

Revisiting Course Evaluations: Provost's Initiative on Evidence of Effective Teaching

Learning Outcomes

Participants will be able to:

1. Describe the Provost's Initiative on Evidence of Teaching Effectiveness
2. Outline best practice research on "course evaluations/student end of semester surveys"
3. Outline what OU and other institutions are currently using for "Student End of Semester Surveys"
4. Review possible recommendations for OU for Student End of Semester Surveys
5. Share experiences and insights on "Student End of Semester Surveys"

Purpose of Provost's Initiative on Evidence of Teaching Effectiveness

To enhance Student, Faculty and Institutional Success

- Improve course delivery, instruction and enhance student learning
- Support the value and importance of teaching and learning
- Develop and implement research-based, best practices for “course evaluations” and additional evidence of teaching effectiveness
- Create a streamlined “course evaluation” process that is effective and efficient
- Use the results and data as evidence for:
 - teaching improvement
 - curriculum improvement
 - tenure and merit decisions

Goals of Initiative

- To revise and enhance the “course evaluation” (student end of semester survey)
- To promote additional evidence-based practices that are used to demonstrate teaching effectiveness for
 - formative (growth and improvement of teaching)
 - summative (decision-making, tenure, promotion and merit) assessments.

Task Forces- Long Term Project

This initiative is a long term (3-5 year) project

- Task Force 1- Course Evaluations Revisited
- Task Force 2- Additional Evidence of Teaching Effectiveness

Task Force 1 -2016-2017- diverse representation from college and professional schools (nominations from deans and then co-chairs select)

Task Force 1-

Course Evaluations Revisited

- Part A- Course Evaluation Questions
 - Rename (move away from Course Evaluation)
 - Ratings- 5 highest- 1 lowest
 - Demographic questions
 - 12-15 core common questions
 - Additional questions per school/department/individual faculty
- Part B- Course Evaluation System
 - Format: paper-pencil, online, Scantron, other
 - How are results analyzed and summarized
 - If online- how to ensure adequate response rate
 - Where is it “housed”- e.g.- Institutional Research and Assessment
 - Who at school/department level is involved

Task Force 2-

Additional Evidence of Teaching Effectiveness

How to use Evidence of Teaching Effectiveness

- Part A- How is information about “Course Evaluation” shared with faculty
 - How is it used for formative (ongoing growth and improvement) and summative (decision-making-PTR, annual reviews, merit, course scheduling)
 - Who meets and shares it with faculty (e.g. mentor vs review committee-conflict of interest)
- Part B- Additional Evidence of Teaching Effectiveness
 - How to triangulate evidence of teaching effectiveness
 - What other evidence can/should be used to demonstrate teaching effectiveness
 - How/who uses this information
 - Develop “peer evaluation” with checklist protocol
- Part C- Teaching and Tenure, Promotion, Merit
 - Formative evaluation- mentorship
 - Summative evaluation- how to use information for decision-making
 - Review and revise tenure documents, merit, etc

Best Practices in Evidence of Teaching Effectiveness

For make formative (improvement) and summative (tenure, decision-making, merit) evaluation of teaching

- There should be a triangulation of evidence (multiple types of examples- student surveys, narratives, protocol observations, portfolios, syllabi, exit interviews, etc.)
- What we call “course evaluations” should not be used as the sole or major determining factor in decision-making
- What we call “course evaluations” are actually “student satisfaction surveys” and provide important information but do not provide a “complete, accurate, valid” data if used alone

Current State of Affairs at OU: “Course Evaluations” - Student End of Semester Surveys

- More than 25 different “evaluations” with no standardization including a range of online (including five different systems), Scantron, paper-and-pencil
- Broad range of types of questions
 - Questions that are rated plus comments
 - Some are only questions
 - Some are only comments or essays style questions
- Most have scoring of 1-5
 - Most have 5 as the highest
 - Some have 1 as the highest
 - Some have no ratings

How it is used

- Range of how information is shared back with faculty
 - Some get raw data (including handwritten comments)
 - Some get it analyzed
- Range of when faculty see results
 - Some given it during next semester
 - Some can access it online
 - Some only see it when they ask for it
 - Some do not even know they can see it
- Range of who and how it is administered
 - One school does not allow chairs/dean to view
- Range of how it is used
 - Few use it for formative reviews with chairs or mentors
 - All use it for tenure purposes-
 - Some only use the “course evaluation” survey for decision-making
 - Some use it with other data
 - Some only use individual questions

Faculty Focus Group

- In winter 2015 held faculty focus group to hear their perceptions on “course evaluations”
 - Majority did not realize all these variations
 - Majority felt that we should revisit and revise what and how we are using them

How are course evaluations/student end of semester surveys being used in your department?

