations for teaching, shared standards of what constitutes good teaching, and standards for measuring teaching effectiveness based on best practices). There is considerable variation in student evaluations of teaching. These evaluations demonstrate consistent bias, particularly against women and underrepresented instructors. This committee studied the current student course evaluation system used in schools and departments across the University. We gathered all of the course evaluations forms used at VCU to ascertain the questions asked across units. Our results indicate great variability across units –both in terms of length and types of questions asked. Additionally, we found that some units have not updated evaluation questions in this century. Some units include additional items for evaluation of faculty teaching, such as peer reviews, scholarship on pedagogy, curriculum development, and student mentorship in faculty evaluation. Yet, it is unclear how often and to what degree these items are used in annual evaluations and whether these evaluations are conducted using best practices. We concluded that greater clarity, consistency, and transparency in the process to evaluate teaching is necessary to further equity among faculty—especially, for annual review, promotion, and tenure purposes. We recommend that best practices be used to create a context for teaching and to create greater clarity around expectations for faculty. Part of the process to decide best practices should include the development of a shared understanding of standards and expectations of faculty around the type of teaching to be expected and encouraged.

Table of Contents

I. Problem statement	4
II. Background	4
III. Methods	5
IV. Results	6
V. Recommendations	8
 Charts and Figures:	
Chart 1 Total Number of Questions Asked on Student Course Evaluations	6
Table 1 Questions Asked Most Often in Student Course Evaluations	7

White Paper on Evaluating Teaching Effectiveness

Problem Statement

As a research university, VCU has a commitment to scientific research and to grounding our actions in evidence. The research on evaluations of faculty teaching has focused on the use of students' course evaluations. The scientific evidence continues to strengthen that students' course evaluations are influenced significantly by the implicit bias based on gender and minority status (Mitchel & Martin, 2018) and that the surveys have zero correlation to teaching effectiveness and student learning outcomes (Uttl, White & Gonzalez, 2017). As the evidence continues to build, it becomes a potential concern that faculty evaluations based on biased information will lead to discrimination against women and minorities potentially opening the University to liability for violations under Title IX (See Mitchel & Martin, 2018). The situation becomes more complicated as promotion and tenure decisions are frequently based on student course evaluations but also because annual assessments have been the primary mechanism to award merit raises, which have been the only raises awarded in the last five years. Having a fair, objective and transparent faculty assessment process is important to improve teaching and to create a place where teaching and research are valued equally. In an effort to help the University move forward on the Quest 2025 to transform teaching and learning at VCU, shared standards and expectations need to be clarified for faculty in relation to teaching, and experimentation will need to be encouraged.

In this white paper, we will provide an overview of the current system used for evaluating teaching effectiveness through student course evaluations. As schools and colleges explore ways to improve the annual review process and increase transparency and fairness in the process, this committee compiled information about VCU's current system, available resources, and methods to create a fairer annual evaluation system. As raises continue to be determined on a merit basis, it is crucial for a fair and transparent system to be developed in order to strengthen the University community and to improve teaching effectiveness for VCU students. We will examine the student course evaluations themselves and compare them across the University. Then we will present our concerns and provide suggestions and resources for schools, colleges, departments, and programs to develop a more holistic and fair evaluation system for teaching effectiveness.

Background

Four years ago, this committee started an examination of the faculty assessment process. We researched the evidence surrounding student course evaluations. The research clearly shows bias based on gender and minority status, which is very concerning to the Faculty Senate. We also discovered that there was no faculty appeal policy in place University-wide. Some schools had an appeal process, but the process ended at the Dean's office, which may choose not to hear the appeal. Consequently, many faculty evaluation issues were brought through the Faculty Mediation and Grievance Policy, using significant University resources. Last year, a [policy](#) was drafted and enacted regarding annual assessment of faculty. This policy included the recognition that because of issues surrounding such surveys, student course evaluations cannot be the sole

determinant of teaching effectiveness. Further, the policy created a clear line of appeal to the Provost office. The procedure has become clarified in process but not necessarily in substance. Therefore, the evaluation of faculty regarding teaching effectiveness continues to be a concern, which has led this committee to examine how faculty are being assessed University-wide and to research the best practices for evaluating faculty.

