ASSESSING FACULTY DEVELOPMENT IN A WRITING/SPEAKING QEP: Direct and Indirect Measures

Julie Howdeshell, Director of Quality Enhancement
THE UNIVERSITY OF SOUTHERN MISSISSIPPI
Session Preview

› Current practices and issues in assessing faculty development
› Experiences of the USM QEP
  • Focus on process
  • Use of indirect and direct measures
  • Connecting faculty development and student learning outcomes
Current Practices in Assessing Faculty Development

- Indirect measures
  - Usage numbers
  - Satisfaction surveys
  - Self-reported behaviors
  - Student ratings of instruction
Benefits of Indirect Measures

- Easy to collect
- Useful
- Provide fast feedback so changes can be implemented more quickly
- Help triangulate data
Call for Direct Measures

- **Centra - 1976:**
  - “Whether institutions will continue to sustain development programs may very well depend on the demonstrated impact of the programs” (p. 79).

- **Hines - 2007:**
  - “a growing interest in measuring outcomes in teaching and learning”
  - overall findings: reliance on satisfaction surveys and self-reported data without linkage to further analysis
Reasons for Lack of Direct Measurement

- Chism and Szabo
  - fairness of evaluating second-order impacts
  - lack of time, resources, or expertise
  - Takes away from focus of developing faculty – “we’re not a research institute”
  - Studies would necessarily include the faculty participants who are already giving of their limited time to participate
Reasons for Lack of Direct Measurement

- Inherent Problems in Design & Methodology of Direct Measures
  - small sample sizes
  - lack of baseline data
  - inability to conduct pre/post tests given that faculty members generally participate one semester and begin implementation the next semester with a new group of students
  - inability to control for the multitude of variables at work
“A Tiny Dot”

“A single faculty development program is likely to produce only a tiny dot on the mosaic of student educational experience…”

“Such dots should be searched for whenever a reasonable possibility exists of finding a relationship between faculty development and impact upon student learning.”

> Elbe and McKeachie (1985)

“Programs directed at improving the teaching of a particular skill, such as writing, probably offer the most promise for such assessment”

“Do the best we can with less than perfect research designs and measures

(Elbe and McKeachie, 1985, p. 178-179)
Assessing Faculty Development at USM

- Nascent Stages of QEP Assessment
- Development of New Assessment Tools and Processes
  - Focus: Connecting Faculty Development Participation to Student Learning Outcomes
QEP at Southern Miss

- Finding a Voice: Improving Oral and Written Competencies
  - Strengthen faculty pedagogies
  - Enhance the learning environment
  - Improve assessment
Faculty Development Seminar

- Intensive, ten-week, graduate-style seminar
- Provide repertoire of strategies
- 24 faculty members participate annually
- $1500 stipend or reassigned time
- Produce portfolio including revised syllabi and speaking and writing assignments
Assessing Faculty Development – Nascent Stages

- Participants:
  - Gave feedback after completing the seminar (open-ended questions)
  - Completed a survey the semester after implementation
  - Rated first drafts of student papers with final drafts of papers and rated student presentations using a standard rubric
Benefits
- Open-ended, immediate feedback
- Built-in focus on revision process

Challenges/Problems
- Only seminar participants were involved
- Data collected every semester for indefinite time
- No pre-seminar information was collected
- Improvement from first to final draft seemed a given
Assessing Faculty Development Now – Indirect Measures

- **Seminar Feedback**
  - readings, content, format, facilitators

- **Pre and Post-Seminar Surveys**
  - use of pedagogies and confidence

- **Annual Survey of Teaching Strategies**
  - university-wide measurement
  - captures cross-fertilization
  - allows for comparisons of QEP/Non-QEP
  - allows for glimpse at long-term use of strategies by QEP alumni (post-post)
Assessing Faculty Development Now – Indirect Measures

- **Benefits**
  - Open-ended, immediate feedback
  - Better understanding of impact of seminar
  - Better measurement of pedagogies used university-wide

- **Challenges/Problems**
  - Not enough time to implement action plans promptly and fully
    - led to fall/spring offerings instead of fall/spring/summer
  - Survey fatigue
Connecting Faculty Development to Student Learning Outcomes - Direct Measures

- **First Iteration**
  - Faculty seminar participants rated papers and presentations using QEP rubric

- **Problems**
  - Lack of inter-rater reliability
  - Lack of representative sample participating
  - “Requiring” use of QEP rubric
  - Time period for requiring participation
Connecting Faculty Development to Student Learning Outcomes - Direct Measures