- Discuss
- Share

How do you feel about the “course evaluations/student end of semester surveys”

- Share
- Discuss

Are they helpful for Formative and Summative Evaluation

- Share
- Discuss

Best Practice Recommends (Arreola, R. ;

2007; Berk, 2013, 2008; Blumberg, P. 2014; Buller, J., 2012; Center for Research in Learning and Teaching ,2012; Centra, J. ,1993; Chism, N. ,2007)

- Multiple Sources of Evidence to demonstrate teaching effectiveness
 - SRI/SETs/SESS
 - Peer Observations using a protocol checklist
 - Reflective Practice- narrative/dossier- instructor's explanation and descriptions
 - Syllabus, assignments, assessments, learning outcomes
 - Grade distribution, DFWI
 - Scholarship of teaching and learning (SoTL)
 - Grants, awards, publications on teaching
 - Other samples

Research on “Course Evaluations”

- Student satisfaction surveys not evaluation of teaching
- Some research suggests that lower ratings are reported for women, under-represented faculty, required courses, challenging courses, students who do not do well in courses

(Benton & Cashin, 2014; Boring, 2016; Centra, 2005; Centra, 2000; Isey, 2007; Shevlin, 2000; Steiner, 2006; Superson, 1999; Zabaleta, 2007)

Research findings acknowledge

(Arreola, R. ,2007; Benton, S. & Cashin, W. 2012; Buller, J., 2012; Hativa, 2013)

- SRI/SETs/Student Surveys are the most commonly used sample of evidence of “teaching effectiveness” for decision-making purposes at universities
- Frequently called “course evaluations” but in reality students do not evaluate courses or instructors, rather they provide their feedback, perceptions and satisfaction
- Yet, NOT the best indicator as SRI/SETs are more of a “student satisfaction survey” than measure of teaching effectiveness or whether students actually learned. Some research suggests that:
 - Easy grading= higher scores
 - Elective courses (vs. required courses)= higher scores
 - Online responses- lower return rate
 - More advanced and committed students (e.g. seniors and graduate students) = higher scores

Summary of Results of Interviews with 8 other Institutions

- All institutions surveyed have recently or are currently reviewing and revising systems and surveys (one of the most discussed issues nationwide)
- Majority are all online or shifting to online surveys
- Range of Surveys- All have either “standard survey” for university or “some standard questions” for university
- All stress importance of multiple measures of evidence to evaluate teaching, however, majority still find that the “survey” is predominantly used and over emphasized
- Range of “who” oversees it-
 - Office of Institutional Research and Assessment
 - Office of Online Learning

What would you like to see happen at OU with course evaluations/student end of semester surveys?

- Share
- Discuss

Possible Recommendations based on Best Practice Research- Task Force to Determine

- Rename- not “course evaluation” but perhaps Student End of Semester Survey (SESS)
- SESS is only one of multiple-measures (triangulation of evidence) of teaching effectiveness is used for evaluation for formative (growth and development of teaching) and summative purposes (decision-making, tenure, promotion)
- Move to online surveys with strategies to increase online response rate
- Have central administration system (perhaps through Institutional Research) but departmental level has responsibility and control
- Standardize rating scales- 5 highest- 1 lowest
- Have some questions that are consistent across the university and have additional questions that can be departmental, course, instructor, type of course specific

Evidence of Teaching Effectiveness

PROVOST'S EVIDENCE OF TEACHING EFFECTIVENESS INITIATIVE

The Provost's Evidence of Teaching Effectiveness Initiative at Oakland University has been created to explore and examine best research-based practices for the evaluation of teaching (formative and summative). The goals of this long term project are to:

Task Force 1 – Redesigning "Course Evaluations" Develop recommendations using research-based practices in designing and implementing "Student End of Semester Surveys" (currently referred to most frequently as "Course Evaluations")

Task Force 2 – Designing Other Measurements for Teaching Effectiveness Develop recommendations for the use of multiple measures and triangulation of other evidence of teaching effectiveness for formative and summative evaluation

FORMATIVE EVALUATION OF TEACHING

- Used by faculty (and mentor/coach) to develop and enhance teaching effectiveness
- Ongoing, continual
- Reflective practice

SUMMATIVE EVALUATION OF TEACHING

- Used for decision-making purposes
- Tenure, promotion, merit, course assignments, continuation of employment for part-time

RESEARCH ON EVALUATION OF TEACHING recommends that evidence of teaching effectiveness be triangulated using multiple measures.