Method

This committee collected student course evaluations from over 25 schools, colleges, departments, and programs to ascertain the basis being used to evaluate faculty across the University. Based on our evaluation, the majority of the schools and colleges have not changed their course evaluations in this century. For many units, evaluations have not changed since their inception beyond moving online in 2008/9. At this point, nearly 60% of the course evaluations are conducted online through a new system called Blue. According to the VCU Course Evaluation website, faculty are unable to add or subtract questions relevant to their courses. Evaluation questions are set by the school, college or the department. Yet faculty are responsible for the response rate for the majority of the courses. In fact, one of the ways to encourage students to complete the survey is for faculty to award them extra credit as stated below:

You can provide students with incentives for completing the course evaluations. Since the evaluations are considered confidential, we do not reveal the names of students who have completed the evaluations. Instead you can give incentives based on the overall response rate of the class. Be creative, some faculty use extra points based on achieving a certain level, others come up with amusing awards. Some faculty have more than one class compete and the winner gets a pizza. That may be extreme but it does work.

Again, use an incentive that you are comfortable with. ([VCU Course Evaluation Website](#))
The concern of low response rate is a skewed sample of strongly negative or positive responses only. Some faculty have expressed concern about the ethics of this type approach (e.g. bribe).

Some colleges, schools, and departments have changed their student course evaluations. The University College has altered its evaluations in Focused Inquiry to concentrate on practiced and attained skills learned by students. These new evaluations are used in particular for UNIV 111 and 112. The faculty appeared to have received good feedback from the students and find that the evaluations focus more on student learning. The School of Pharmacy revised its student course evaluations down to nine questions for the majority of their courses. For clinical pharmacy courses, three more questions are asked. The nine questions focus on clarity, assessment, and placement of the course in curriculum. The School of Medicine and the School of Dentistry have adapted their course evaluations to align with program goals and the curriculum. The School of Medicine has initiated a program to track student participation in these surveys and has made participation in evaluations part of the students' required responsibilities. The School of Dentistry has changed its course evaluations; however, it is unclear whether they were deployed properly, as Dentistry does not use the Blue System.

Some units appear to be developing new initiatives. For example, some departments in the College of Humanities and Sciences have started peer observation of teaching and have

developed protocols surrounding such endeavors. The Department of English has also changed its evaluations by reducing the number of questions asked. Additionally, in late 2016, the Center for Teaching and Learning Excellence (CTLE) was opened. This year, the CTLE has hired new colleagues and continues to develop resources for faculty to improve teaching and scholarship on pedagogy. The CTLE provides a midterm evaluation that involves about 30 minutes of class time when students are placed in groups and discussion as directed by a CTLE representative. Following the evaluation, CTLE staff and the faculty member meet to discuss the results. These evaluations are limited, however, due to lack of resources. (Faculty who are interested should contact the CTLE about scheduling.)

Results

Using online evaluations were available through the Blue system, we tracked frequency of questions in Chart 1 and for type of questions in Table 1. In Chart 1, student course evaluations ranged from the most questions asked, at 32 questions for 6 schools, and the least questions asked, at 9 for the School of Pharmacy. Student course evaluations averaged 20.6 questions. This count does not include questions prompting for comments. Nearly all course evaluations had space at the bottom after all questions for additional comments.

