



# Instructor and Institutional Assessments of Teaching Effectiveness: Faculty Views

Chandra Turpen\*, Charles Henderson\*, and Melissa Dancy†

\*Department of Physics, Western Michigan University, Kalamazoo, MI 49008

†Department of Physics, University of Colorado, Boulder, CO 80309



## Introduction

**Strong assessment practices are broadly recognized as integral to effective instruction** [1]. With the current national emphasis on accountability in higher education [2], we suggest PER contributes to the conversation. The purpose of this project is to present faculty perspectives about the methods that both they as instructors and their institutions use to evaluate teaching effectiveness based on a set of 35 interviews conducted with faculty familiar with Peer Instruction [3].

**RELEVANT FOR EDUCATIONAL CHANGE:** Changing assessment methods is viewed as an important step in promoting the use of more innovative, research-based instructional strategies (RBIS) [4]—an important lever for instigating instructional change. However, we argue that evaluation methods can also play an important role in encouraging (or discouraging) the continued use of RBIS. Specifically, the assessment methods used by institutions and instructors **will influence instructors' determinations about whether their use of RBIS is working or worth the effort**. In this way, **assessment practices influence judgments about the relative advantage and compatibility** [5] of educational innovations, and thus the continuation or discontinuation of RBIS use.

**COMMENTS ON STUDENT EVALUATIONS OF TEACHING (SETs):** There is a large body of literature and considerable disagreement about the value of SETs. Some researchers argue that SETs are valid because they are correlated with student achievement [6, 7]. Others argue that the SETs are merely popularity contests and that the use of SETs is a barrier to more effective teaching, since faculty are hesitant to be more challenging due to fears of lower SETs [8, 9]. **There is almost universal agreement, however, that assessment of something as complicated, nuanced, and important as teaching effectiveness should be made using multiple methods** [e.g., 4, 7, 10].

## Implications for PER

### Summary of Results

1. Faculty and institutions use almost non-overlapping methods to evaluate teaching.
2. Faculty do not believe that the methods used by institutions are valid.
3. Neither faculty nor institutions make much use of standardized measures of student learning (e.g., existing conceptual inventories).
4. Significant fraction of faculty are already engaging in forms of formative assessment.

Both faculty and institutions could benefit from **broadening the assessment sources** that they use to determine teaching effectiveness—specifically through the **use of nationally-normed assessments** that would allow for inter-institutional and across classroom comparisons.

Change agents could work to **pressure university administrators to diversify the metrics used to assess teaching effectiveness**—at least utilizing some measures of student learning (i.e. scholarship of teaching and learning projects, or standardized assessment measures).

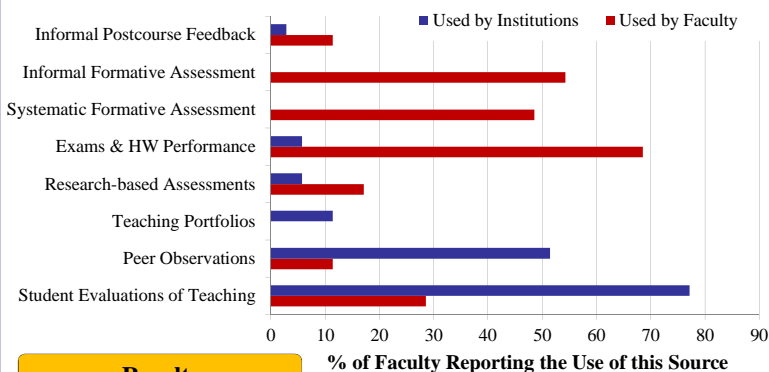
The PER community can lend a **moderating voice encouraging the coordinated use of formative assessment and standardized research-based summative assessments** in determining teaching effectiveness.

Many faculty seek out and value formative assessment sources—this may be a powerful lever for change agents. **Faculty find what students do and say in class as compelling evidence of teaching and learning.**

## What sources of assessment information are used?

### Sources of Assessment Information

<b>Student Evaluations of Teaching:</b> All structured collection of student evaluative feedback about a course.
<b>Peer Observations of Teaching:</b> Having peer faculty or university administrators sit in on an instructor's course and provide feedback (often through written reports and sometimes verbal feedback in a face-to-face meeting).
<b>Teaching Portfolios:</b> Having instructors self-report how they teach, sometimes providing references to the research literature about the evidence for the success of the instructional methods they are using.
<b>Research-based assessments:</b> Typically involving pre/post-testing through the use of MC conceptual inventories.
<b>Student Performance on Exams, Quizzes, or Homework:</b> Using students' performance on exams, quizzes or homework and taking this performance as an indicator of the success of instruction.
<b>Systematic Formative Assessment:</b> Gathering a sampling of students' performance on an in-class task as an indicator of teaching effectiveness (e.g. having students submit votes, walking around the room and observing the solutions of multiple groups).
<b>Informal Formative Assessment:</b> All other forms of formative assessment, such as verbal comments in class or office hours, the look of confusion in students' eyes, whether or not students are awake, or whether or not students are asking questions.
<b>Informal Post-course Feedback from Students:</b> Spontaneous informal comments from student(s) semesters or years after the student(s) have left the class with feedback on the course or the instructor.