STUDENT END-OF-SEMESTER (SES) SURVEYS

- While an important feedback tool, students do not "evaluate" faculty; they provide information on their perceptions and satisfaction of courses and instruction
- These ratings should only be used as one of multiple measures of teaching effectiveness

CONTEXTUAL NARRATIVE

- Accompanies SES Surveys to address factors that can impact ratings (e.g., teaching methodology, gender, race, class size, controversial topics, and grading)
- Analyzes trends over time

PEER PROTOCOL OBSERVATIONS

- Colleagues use protocol prompts
- Observation evaluates teaching practices grounded in research

TEACHING PORTFOLIO

- Teaching philosophy
- How learning outcomes have been developed and met
- Syllabi
- Assignments/Rubrics
- Course (re)design
- Program development
- Mentoring/Advising
- How teaching, research and service are integrated
- Letters of support and recommendations
- Awards

FACULTY DEVELOPMENT

- Participation in workshops, seminars and conferences on teaching and learning
- Service involvement in teaching and learning

SCHOLARSHIP OF TEACHING AND LEARNING

- Teaching and learning research
- Research on their own teaching practice
- Grants for teaching and learning
- Publications in teaching and learning
- Conference presentations in teaching and learning

"Course Evaluation" Revisited

What are they called?

"Course evaluations" is often the term used for **student end-of-semester (SES) surveys**.

Preferred Terms

SES Surveys (Student End-of-Semester Surveys)

SRI: Student Ratings of Instruction

SET: Student Evaluation of Teaching

Additional Terms

SRT: Student Ratings of Teaching

SIR: Student Instructional Report

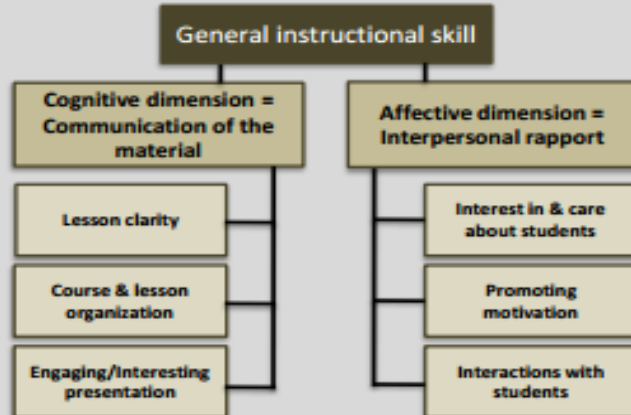
SRI: Student Response to Instruction

S RTE: Student Ratings of Teaching Effectiveness

Recommendations for Evaluating Teaching Effectiveness

Students can provide insight and important feedback about a course and instructor, and may be a reliable and valid measure if it is part of a larger, comprehensive evaluation system that triangulates evidence of teaching effectiveness. A student survey alone, however, should not be seen as the single measure of teaching effectiveness. Students do provide us with the information, but it is up to faculty, departments and tenure committees to provide the evaluation component and decision-making.

Model 1 – Dimensions of Effective Teaching



Hativa, N. (2013). *Student Ratings of Instruction: Recognizing Effective Teaching*.

Model 2 – Student Ratings

Instructional design

Instructional delivery

Instructional assessment

Course management

Self-reported course impact on student

Alternative and supplementary teaching/learning environments

Arreola, R. (2007). *Developing a Comprehensive Faculty Evaluation System*.

Recommended Process for University-Wide Student End-of-Semester (SES) Surveys: Sample of Revised System

CENTRALIZED SYSTEM

A centralized system creates consistency in how students provide feedback on courses. This consistency could include:

Online Format

Taken online, with analysis available for faculty.

Standardized and Individual Questions

Likert scale questions and open answers.
Include consistent university-wide questions and specific questions by department, course, and modality.



Centralized Delivery of Online Student End-of-Semester Surveys

Centralized system is solely for organization, dissemination and data analysis.

- Available for one week
- Three reminders that survey is open
- For face-to-face classes, in-class time is provided to students to complete on a tablet, smartphone or laptop.



Centralized Analysis of Data

The online format allows for the most efficient and effective summary of data and comments.

Centralized Dissemination

Results are sent to faculty within a short period of time following the end of the semester.



Faculty and Department Use Student End-of-Semester Surveys as ONE PIECE OF EVIDENCE among triangulation of additional samples of teaching effectiveness (peer observations, syllabi, etc.).

These data should be used for two kinds of evaluation:

- **Formative** – ongoing growth and improvement
- **Summative** – decision-making (tenure and promotion, reappointment)

Department and faculty are ultimately responsible for and have control of all decision-making using data.

