Developing CourseEvalUM: Our Process, Challenges, and New Directions

Sharon A. La Voy & Jessica Mislevy
Institutional Research, Planning and Assessment

What We’ll Cover Today

- History of the project
- Decisions made along the way
- Instrument development
- Reporting
- Administration and implementation
- Next steps
Project History

- Decentralized by college and even department
- Students interested in seeing results
- Senate action – committees, task forces, etc.
- Assigned to IRPA to design and administer
- Technical development assigned to OIT

Project History - Timeline

April 2005: Task Force submits final recommendations to Campus Senate
April 2006: Implementation Committee submits plan
December 2006: Items piloted
Summer 2007: First level of technology piloted
December 2007: All courses participate
May 2008: College level & multiple instructors added
Decisions

- Senate task forces envisioned dynamic system with hierarchy for university, college, department, and instructor items
- Lawyers required administrative vs. student views for personnel file reasons
- Comments go to administrators, and instructors see all
- Off-the-shelf products did not have such a system available, could develop for $$

Decisions (cont.)

- OIT heard of an open-source product being designed with these specifics, in Sakai
- Student participation key
  - 70% response rate necessary for course display
  - No access without participation
- Needed a coordinator
Instrument Development

- Senate task force suggested 16 universal items
- Piloted the items to see how they function
  - Partnered with units to replace with or add
    items to existing systems in Fall’06
- Conducted subsequent pilots to test new
  technology system

Qualitative Analysis of Pilot Data

- Respondents asked to comment on items
  which seemed unclear, were hard to answer,
  or did not seem to apply to them
- Most feedback was positive but some items
  were identified as problematic
  - Slight changes made to item wording and
    response options
Example of Item Changes

- Original item: “The grading in this course was fair.” (Strongly Disagree to Strongly Agree)
- Several issues raised by respondents
  - Hard to answer
  - Non-directional
- Revised item: “Based on the quality of my work in this course, the grades I earned were...” (Too Low, Appropriate, Too High)

Quantitative Analyses of Pilot Data

- Descriptive statistics:
  Respondents tend to use positive end of scale
- Dimensionality:
  Items tap single dimension
Quantitative Analyses (cont.)

Reliability:
- Responses are highly consistent across the set of items
- Student and administrator item sub-sets function the same
- Responses remain stable across administrations

Reporting of Results
- Types of reports currently available
  - Course-section reports
  - Summary reports
  - Large lecture roll-up reports
- Calculation and display of results
  - Access to results
  - Summary measures
- Methodological decisions and challenges
Course-Section Reports

Display of results depends on item type:

- Access to results - student, admin, or instructor
- Focus of the item - instructor or course item
- Response scale - interval, ordinal, or text
- Hierarchy level - university or college-level

Sample Course-Section Report

**ADMINISTRATOR UNIVERSITY-WIDE INSTRUCTOR ITEMS:**
Questions for use by faculty/instructors and for administrative purposes
N/A responses have been excluded from the following calculations.

Instructor: Professor A

<table>
<thead>
<tr>
<th>Instructor</th>
<th>FAKE100 0101 COLLEGE COMPARISON*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of Responses</td>
<td>% Strongly Disagree</td>
</tr>
<tr>
<td>The instructor treated students with respect.</td>
<td>65</td>
</tr>
<tr>
<td>The instructor was well-prepared for class.</td>
<td>65</td>
</tr>
<tr>
<td>Overall, this instructor was an effective teacher.</td>
<td>65</td>
</tr>
</tbody>
</table>

* Average rating for all similarly leveled course sections (e.g., all 200-level course sections) in this college.

**AVERAGE OF FIVE ADMINISTRATOR AGREE/DISAGREE QUESTIONS: 3.58 / 4.00**
Scaled 0-4: Strongly Disagree=0; Strongly Agree=4. N/A is not in the average.