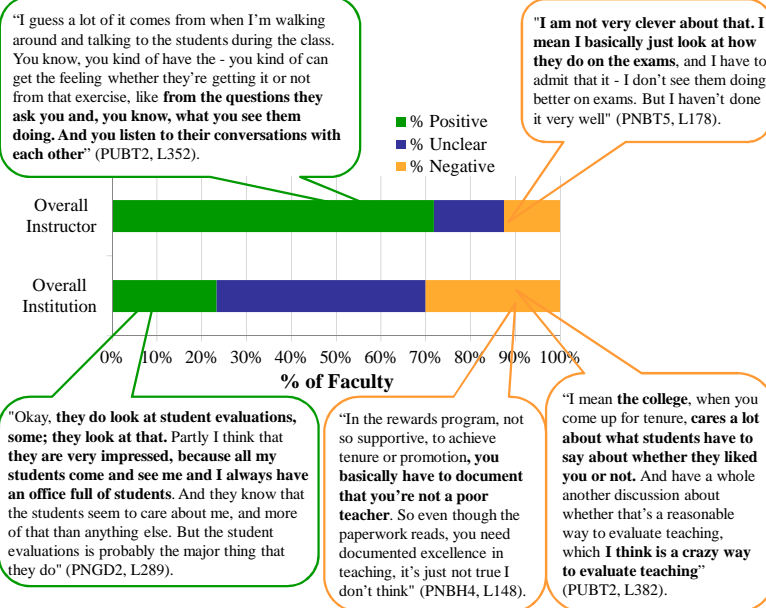


## Results

- Institutions primarily rely on student evaluations of teaching (particularly the quantitative measures) to determine teaching effectiveness.
- Faculty primarily rely on systematic and informal formative assessment as well as students' performance on exams, quizzes, and homework to determine teaching effectiveness.
- **Faculty and institutions use almost non-overlapping methods to evaluate teaching.**
- **Neither faculty nor institutions make much use of standardized measures of student learning.**
- Even though information from SETs is made available to faculty, very few faculty find this information useful (those that do focus on long-answer student comments).

## How do faculty evaluate assessment practices?

The researchers holistically assessed whether the interviewee felt positively or negatively about how well they are able to evaluate their own teaching effectiveness. Similarly, the researchers holistically assessed whether the interviewee felt positively or negatively about how well they think their institutions are able to evaluate teaching effectiveness.



## Results

- Faculty seem to have more confidence in their own assessment practices than they do in the assessment practices of the institution.
- Of the 30 faculty who discussed institutional assessment, one third said their institution uses only one measure of teaching effectiveness. SETs were reported to be the most common single source that institutions relied upon (8/10).
- An additional 15 faculty described their institutions as only relying on two assessment measures for judging teaching effectiveness. The most common combination of two assessment measures for institutions was SETs paired with peer observations (12/15).
- Our preliminary findings suggest that even institutions that use multiple measures tend to give significant (or sole) prominence to SETs over other sources.

## References & Acknowledgements

This presentation is based upon work supported by the National Science Foundation under Grant No. 0715698. We thank all of the professors who dedicated their time and shared their experiences with us as part of this research study. We are grateful to our advisory board (Dr. Robert Beichner, Dr. Paula Heron, Dr. Priscilla Laws, and Dr. David Maloney) for their early feedback on the procedures for sampling interviewees and the interview protocol used. We also appreciate feedback from the PER group at University of Colorado at Boulder on this work.



1. J. M. Alkins and J. E. Coffey (eds.), *Everyday Assessment in the Science Classroom*, Arlington, VA: National Science Teachers Association Press, 2003.
2. P. T. Ewell, *Making the Grade: How Boards Can Ensure Academic Quality*, Washington, DC: Association of Governing Boards of Universities, 2006.
3. E. Mazur, *Peer Instruction: A User's Manual*, (Prentice Hall, Upper Saddle River NJ, 1997).
4. E. Seymour, "Tracking the process of change in US undergraduate education in science, mathematics, engineering, and technology," *Science Education*, **86**(1), 79-105 (2002).
5. E. M. Rogers, *Diffusion of Innovations*, 5<sup>th</sup> ed., New York: Free Press, 2003.
6. H. Wächter, "Student Evaluation of College Teaching Effectiveness: a brief review," *Assessment & Evaluation in Higher Education*, **23**(2), 191-212 (1998).
7. K. A. Feldman, "Identifying Exemplary Teachers and Teaching: Evidence from Student Ratings." In R. P. Perry & J. C. Smart (Eds.), *The Scholarship of Teaching and Learning in Higher Education: An Evidence-based Perspective* (pp. 93-143), Dordrecht: Springer (2007).
8. R. Arum and J. Roska, *Academically Adrift*, Chicago: University of Chicago Press, 2011.
9. V. E. Johnson, *Grade Inflation: A Crisis in College Education*, New York: Springer (2003).
10. W. A. Anderson et al., *Science* **331**(6014), 152-153 (2011).