**How to Use Student Evaluations Wisely**



June 16, 2015

[*Image:*](https://www.flickr.com/photos/seattlemunicipalarchives/3024896439/in/photolist-5BiojT-snv813-3QUksd-okH7TR-2XjzSd-8vXEdm-8qLXsP-by4Hqj-7DQWhU-9NAsCc-7NMjU4-by4Fcm-bLYof6-bVjdcF-7dweuQ-9NEdXe-xx985-4gsKCR-8vUjzr-9NEykh-9Nz3As-roCaST-pcmUPh-9NFhVS-9NF1UU-8vUmHF-6ArRC-7d4XnB-9NJybd-8vXdu7-47TrtC-9NEtYQ-9M2DZZ-9NFVNX-dUSc9a-9NGKrb-69pXub-osALHa-9NH9LR-7ReoEu-8pinKn-tB9NU-9NCNYS-9NxxHd-aqbLwt-4EScCy-9NKJWL-t7D4t-haeGC-g34xP) *Scales condemned by Department of Weights and Measures, 1917 (Seattle Municipal Archives)*

When I was a doctoral student, nervously facing my first set of student evaluations, I turned for advice to my father, who was already a professor when those evaluations were first introduced. “We *should* be polling students to see what they thought of our classes,” he insisted. “Of course, their evaluations can’t signify the be-all and end-all for what constitutes effective teaching.” His position sounded sensible to me then -- and still does, now that I am a dean.

And yet -- as Stacey Patton’s recent essay, [“Student Evaluations: Feared, Loathed, and Not Going Anywhere,”](https://chroniclevitae.com/news/1011-student-evaluations-feared-loathed-and-not-going-anywhere?cid=at&utm_source=at&utm_medium=en) demonstrates -- many administrators make the mistake of using those evaluations as the sole, definitive, and objective measure of teacher quality.

I’m well aware that the value of student evaluations [is contested](http://blogs.berkeley.edu/2013/10/21/what-exactly-do-student-evaluations-measure/), but like my father, I’ve also found that they can be useful tools. At my own institution, I inherited a nuanced set of faculty-designed guidelines on the use of student evals. Among the best ideas I found in them:

* We never judge an instructor’s teaching solely on the basis of student reviews, but use them as one metric of several.
* We regularly conduct peer evaluations of teaching, even for senior professors.
* Faculty review each other’s syllabi and learning outcomes.
* We also ask recent graduates to tell us which of their courses were most helpful.
* Faculty can create their own evaluations to supplement (but not replace) the college’s standardized form. So if you as an instructor feel that a particular course might have some special circumstances (from field trips to research lab exercises) that might not be reflected in the standard evaluation, you have the option of providing us with further data.

As a dean, I’ve found that student evaluations can highlight trends and patterns in a faculty member’s teaching. For example, if, on a scale of 1-to-5, the median teaching score in a department falls in the high 4s but an instructor consistently gets 2s, then that might be a red flag worth paying attention to. Likewise, if dozens of students in a course call the exam questions “confusing,” that probably means the professor should think about rewording them. And in fact, I know many faculty members who have worked hard to resolve common criticisms raised by students in course evaluations (about things like vaguely worded test questions).

That said, the problem with the way in which student evaluations are used nationally is that they have devolved solely into a method of appraising *individual* teaching. You get ratings for *your* class; administrators review *your* performance. But few people are looking beyond those silos, and that’s not a particularly effective way to gauge teaching (or learning, for that matter). Evaluations from one course ought to be compared and contrasted with those from similar courses. Why? Because what we discover about learning in one course may be applicable to others.

On one instructor’s recent set of evals, for example, I noted that many students had praised a particular assignment: They wrote things like “it was difficult and challenging” and “yet I learned so much from it.” The instructor and I agreed: That assignment was a hit worth repeating in future iterations of the course.

Moreover, students’ comments on course evaluations can prompt more general discussions about what’s working in the curriculum -- and what’s not. If, for instance, most of the students who took “Digital Media Production I” felt they lacked the foundation they needed for “Digital Media Production II,” then we might need to make some changes. Drill down into the comments and you might find many students feeling “there is too much material; we are too rushed. We don’t have time to learn fully one thing before we are off to the next.” If you read such responses over and over, no matter who teaches the first course in the sequence, then maybe it’s time to rethink how the courses in that sequence are organized.

So student evaluations can be [useful tools](http://chronicle.com/article/How-to-Read-a-Student/129553/). But administrators should never conclude that high scores automatically designate a great teacher and low scores a bad one. Instead, they should look at other variables: What were the grades in the course? Were they much higher -- or lower -- than usual for that class and in that department? Were peer evaluations done, and, if so, what did they reveal?

Likewise we know that some students exaggerate; and aggrieved students may even lie on evals in retaliation for a perceived injustice, like a low grade. However, if 20 students in a class of 25 note -- without prompting -- that an instructor “is never at her office hours” or “keeps showing up late to his own class,” then there ought to be an investigation into those allegations. Alternately, if a large number of students say an instructor is always available and helpful during office hours, then, cheers for that faculty member.

The most important rule for administrators who are reviewing course evaluations: Follow your institution’s own procedures -- *always.* In 20 years of writing about promotion and tenure, I have heard countless complaints about what I call “purposive subjectivity.” A case in point: A tenure tracker at a community college who told me, “My chair decided I was not a good enough teacher, so if there is one bad comment out of a 100 on my evals, he claims it’s indicative of general failure on my part. Yet one of his favorites gets much lower scores and more negative comments and the chair spins them away.” I believed that faculty member; such practices are, sadly, common.

Hence the need for rigor in evaluating the evaluation process. How, exactly, do the course evaluations “count”? What does a “4” mean versus a “3”? How many students spoke up about a certain problem? Those issues need to be thoroughly discussed between administrators and faculty members.

Unfortunately, student evaluations provide the overworked and distracted administrator with an easy metric. But while the temptation to be reductive is understandable, doing so fails both faculty *and* students. If effective teaching is a universal priority (as it should be), then we need a variety of ways to measure it quantitatively and qualitatively.

Student evaluation of teaching should be one of our tools, not the only one.

[David D. Perlmutter](https://chroniclevitae.com/people/17840-david-d-perlmutter) is a professor at and dean of the College of Media & Communication at Texas Tech University. He writes the "Career Confidential" advice column for *The Chronicle.* His [book on promotion and tenure](http://www.hup.harvard.edu/catalog.php?isbn=9780674048782) was published by Harvard University Press in 2010.