The standards the instructor set for students were ... (Number of Responses 65)

0% Too Low 94% Appropriate 6% Too High
Summary Measures

- Instructor score
- College comparison mean
- Large lecture roll-up
- Unit summaries
  - Department, college, and university level
  - Results by sub-unit and by course level

Sample Summary Report

College-level Results by Department and by Course Level

Fake College Department-Level Results:

<table>
<thead>
<tr>
<th>Department</th>
<th>Score*</th>
<th>The instructor treated students with respect.</th>
<th>The instructor was well-prepared for class.</th>
<th>The course was intellectually challenging.</th>
<th>I learned a lot from this course.</th>
<th>Overall, this instructor was an effective teacher.</th>
<th>Number of Course Section Units Included in Calculations</th>
<th>Total Number of Evaluations Submitted</th>
<th>Total Number Enrolled</th>
<th>Overall Response Rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dept A</td>
<td>3.53</td>
<td>3.60</td>
<td>3.73</td>
<td>3.23</td>
<td>3.41</td>
<td>3.70</td>
<td>11</td>
<td>112</td>
<td>201</td>
<td>55.7%</td>
</tr>
<tr>
<td>Dept B</td>
<td>3.10</td>
<td>3.51</td>
<td>3.46</td>
<td>2.59</td>
<td>2.74</td>
<td>3.22</td>
<td>11</td>
<td>117</td>
<td>175</td>
<td>66.9%</td>
</tr>
<tr>
<td>Dept C</td>
<td>3.33</td>
<td>3.55</td>
<td>3.46</td>
<td>3.14</td>
<td>3.20</td>
<td>3.30</td>
<td>73</td>
<td>995</td>
<td>1,301</td>
<td>76.5%</td>
</tr>
<tr>
<td>Dept D</td>
<td>3.28</td>
<td>3.63</td>
<td>3.53</td>
<td>2.71</td>
<td>3.08</td>
<td>3.45</td>
<td>40</td>
<td>527</td>
<td>839</td>
<td>62.8%</td>
</tr>
</tbody>
</table>

* Average of Instructor Scores from course-section units across all course sections in department
Instructor Scores are scaled 0-4: Strongly Disagree=0; Strongly Agree=4. N/A is not in the average.
Methodological Challenges

- Deciding what to display and how to calculate it
  - Specific descriptions and formulas
  - Handling NA and missing responses
- Rolling up the data across courses
  - Unit of analysis
- Defining a course section unit
  - Large-lecture and multiple-instructor courses

Administration and Implementation

- Senate interest
  "Must be implemented next semester"
- IRPA Coordinator as campus “point-person”
- IRPA-OIT planning and development team
Administration

College Liaisons
- Communication within colleges
- On-going communication with IRPA

Schedulers
- IRPA Liaison with registration to department schedulers
- SIS indicators for Yes or No Evaluation per course and instructor

Administration (cont.)

Advisory Group
- Colleges, SGA, GSG, Academic Affairs, OIT, IRPA, Graduate School
- Policy and development recommendations

OIT Help Desk
- Responding to issues they cannot address

Student Government and other groups
Implementation

- On-going development of application
  - Sakai development vs. reporting
  - Open source challenges
- Faculty buy-in
  - Varies by college and unit
  - Individual instructors make the difference

Implementation (cont.)

- Student participation challenges
  - About 61-66% overall each semester
  - Confusion over “shadow systems”
  - 2/3 submit and 1/3 did not
- Need to understand why, and to what effect
Response Rate Qualitative Study

Best Practices for Faculty:
- Almost all - verbal reminders
- Almost half - electronic reminders
- Several - section-specific response rates, some for friendly competitions among sections
- Only some - mention on syllabus
- Theme - high importance placed on teaching

Response Rate Qualitative Study

Non-Participating Students:
- Almost all were aware of CourseEvalUM
- Most said they were too busy or ran out of time
- Others said they filled out another eval, they forgot, were too lazy, or didn’t want to participate
Response Rate Quantitative Study

Study on relationship between response rates and instructor scores:
- Less than 1% of variance in scores explained by response rate
- Less than 5% of variance in scores explained by response rate, class size, course level, and academic discipline
- Leaving 95% unexplained by known course characteristics

Continuing Development
- Competing desires for enhancements
- Eliminating “shadow systems”
- Acknowledging frustrations with iterative process of development
Data Warehousing

- Need for direct access to data
- IRPA long-term assessment/research interests
- “Local” assessment interests
- Current vs. frozen
  - Users would query current
  - IRPA would use frozen
- Methodological challenges – calculation error

See our CourseEvalUM website:
https://www.irpa.umd.edu/Assessment/crs_eval.shtml
Email: course-eval-admin@umd.edu

Fear the Turtle