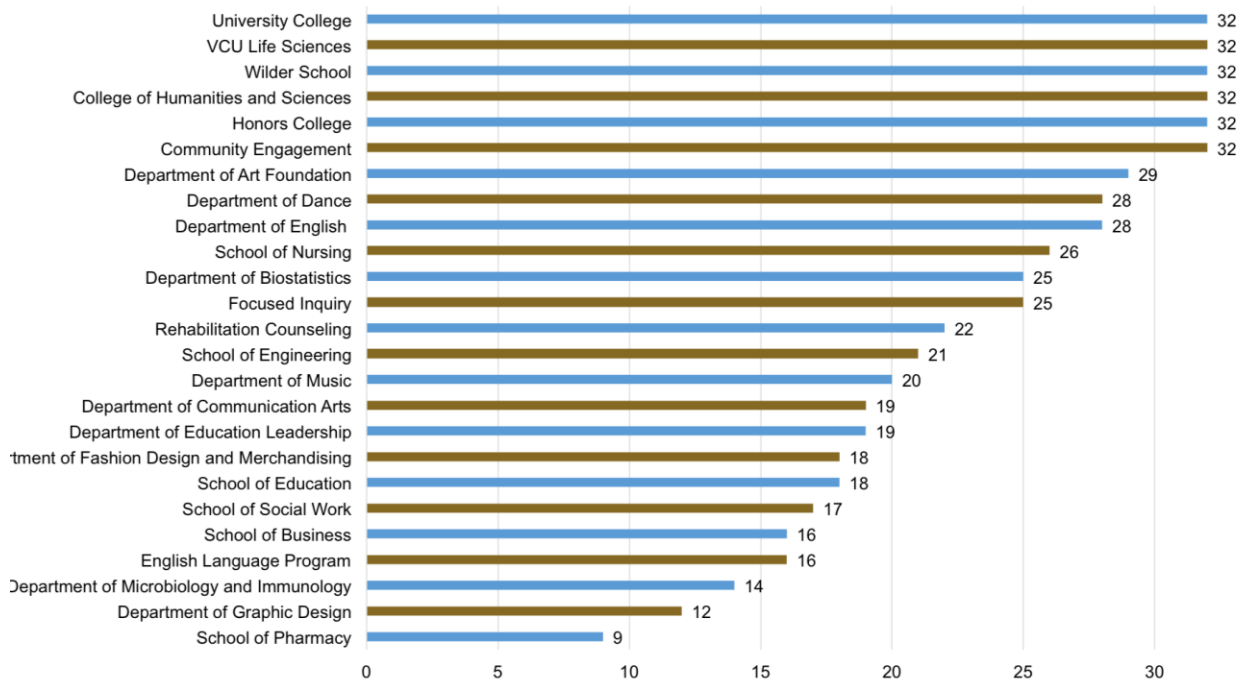


Chart 1 Total Number of Questions Asked on Student Course Evaluations

The evaluations with 32 questions have been in existence since the 1980s according to faculty who have been at VCU since that time. Beyond the change from paper to electronic evaluations, all other aspects of these 32-question evaluations have remained constant.

We also examined the text of the evaluation questions. Thirteen questions are present on the majority of evaluations. The questions do not provide a context for the course, and in some cases are ambiguous. As seen in Table 1, the most common questions do not relate to student learning outcomes, skill attainment and/or practice in the classroom. Studies have shown that a student's grade in a course has correlated highly with rating for instructor (Boring, Ottoboni & Stark, 2017; Brockx, Spooen, & Mortelmans, 2011; Fenn, 2015; Svanum & Aigner, 2011). Answers to these questions may be easily influenced by bias based on gender and minority status (Boring, Ottoboni & Stark, 2017; Fenn, 2015; MacNell, Driscoll & Hunt, 2015; Potvin & Hazari, 2016; Rosen, 2018; Wagner, Rieger, & Voorvelt, 2016).

Table 1 Questions Asked Most Often in Student Course Evaluations

Course Questions	Rate
What grade do you expect to earn in this course?	65%
The course was well organized.	65%
The instructor was readily available for consultation with students during office hours or by appointment.	62%
How would you rate this professor overall?	62%
Student responsibilities (being prepared for class, classroom participation, group projects, etc.) were well defined in this course.	62%
The instructor treated students with respect.	58%
What was your overall grade point average (GPA) at the beginning of the semester?	58%
What best describes why you are taking this course?	58%
How would you rate this course overall?	58%
The instructor was well prepared for each class	54%
The instructor encouraged students to feel free to ask questions.	54%
What is your class level? (If you are a special student, please choose the category that best describes you.)	54%
How would you rate the demands which this course made upon you?	54%

Current students' course evaluation system

Based on our research and discussions, the current student course evaluations do not provide sufficient context for summative assessment of teaching effectiveness. Yet, the numbers obtained from student evaluations often used by supervisors for measuring teaching effectiveness without a context of the course or a holistic understanding of faculty teaching. The majority of the student course evaluations have too many questions. Additionally, these questions need to be should not contain ambiguous and vague interrogatories, such as “How would you rate this professor overall?” This question, particularly, is open for bias since it lacks the focus on specific behavior. Such a far-reaching and vague question has been found in research to have zero correlation to teacher effectiveness and student learning (Uttl et al., 2017). At this point, there does not appear to be any evidence to support the use of student course evaluations for summative assessment of faculty.

Recommendations

The recommendations from our work focus on two different areas: student course evaluations and best practices in holistic assessment of faculty teaching. The student course evaluation recommendations concentrate on addressing the inadequacies of the current system and responding to concerns expressed by faculty across VCU. Recommended holistic assessment options arose from research, work with Enoch Hale, Director of the Center for Teaching and Learning Excellence, and committee discussions surrounding ways to improve the current system.

Committee's Recommendation for Student Course Evaluations

Although student course evaluations are not useful in summative assessment of faculty, they may provide insight for instructors as formative assessment of teaching if the questions are customized and responsive to courses and programs about methods and learning. Formative assessment guides instructors to evaluate teaching methods and student learning in order to make decisions on future pedagogy. To that endeavor, we suggest that evaluations contain fewer questions with more focus on skill attainment. Additionally, these evaluations should allow for comments and for students to provide examples. For instance, if a student gives a low rating on “respect toward students,” the survey should ask for an example in comment. Currently, nearly all comments in the evaluations are only at bottom of the forms. To provide a more complete picture, it is important to include questions regarding students' actions including—whether students purchased course materials, turned in assignments, etc. It is important to gauge how much the student engaged with the course. Students who are truly engaged can provide the best feedback to improve the course for professors.

