

---

**FYI: Stop Relying on Student Evaluations**

1 message

**Brad Reid** <brad.reid@lipscomb.edu>

Mon, Apr 30, 2018 at 10:43 AM

To: Craig Bledsoe <bledsoewc@lipscomb.edu>, Educational\_News <educational\_news@lipscomb.edu>, Educational\_News@yahoogroups.com, Sarah E Gibson <sarah.gibson@lipscomb.edu>, Norma Burgess <njburgess@lipscomb.edu>, Mike Fernandez <mike.fernandez@lipscomb.edu>

## Why We Must Stop Relying on Student Ratings of Teaching



iStock

By Michelle Falkoff | APRIL 25, 2018

**F**ew academics will be surprised to hear that more evidence has come out showing that student evaluations of teaching are often biased.

The [latest study](#), released this year by the American Political Science Association, found that the "language students use in evaluations regarding male professors is significantly different than language used in evaluating female professors." The study also showed that "a male instructor administering an identical online course as a female instructor receives higher ordinal scores in teaching evaluations, even when questions are not instructor-specific."

Kristina Mitchell, one of the study's authors, summarized its findings [in Slate last month](#) and concluded: "Our research shows they're biased against women. That means using them is illegal." Academic institutions must stop giving an inordinate amount of weight to student evaluations when making employment decisions, she argued, until the institutions can account for, address, and eliminate bias.

Unfortunately, there's no consensus on how best to do that, and gender isn't the only kind of bias at issue. Still, it's time for academic institutions to do better on this front.

Evidence of gender bias has been available for a long time. Even the most cursory search reveals [multiple studies](#), going back to the 1980s, of the role of gender bias in academic evaluations. More recently, researchers have investigated the effect of other kinds of bias, such as [racial](#) and [ethnic](#), and have found equally problematic outcomes.

Even biases that fall outside traditional categories of discrimination — such as student negativity toward classes they perceive as overly challenging or taxing — harm an institution's ability to use student evaluations to gauge instructors' effectiveness. Professors who are perceived to be difficult, or who teach difficult material, may receive lower evaluations despite students' often having greater success in later courses based on what they learned from those professors, as [one study](#) found.

Student evaluations have also become less reliable over the years because most institutions have switched to online systems. In 2016 the American Association of University Professors

**Angry evaluations are not that helpful in assessing what faculty members do well or where they need improvement.**

released [a comprehensive survey](#) of faculty members about teaching evaluations. which found that:

- The rate at which students were filling out evaluations has gone down precipitously in the electronic age.
- The tone of their comments has started to resemble that of internet message boards, with more abuse and bullying.
- Students who were aware of some or all of their grades tended to be harder on faculty members in both written comments and numerical assessment.

That decrease in reliability and consistency contributes to the ineffectiveness of student evaluations as a primary metric for faculty assessment.

It may be possible to improve evaluations by trying to account for, or eliminate, bias — institutions could, for example, change the questions or discount the numbers to account for bias. But the better approach is to look at alternative means of assessing faculty performance.

In particular, it's time to stop relying primarily on one approach — in this case, student evaluations of teaching — and move to a more holistic strategy in which multiple factors contribute to a more accurate, consistent, and well-rounded assessment.

I was convinced of that by my experience as director of the program in communication and legal reasoning at Northwestern University's law school. A majority of the program's faculty members are women, and our primary responsibility involves teaching a required first-year course ("Communication and Legal Reasoning") on legal analysis, writing, and research. Students don't get to choose their professor, and they receive extensive critical feedback during the semester before they fill out their course evaluations. Perhaps not surprisingly, they often use the evaluation to vent — which typically involves lashing out at the professor.

---

#### RELATED CONTENT

**[Do Their Stereotypes Affect Your Teaching?](#)**

**[How, and Why, You Should Have Students Assess Themselves](#)**

**[How to Use Student Evaluations Wisely](#)**

Angry evaluations are not that helpful in assessing what faculty members do well or where they need improvement. It can be difficult to tell whether a student's frustrations are a natural byproduct of the difficulty of the course or reveal actual teaching issues that impede learning.

## **The Inherent Unreliability of Student Evaluations**

### **How to Read a Student Evaluation**

### **Do Students' Online Ratings of Courses 'Suck' (or 'Rock')?**

---

Several studies advise a big-picture approach. The authors of the AAUP report on course evaluations, for example, recommended clearer institutional policies, more mentoring of new instructors, and multiple sources of assessment. Likewise, the University of

Michigan's [Center for Research on Learning and Teaching](#) emphasizes the importance of using more than one method — evaluating how faculty members deliver instruction, how they plan their courses, how they assess their students — and gathering feedback from students, colleagues, and supervisors.

In my own program, I've found alternative methods of assessing teaching to be extremely effective: watching faculty members teach (whether via video or in person), reviewing their course materials, reading faculty self-evaluations, and meeting with them one-on-one to discuss performance.

With a clear sense of how faculty members perceive their own courses, student feedback is easier to contextualize. It becomes possible to determine whether student concerns are legitimate or just typical of a demanding course (as first-year law-school classes tend to be).

Holding instructors to high standards is important, and student feedback is relevant. But if academic institutions do not take steps to assess teaching more holistically, they run the risk of losing talented faculty members for reasons that are not only inappropriate but may well be illegal. Moving beyond reliance on student evaluations may take more time and effort, but it will also help us ensure that we are helping instructors succeed while eliminating the possibility that bias will play a role in making or breaking their careers.

*Michelle Falkoff is a clinical associate professor of law and director of communication and legal reasoning at Northwestern University Pritzker School of Law, where she is a Public Voices Fellow through the OpEd Project.*

--

You received this message because you are subscribed to the Google Groups "Educational\_News" group.

To unsubscribe from this group and stop receiving emails from it, send an email to [educational\\_news+unsubscribe@lipscomb.edu](mailto:educational_news+unsubscribe@lipscomb.edu).