Building a Certified Peer Observation Program: *Improving Teaching and Removing Biases*

Rebecca Johnston, Associate Director
Roger Runquist, Director
Carl Ohrenberg, Assistant Director
Lindsay Linsky, Assistant Director

Center for Teaching, Learning, and Leadership
Your class is going to be formally observed today.

What is your reaction?

PollEv.com/catherinelin547
Purpose

• Examine Student Evaluations of Teaching (SETs) in extant literature
• Discuss benefits of peer observation
• Explain genesis of program development
• Present elements of the program
• Provide opportunity for questions and discussion
Activity:
https://padlet.com/cllinsky/9v0a rb4i8lqkalm8

List something strange, unusual, or just plain crazy that you’ve seen or heard of on any student evaluation.
Student Evaluations of Teaching

• SETs primary source of data to evaluate teaching, merit raises, and promotion/tenure

• Use of SETs as primary high-stakes measure is problematic
  – Issues include invalidity, unreliability, bias, and small sample size

• Despite this, SETs remain the primary tool
  (Lince, 2017; Boring, Ottoboni, & Stark, 2016; Spooren, Brockx, & Mortelmans, 2013; Kelly, 2012; Storage, Horne, Cimpian, & Leslie, 2016; MacNell, Driscoll, & Hunt, 2015; Hornstein, 2017; Braga, Paccagnella, & Pellizzari, 2014)
Problems with SETs

• Do not legitimately assess *teaching effectiveness*
  
  – Measure *student opinions* of teaching effectiveness

  – Students are *not qualified* to assess teaching effectiveness

  – SETs gather collective views of student experience in a *single course* with a particular faculty
  
  • Not a global evaluation

(Hornstein, 2017; Arreola, 2007; Hativa, 2013; Linse, 2017)
Problems with SETs

• Many behaviors and skills are involved in teaching effectiveness
  – Knowledge and content expertise
  – Teaching methods
  – Course design and organization
  – Quality of course materials
  – Assessment instruments and methods
  – Grading practices
  – *Students not qualified to assess (not trained in course design and instruction)*
    (Marsh, 2007; Svinicki & McKeachie, 2011; Berk, 2013)
Problems with SETs

• Female instructors evaluated more critically
  – Gender bias accounts for .50 on a 5-point scale
  – Female instructors evaluated more critically even when gender randomly assigned
  – Bias influences even objective measures, i.e. time to return work
    (Macnell, Driscoll and Hunt, 2015)
Problems with SETs

- Other documented biases:
  - Faculty rank
  - Student motivation
  - If course is required or elective
  - Anticipated grade
  - Upper vs. lower division
  - Class size
  - Academic discipline
  - Student workload

(Braskamp & Ory, 1994; Marsh & Dunkin, 1992, 1997; Centra, 2003, 2009; Hoyt & Lee, 2002a, 2002b)
Problems with SETs

• Valid evaluations cannot be made using parametric statistics
  – Data is categorical
  – Parametric analyses make assumption of symmetry in distribution not reflected in SETs samples
Problems with SETs

• Possible to have instructors very effective at teaching, but very low scores on SETs

• The reverse is also true

(Steury, et al., Auburn University)
Problems with SETs

• Use of SETs for faculty hiring, promotions, merit increase
  – Can encourage:
    • poor teaching
    • result in grade inflation
    • empower students to shape faculty behavior
      (Stroebe, 2020)

• Higher ed. continues to use SETs regardless of problems
  (Hamermesh & Parker, 2005)
Studies in Support

• SETs positively correlated with teaching effectiveness and student learning
  – Useful for measuring *in aggregate*

– To compare multiple individuals across departments or programs
  • Do not distinguish among individual teachers, especially through a single class
Studies in Support

• Scores from multiple courses taught over multiple years
  – May give some indication of individual teacher effectiveness
  – Data over time
  – Trends
History of SETs

• In the 1970s, used primarily for *formative* assessment
  – How can teaching be improved?

• Since then, have become primary means of *summative* assessment
  – Single high-stakes summations
  – *Absolute* measures of quality of teaching

(Hornstein, 2017)
Recommendations

• SETs are poor indicator of individual instructor’s overall effectiveness
• Should be used formatively
• Ratings should be evaluated from multiple courses across time
• If poor participation rate, should not be used
• Should be used alongside other forms of evaluation
Why are They Still Primary Measures?

• Easily quantifiable data

• Easily repeatable

• What are alternate forms of evaluation?
Alternate Forms of Evaluation

- Peer review of course material
- Peer review of instruction
- Review by expert outside evaluators
  - CTL
- Teaching scholarship
- Learning outcome measures
- Teaching portfolios
Benefits of Peer Observation

• Creates an opportunity for reflection
  – For both faculty and CPO
• Provides a data point to balance SETs
• Can be used to document teaching effectiveness
• Formative focus on faculty growth
• Trained observers focused on analysis of teaching behaviors
Program Development

- Our CTL has often offered to do an observation but there has been no formalized process

- A presentation by St. Leo at SoTL Commons offered insights into an effective program

- Faculty interest at UNG to have a counterpoint for SETs
Developing a Peer Assessment Program
CPO Development Process

Developing a Peer Assessment Program

- Why Create the Program?
- Who is Involved?
- What is the Observation Process?
- What Logistics Do You Need to Consider?
- How Will You Communicate the Results?
CPO Development Process

Developing a Peer Assessment Program

Why Create the Program?

- What is the Purpose of the Program?
- Summative or Formative Assessment?
CPO Development Process

Developing a Peer Assessment Program

Who is Involved?

- Who is in Charge of the Program?
- Who Will Be the Observers?
- What Faculty Will Be Observed?
CPO Development Process

Developing a Peer Assessment Program

- What is the Observation Process?
  - How and When Will the Observer Meet With the Faculty?
  - What Evaluation Tool Will Be Used?
CPO Development Process

Developing a Peer Assessment Program

What Logistics Do You Need to Consider?

How Will You Train the Observers?
What are the Time Commitments?
Other Considerations?
CPO Development Process

Developing a Peer Assessment Program

How Will You Communicate the Results?

- How Will the Results Be Reported Out?
- Who Will Have Access to the Results?
Our Process
Factors for our program:

- Faculty development opportunity
  - Enhanced by self-awareness
- Flexibility to use the review is up to the faculty member
  - Remember, it is formative
CPO Training

Training

CPO Last Year

No

Terminology

Overview of the Process

Avoiding Biases

Yes

Introduction of the Instrument

Pre-Meeting Processes

Case Study and Alignment/Reliability

Add to List of Available CPOs

UNG UNIVERSITY OF NORTH GEORGIA
Observation

1. Observation Process
2. CPO Makes Initial Contact, Send Pre-Observation Form
3. Pre-Meeting and setting date of observation
4. Perform Observation, Complete Observation Form
5. Send Post-Observation to Faculty, Set Meeting Date
6. Meet and Discuss Observation and Post Observation
7. Repeat?
8. Yes, Repeat Min. of 1 Time
9. Provide information for internal analysis
Activity:

https://padlet.com/cllinsky/4e9rm

t4azm43bluq

Other than SETs what experiences have helped *the most* to improve your teaching?
Discussion

• Do you currently have a peer observation program at your institution? If so, please share!

• If you do not have a CPO program, which entities at your institution could help to develop one?

• Could you develop a CPO program within your department? What might be some barriers and/or solutions?
References


References


Hoyt, D.P., & Lee, E. (2002a). *Basic data for the revised IDEA system* (IDEA Paper No. 12). Manhattan, KS: The IDEA Center at Kansas State University, Center for Faculty Education and Development.

References