To summarize, we recommend the following actions regarding student course evaluations:

1. All surveys should contain 15 or fewer questions.

2. Units that have not changed or adapted their student course evaluations in the last eight years should examine their surveys to determine their current relevance and effectiveness for teaching and learning.
3. Units should eliminate questions that are vague—e.g. What is your overall rating of the instructor? What is your overall rating for this course? Research has shown zero correlation between these questions and teaching effectiveness
4. If a student provides a low rating on a key question, the survey should require students to provide a comment that gives an example to explain the rating. Comments should be allowed throughout the survey to provide students an opportunity to give context to their quantifications.
5. Units should remove questions that do not relate to teaching and learning within the control of the student and the professor; e.g. How would you rate the physical environment of the class? Facilities questions may be better served by a separate survey.
6. Evaluations should be restructured to present all teaching questions in one section and all student questions in another.
7. Evaluations should allow some customization so instructors can ask specific questions to improve pedagogy.

Suggestions to develop a more holistic approach to evaluating teaching effectiveness

In this report, we have provided several suggestions to create a more holistic evaluation of teaching effectiveness. Some items listed are methods to promote the development and innovation in teaching methods. Others are ways to gather and present data to show teaching activities and effectiveness. Through these types of contextual assessments, the University can provide a path for faculty to show adaptation, innovation, and improvement in teaching methods from multiple perspectives. These contextual assessments can better inform department chairs about the quality and intentions informing faculty teaching efforts.

Teaching effectiveness is a fluid endeavor as each semester brings new students with different strengths and weaknesses. A key component for evaluation should be evidence that shows efforts to improve and innovate teaching through self-reflection, pedagogical training, and experimentation. Sometimes experimentation and innovation works and sometimes it does not. By providing faculty the means to show innovation, the University can encourage the development of teaching styles consistent with Quest for Distinction Theme 1. In the Quest 2025, innovation and transformation to lead to student success will be supported well by re-thinking how we develop, assess, and evaluate teaching.

- ★ *Focus group evaluation:* For any class over a duration of time that sufficiently accounts for changes in student populations (CTLE completes these evaluation). This type of evaluation, usually conducted at the midpoint of the semester, gathers anonymous feedback from students about what is helping them learn and what is not. Focus group evaluation provides not only for student feedback on the impact of the teaching, but these evaluations also show how the professor responds to student feedback. This is considered

best practice throughout centers for teaching and learning nationally. [Vanderbilt](#) provides one example of a similar program.

- ★ *Peer observations:* When a faculty member's teaching evaluation is conducted by colleagues who can best relate to the instructional context, a culture of equity and continual growth become the norm rather than the exception. Example: [Teaching Triangles](#). Teaching triangles: 1) identify areas for enhancement and effectiveness in teaching techniques, 2) utilize meaningful feedback methods, and offer ways to improve teaching while providing a structured means of reflection with a colleague. This type of peer review is more formative in nature and will be helpful with innovating and improving teaching at VCU.
- ★ *Summative peer reviews:* We recommend that peer reviewers be trained in appropriate practice and expectations for best practice. Peer reviews should not be a surprise or ambush-type of evaluation but should align with best practices to a three step approach (Golparian, Chan, & Cassidy, 2015). A good peer review provides a true context for teaching. At a minimum, three meetings should occur including a pre-observation meeting during which teaching philosophy and approach are discussed and materials for the course are shared, a class observation (a minimum of one class observation, but two provides more balanced data), and a post-meeting debriefing. A report should then be written and provided to the observed professor. The CTLE can assist in the training of peer reviewers. The [Center for Teaching at Vanderbilt](#) sets out the best practice and has excellent resources concerning peer reviews. Another resource is [Cornell Center for Teaching Innovation](#). Both resources point out that peer reviews are more than one or two visits to the classroom and a report generated for promotion and tenure. Best practices included trained observers and a three-step process: pre-observation meeting, class observation, and post-observation debriefing (Golparian, Chan, & Cassidy, 2015).
- ★ *Self-assessment:* Evaluations should use specific criteria from the faculty member's field (including utilization of student voice in some form) to show development of teaching to meet students' needs. The goal for this portion of the evaluation is to show a professor's development as a teacher. Vanderbilt has extensive resources about [self-assessment of teaching and developing teaching portfolio](#) options.
- ★ *Long-term student outcomes:* These outcomes should be tracked in order to assess effectiveness in teaching and course design. For example: follow students in from gateway courses to 300/400 level courses and see how the students perform. Another example: follow alumni and send out surveys to see how well prepared they were after being in the field for two to five years.
- ★ *Increasing teaching pedagogy and training:* The University should create resources to help professors innovate and adapt their teaching to meet new students' needs. Example: a program decides to revamp its approach to a program goal and the CTLE works with them to develop the new modality through the curriculum. The creation of Faculty Learning Communities (FLC) around specific topics, such as high impact practices, large

lecture classes, discussion and seminar courses, etc., would provide faculty with more resources in their teaching methods.