FACULTY
REPRESENTATION
through a narrative
or dossier.

DEPARTMENT
REPRESENTATION
through a narrative
or dossier.

Debrief and Wrap-Up

- Questions
- Comments
- Take-aways

References

- AAU&P (2005) Observations on the Association's Statement on Teaching Evaluations"
- Arreola, R. (2007) Developing a Comprehensive Faculty Evaluation System: A Guide to Designing, and Operating Large-Scale Faculty Evaluation Systems. San Francisco, CA: Jossey-Bass Inc. Publishers.
- Benton, S. & Cashin, W. (2012) "Student ratings of teaching: A summary of research and literature" *IDEA Center*. Manhattan, Kansas: The IDEA Center
- Berk, R. (2013) Top 10 Flashpoints in Student Ratings and the Evaluation of Teaching. Sterling, VA: Stylus Publishing, LLC.
- Berk, R. (2006) Thirteen Strategies to Measure College Teaching. Sterling, VA: Stylus Publishing, LLC.
- Bess, J. & Dee, J. (2012) Understanding College and University Organization: Theories for Effective Policy and Practice. Sterling, VA: Stylus Publishing, LLC.
- Blumberg, P. (2014) Assessing and Improving Your Teaching: Strategies and Rubrics for Faculty Growth and Student Learning. San Francisco, CA: Jossey-Bass Inc. Publishers.
- Boyer, E. (1990) Scholarship Reconsidered: Priorities of the Professoriate. San Francisco, CA: Jossey-Bass Inc. Publishers.
- Boring, A., Ottoboni, K, Stark, P. B. (2016). Student evaluations of teaching (mostly) do not measure teaching effectiveness. *ScienceOpen Research*.
- Braskamp, L. & Ory, J. (1994) Assessing Faculty Work: Enhancing Individual and Institutional Performance. San Francisco, CA: Jossey-Bass Inc. Publishers.
- Buller, J. (2012) Best Practices in Faculty Evaluation: a practical guide for academic leaders. San Francisco, CA: Jossey-Bass Inc. Publishers.

Cont.

- Cambridge, B. & Babb, M. & Cook, C. & Gale, R. & Lieberman, D. & Roen, D. & Wert, E. (eds.)(2004) Campus Progress: Supporting the Scholarship of Teaching and Learning. Washington, DC: American Association for Higher Education.
- Center for Research in Learning and Teaching (2012) "Guidelines for evaluating teaching". University of Michigan, Ann Arbor, Michigan: CRTL
- Centra, J. A., Gaubatz, N. B. (2000). Is there gender bias in student evaluations of teaching? *The Journal of Higher Education*, 71(1), 17-33.
- Centra, J. A. (2003). Will teachers receive higher student evaluations by giving higher grades and less coursework? *Research in Higher Education* 44(5), 495-518.
- Centra, J. (1993) Reflective Faculty Evaluation: Enhancing Teaching and Determining Faculty Effectiveness. San Francisco, CA: Jossey-Bass Inc. Publishers.
- Chism, N. (2007) Peer Review of Teaching: A Sourcebook. San Francisco, CA: Jossey-Bass Inc. Publishers.
- Cook, C. & Kaplan, M. (eds.) (2011) Advancing the Culture of Teaching on Campus: How A Teaching Center Can Make A Difference. Sterling, VA: Stylus Publishing, LLC.
- Felder, R. & Brent, R. (2004) "How to evaluate teaching". *Chemical Engineering Education*, 38 (3) pg. 200-202
- Fry, H. & Ketteridge, S. & Marshall, S. (eds.) (2009) Teaching and Learning in Higher Education: Enhancing Academic Practice. Abingdon, Oxon UK: Routledge Publisher.
- Isley, P., & Singh, H. (2007). Does faculty rank influence student teaching evaluations? Implications for assessing instructor effectiveness. *Business Education Digest*, 16, 47-59.
- Seldin, P. (2006) Evaluating Faculty Performance: A Practical Guide to Assessing Teaching, Research, and Service. San Francisco, CA: Jossey-Bass Inc. Publishers.
- Shevlin, M., Banyard, P., Davies, M., & Griffiths, M. (2000). The validity of student evaluation of teaching in higher education: Love me, love my lectures? *Assessment & Evaluation in Higher Education*, 25, 397-405.
- Superson, A. M. (1999). Sexism in the classroom: The role of gender stereotypes in the evaluation of female faculty. *APA Newsletter on Feminism and Philosophy*, 99(1), 46-51.
- Zabaleta, F. (2007). The use and misuse of student evaluations of teaching. *Teaching in Higher Education*, 12(1), 55-76.