- **Second Iteration (Pilot)**
  - Faculty from the QEP Assessment Committee served as raters

- **Problems**
  - Lack of inter-rater reliability
  - Faculty found it problematic to rate papers from other disciplines
Focus on Student Learning Outcomes

- U. of Houston Undergraduate Writing Assessment: Spring 2006 Report
  (www.uh.edu/writecen/Resources/UndergraduateWritingAssessmentSpr06.pdf)

- Examined five factors:
  - past performance (GPA)
  - explicitness of various traits in the writing assignment
  - student attitudes and beliefs about writing
  - the age student began learning English
  - transfer status
Adapted U. of Houston model

Goals:
- Measure Student Learning Outcomes
- Compare differences between classes led by QEP faculty seminar participants and non-participants
- Examine role of other variables
- Add more dots in the mosaic
Data Collection

- **Who**
  - **All** capstone faculty and students

- **What**
  - Syllabi, assignment instructions, ungraded student papers, student recordings, student surveys

- **Where**
  - Students upload papers via Blackboard
  - Recordings are conducted in Speaking Center or submitted electronically

- **When**
  - Annually in the spring only

- **How**
  - Requirements communicated through faculty to students
Data Preparation

- Identifying a Representative Sample
  - Complete sets of data
  - ~400 papers and presentations
    - 5 colleges –
      - 80 papers and presentations from each college
    - Number/course based on proportion of enrollment in the college
Data Preparation

- Preparing artifacts
  - Remove student identifiers
  - Replace with unique artifact codes
    - Link back to survey data
    - Full anonymity not possible with presentations
  - Add course codes so that after ratings, courses can be linked back to faculty development participation (not known during rating process)
Selecting the Raters

- 4 graduate students for each area
- Experience tutoring in the centers and/or teaching writing or speaking
- Identified by the center directors
- Pay them well and “by the project”

For more on selecting the raters, see “Writing Center-housed Writing Assessment” by Dr. Melanie Barthelme (last pages of handout).
Training the Raters

› Training is critical
› Review rubrics and trait analysis and apply to artifacts
› Discuss issues unique to particular disciplines when possible
› Review FERPA and confidentiality issues
› Establish inter-rater reliability
› Discuss procedures for entering scores

For more on selecting the raters, see “Writing Center-housed Writing Assessment” by Dr. Melanie Barthelme (last pages of handout).
The Rating Process

- Establishing context
  > Each rater reads the syllabus and assignment instructions and analyzes it using the Assignment Trait Analysis.

- Measuring student outcomes
  > Each paper/presentation is rated by two raters using the QEP rubrics.

- Recording and Entering Scores
  > Electronic and paper
    • Minimizes error
    • Allows for monitoring of pace of progress
    • Allows for monitoring inter-rater drift
    • Provides for back-up
Obtain a final rating for each criterion by averaging the two scores.

Overall scores determined by averaging the category scores.

Use of third rater is recommended.
Analyzing Data

- Compare mean scores by criterion and overall for QEP faculty participants and non-QEP faculty.

- Frequencies of proficiency levels are also analyzed and compared.

- Examine other variables.
Reporting and Using Findings

- Useful for identifying areas of improvements
- Useful in seeing differences in faculty development, especially when other forms of baseline data were not established
- Shared in aggregate by overall, QEP, non-QEP with Advisory Board, GEC, and Deans
- Aggregate numbers are shared by college
- Course-level information is only given to the instructor upon request
Examples of Use of Results

- Developed “Best Practices Guide”
- Added resources specifically on documentation of sources
- Offered more faculty workshops on designing writing assignments and speaking assignments
- Held alternative delivery model of seminar to reach more faculty
- Partnerships with colleges
Limitations and Challenges

- One dot
- Other variables
- Faculty who participate are more likely to be good but want to be better
- Time and labor intensive
- Turn-around time in using results is especially difficult
Recommendations

- Use both indirect and direct measures
- Involve everyone not just a sample
- Conduct periodically
  - “We know that Clorox bleaches. We don’t have to restudy this before we do every wash” (Chism and Szabo, 1997).
- Involve experts/resources available to you
- Connect direct to indirect to see relationships
A Sunday on La Grande Jatte by Georges Seurat

http://teachers.henrico.k12.va.us/godwin/hoen_I/artinsight1/art1projects/PointalistFruit/A%20Sunday%20on%20La%20Grande%20Jatte.jpg


U. of Houston Undergraduate Writing Assessment: Spring 2006 Report (www.uh.edu/writecen/Resources/UndergraduateWritingAssessmentSpr06.pdf)