- ★ *Digital teaching portfolio* (presentation by CTLE): The portfolio provides a longitudinal view of the course evaluations for one course plotted by years (x axis) and ratings (y axis). Example: as one hovers over the rating, the context is displayed—change in class size, changing a book and/or assignments, changing modality, writing intensive characteristics, other works completed that year (one published a book or many articles), etc. Such a system would allow for innovation to be rewarded even when it does not work well.
- ★ *Inclusion of teaching dossiers*: Teaching dossiers, while somewhat like the portfolio above but not quite so quantitative data-based, contain student-outcomes and products to provide a more qualitative analysis to the professor's work. Example: out of 120 students last year, 71 submitted to the National Conference for Undergraduate Research and 60 were accepted. The faculty would record that in the portfolio because that is a very high acceptance rate for that particular conference. The faculty member would also collect publications that grew out of his/her course, etc. This type of evidence would vary widely but would allow professors to tailor the expectations and outcomes of how their course is paying off for students. This evidence would be highly customized to the field and a way to promote and reward Relevant Experiential and Applied Learning (REAL) teaching endeavors. Universities in Canada are using these types of evidence for teaching. [University of Victoria's Learning and Teaching Centre](#) has a worksheet to help professors develop and customize their teaching dossiers. This link is for the [Teaching Dossier: Organizational Matrix](#).

References

- Boring, A., Ottoboni, K. & Stark, P.B. (2017). Student evaluations of teaching (mostly) do not measure teaching effectiveness. *ScienceOpen Research*, 2016(1), 1-11. Retrieved from <https://www.math.upenn.edu/~pemantle/active-papers/Evals/stark2016.pdf>
- Brockx, B., Spooren, P., & Mortelmans, D. (2011). Taking the grading leniency story to the edge: The influence of student, teacher, and course characteristics on student evaluations of teaching in higher education. *Educational Assessment, Evaluation and Accountability*, 23(4), 289-306. doi:10.1007/s11092-011-9126-2
- Fenn, A. (2015). Student evaluation based indicators of teaching excellence from a highly selective liberal arts college. *International Review of Economics Education*, 18, 11-24. <https://doi.org/10.1016/j.iree.2014.11.001>
- Golparian, S., Chan, J., & Cassidy, A. (2015). Peer review of teaching: Sharing best practices. *Collected Essays on Learning and Teaching*, 8, 211-218. doi:10.22329/celt.v8i0.4239
- MacNell, L., Driscoll, A., & Hunt, A. (2015). What's in a name: Exposing gender bias in student ratings of teaching. *Innovative Higher Education*, 40(4), 291-303. doi:10.1007/s10755-014-9313-4
- Mitchell, K., & Martin, J. (2018). Gender bias in student evaluations. *PS: Political Science & Politics*, 1-5. doi:10.1017/S104909651800001X
- Svanum, S., & Aigner, C. (2011). The influences of course effort, mastery and performance goals, grade expectancies, and earned course grades on student ratings of course satisfaction. *British Journal of Educational Psychology*, 81(4), 667-679. doi:10.1111/j.2044-8279.2010.02011.x.

- Potvin, G., & Hazari, Z. (2016). Student evaluations of physics teachers: On the stability and persistence of gender bias. *Physical Review Physics Education Research*, 12(2), 020107-1-020107-9. doi:10.1103/PhysRevPhysEducRes.12.020107
- Rosen, A. (2018). Correlations, trends and potential biases among publicly accessible web-based student evaluations of teaching: A large-scale study of RateMyProfessors.com data. *Assessment & Evaluation in Higher Education*, 43(1), 31-44.
- Uttl, White, & Gonzalez. (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. <https://doi.org/10.1016/j.stueduc.2016.08.007>
- Wagner, N., Rieger, M. & Voorvelt, K. (2016). Gender, ethnicity and teaching evaluations: Evidence from mixed teaching teams. *Economics of Education Review*, 54, 79-94. <https://doi.org/10.1016/j.econedurev.2016.06.004>