



Evaluating Teaching Effectiveness Task Force

December 30, 2015

Review: Research on Student Course Surveys

Student course surveys are commonly used, at CSU and nationally, to assess teaching effectiveness. Here and elsewhere, some departments employ course surveys as the primary, or exclusive, means of conducting such assessment. Some departments use only a few questions, such as how students rank what they learned in a course and how they rank an instructor's overall effectiveness. Departments need assessments for high-stakes purposes, such as tenure and promotion decisions, contract renewal decisions for contingent faculty, and salary reviews. Given the importance of such decisions to both the institution and instructors, the attractions of this approach include a seemingly objective quantitative measure, the ready availability of data, and ease of score comparisons.

However, the research literature on course surveys has raised serious questions about whether they are valid measures of teaching effectiveness. Two salient points emerge from this research. First, a number of studies have questioned the validity of using course surveys for this purpose (Beleche, Farris, & Marks, 2012; Braga, Paccagnella, & Pellizzari, 2014; Carrell & West, 2010; Langbein, 2008; Stark & Freishtat, 2014; Weinberg, Hashimoto, & Fleisher, 2009). These studies showed that positive course survey responses correlate positively with students' expected grades in the course but inversely with learning, which was typically measured through students' grades in subsequent courses that depend upon the learning gained in the initial course. For example, Carrell and West, who examined course surveys from 10,534 students at the U.S. Air Force Academy (USAFA), based their study design on USAFA's random assignment of students across sections and its standardized core curriculum, which uses shared syllabi, examinations, and cumulative courses required of all students. They found that students in introductory courses who rated professors more positively received higher grades in those courses but lower grades in subsequent courses. Further, less experienced instructors' students consistently received higher grades in introductory courses and lower grades in subsequent courses, while more experienced instructors' students consistently received lower grades in introductory courses and higher grades in subsequent courses. "Students appear to reward higher grades in the introductory courses but punish professors who increase deep learning [as demonstrated by higher grades in subsequent courses]," the authors indicated, concluding that this fact "draws into question the value and accuracy of this practice [of using course surveys to measure teaching effectiveness]" (p. 412). Similarly, Braga, Paccagnella, & Pellizzari, who found similar patterns in course survey responses, concluded, "good teachers are those who require their students to exert effort; students dislike it, especially the least able ones, and their evaluations reflect the utility [or desired ease] they enjoyed from the course" (p. 85).

Second, some studies demonstrated consistent patterns of bias in course survey responses, patterns that negatively impact instructors of color and female instructors (MacNeil, Driscoll, & Hunt, 2014; Weinberg, Hashimoto, & Fleisher, 2009). For instance, MacNeil, Driscoll, & Hunt used the absence of in-person interaction between instructor and students in online courses to test whether course survey responses differed based on instructors' perceived gender identities. To do so, they assigned two online assistant instructors, one male and one female, two different gender identities, but equivalent qualifications, in their instructor biographies for a course. Thus each instructor taught half of his or her students identified as a male and half identified as a female, with no other differences in teaching approach. The authors found that both instructors of both genders received consistently higher course survey scores when identified as male, with gender identity the only difference. "Regardless of actual gender or performance, students rated the perceived female instructor significantly more harshly than the perceived male instructor," they said, concluding, "the continued use of student ratings of teaching as a primary means of assessing the quality of an instructor's teaching systematically disadvantages women in academia" (p. XX).

Thus using course surveys as a primary measure of teaching effectiveness poses three problems. First, much research suggests that such use is not valid. Second, course surveys produce systematically biased response patterns that negatively impact instructors of color and female instructors. Third, emphasizing course surveys in assessing teaching effectiveness undermines a crucial means of improving this effectiveness, namely fostering ongoing reflection on teaching. Such reflection is a crucial component of teaching development and should be supported by the use of various data types, consistent engagement with peers in discussions of teaching, and other forms of instructor professional development (e.g., workshops, seminars, short courses, and conferences).

However, there are potentially valid uses of course surveys that can contribute importantly to assessing teaching effectiveness and supporting teachers' development. Reid (2011) argued that course surveys should be more focused on curricular goals. In addition to supporting curricular improvement, course surveys can provide information instructors can use to improve instructional approaches. When course survey responses offer information on issues students can evaluate, such as clarity of presentation, pacing and accessibility of assignments, and perceived relevance of course texts, they often provide instructors with information that can be used to improve curriculum design or instructional delivery. For departments needing to replace the direct use of course surveys with a more valid – but still manageable – measure of teaching effectiveness, this feature of course surveys makes them a potentially valuable tool, when used to prompt instructor reflection and revision of course design or instructional delivery.

Instructors can be asked to summarize their course survey responses, briefly explain changes they've made to address issues needing improvement, and summarize any relevant responses from subsequent surveys. This material can be provided in one page per academic year. By asking instructors to produce such surveys each year, departments can track progress (or lack thereof) in instructors' course survey responses and in their efforts to improve teaching effectiveness. Further, use of this approach will promote sustained, cumulative reflection on teaching. It may scaffold instructors' and departments' work toward more robust forms of reflection on teaching, such as teaching portfolios. It is likely to be

most effective when supported by professional development endeavors such as teaching effectiveness consultations, workshops, seminars, conferences, and the like.

REFERENCES

Beleche, T., Farris, D., & Marks, M. (2012). Do course evaluations truly reflect student learning? Evidence from an objectively graded post-test. *Economics of Education Review, 31*, 709-719.

Braga, M., Paccagnella, M., & Pellizzari, M. (2014). Evaluating students' evaluations of professors. *Economics of Education Review, 41*, 71-88.

Carrell, S. E., & West, J. E. (2010). Does professor quality matter? Evidence from random assignment of students to professors. *Journal of Political Economy, 118.3*, 409-432.

Langbein, L. (2008). Management by results: Student evaluation of faculty teaching and the mis-measurement of performance. *Economics of Education Review, 417-428*.

MacNeil, L., Driscoll, A., & Hunt, A. N. (2014). What's in a name: Exposing gender bias in student ratings of teaching. *Innovation of Higher Education, online, 05 December 2014*.

Stark, P. & Freishtat, R. (2014). An evaluation of course evaluations. Unpublished.1-15.

Weinberg, B. A., Hashimoto, M. & Fleisher, B. M. (2009). Evaluating teaching in higher education. *Journal of Economic Education, Summer 2009, 227-261*.