Assessment for Course and Program Improvement

Charles Henderson, Western Michigan University

Physics Department Chairs Conference, June 6-8, 2014
Panelists

- **David Kuehn, Pittsburg State University**
  - Assessment at course and program level

- **Jesus Pando, DePaul University**
  - Assessment that meets both the department and university interests

- **Peter Saeta, Harvey Mudd College**
  - Assessing the efficacy of a “sidecar” support course in parallel with 1st year Mechanics
Some Assessment-Related Pressures

Shift in thinking about faculty role
- teacher-centered to student-centered

Changes in Regional Accreditation
- Outcomes rather than inputs
- Continuous improvement rather than minimal standards

Concern from Public and Policy Makers
- Need to justify value of higher education

Education Research
- New measurement techniques
Is this what assessment looks like at your institution?

“The program review process is seen as a perfunctory exercise to be performed at specific predetermined intervals to meet the requirements of an external authority or institution. The process generates reams of paper, which while satisfying the needs of the external authority, have little or not impact on the day-to-day life of the academic unit. The process, like other aspects of accreditation, is often seen by faculty as busy work, and has very little to do with the units’ academic goals or processes of continual renewal.” (p. 73)
Redirect rather than fight

Aikido is performed by blending with the motion of the attacker and redirecting the force of the attack rather than opposing it head-on.
Assessment is a Simple Idea

Assessment Questions

1. What are the major goals?
2. Have they been met?
3. How do we know (evidence)?
Assessment is part of the continuous improvement cycle
Assessment is also used to judge performance
Levels of Assessment

Diagram showing the levels of assessment:
- University
- Program
- Course
# Purposes of Assessment and Relationships

<table>
<thead>
<tr>
<th>Improve</th>
<th>Judge</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Course</strong></td>
<td><strong>Personnel Decisions:</strong></td>
</tr>
<tr>
<td>Course Improvement (Instructor, Department)</td>
<td>(Department, Institution)</td>
</tr>
<tr>
<td><strong>Program</strong></td>
<td><strong>Program Review:</strong></td>
</tr>
<tr>
<td>Program Improvement (Department)</td>
<td>Accountability and Allocation of Resources (Department, Institution)</td>
</tr>
</tbody>
</table>

=Simplify/Summarize
Course Level
Data Source

- **N=72** physics faculty
- Semi-structured telephone interviews
- Assessment-related data from throughout the interview
- Specific questions about assessment
  - How do you know if your instruction is working?
  - What criteria does your institution use to evaluate teaching?

What Assessment Sources are Currently Used?
(Faculty perceptions inferred from interviews, N=72)

- Student Evaluations of Teaching
- Peer Observations
- Teaching Portfolios
- Research-based Assessments
- Exam and HW Performance
- Systematic Formative Assessment
- Informal Formative Assessment
- Post-Course Feedback

Percentage of Faculty Reporting Use

- Used by Faculty
- Used by Institutions
Institutions and departments typically base most or all of their assessment of teaching effectiveness on the numerical ratings from SETs, a measure that many faculty are skeptical of. Nobody thinks this is a good idea. SETs could be improved with existing knowledge, (e.g., salgsite.org).
When peer observations are used, there are no predetermined criteria.

Seven guidelines for useful peer observations:
1) observers receive training
2) a single classroom observation is not sufficient
3) purpose of the observation should be clear at the course and should help the observer achieve necessary goals
4) the approach used by an observer should help to focus the observations (e.g., via a checklist or rating form)

Would be useful for departments to agree on purpose and procedures for peer observations
Not common to use available nationally-normed research-based assessments (such as the FCI).

This is the easiest course-level evidence to summarize for higher levels.
Faculty base much of their assessment of teaching effectiveness on student test performance. Institutions and departments rarely use this information.

These can be summarized for course judgment and build to program level.
## Purposes of Assessment and Relationships

<table>
<thead>
<tr>
<th>Improve</th>
<th>Judge</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Course</strong></td>
<td><strong>Program Review:</strong></td>
</tr>
<tr>
<td><strong>Course Improvement</strong> (Instructor, Department)</td>
<td><strong>Accountability and Allocation of Resources</strong> (Department, Institution)</td>
</tr>
<tr>
<td><strong>Mainly Informal Measures</strong></td>
<td><strong>Limited Measures</strong></td>
</tr>
<tr>
<td><strong>Program Improvement</strong> (Department)</td>
<td></td>
</tr>
</tbody>
</table>

Many missed opportunities to use measures that can be summarized for higher levels.
Program Level

Actual Situation Not Well Studied

- Weak measures typically used
  - Number of graduates
  - Standardized exams for physics majors
  - Capstone experience (usually assessed informally)
Meaningful Program Assessment Requires Faculty Input

Two Examples

- Wieman Course Transformation Model
- Marbach-Ad Research Group Model

Both involve faculty groups developing goals and measures.
Wieman Course Transformation Model
Start with Course Level

- Identify Course Goals (not topics)
- Develop Assessment
- Make Course Revisions
- Assess

For Upper Division E&M Course, 13 instructors met 7 times to set goals. (Supported by Science Teaching Fellow)

Developed diagnostic test (CUE).

Core Question: “What is junior E&M about? How is it different from the introductory E&M course?”

Course Level Led to Broader Program Level Goals

Broad Learning Goals for Upper-Level Physics
1. Math/Physics Connection
2. Visualization
3. Knowledge Organization
4. Communication
5. Problem-Solving Techniques
6. Problem-Solving Strategies
7. Expecting and Checking Solution
8. Intellectual Maturity

http://www.colorado.edu/sei/departments/physics_learning.htm
Marbach-Ad Research Group Model

Start with Important Topic Area

- Focus on 7 microbiology courses

Goals
- Minimize overlap, allow courses to build on one another
- Develop assessment tools
Marbach-Ad Research Group Model

12 instructors meet monthly. Supported by a graduate student.

Instructors change their courses and discuss experiences with group.

Core Question: “What do we want our students to truly understand and remember 5 years after they have completed our set of our courses?”
### Curricular Alignment

#### Table 4. Example of curricular Alignment Matrix for eight of our courses and 16 questions

<table>
<thead>
<tr>
<th>Question</th>
<th>Concept*</th>
<th>BSCI223</th>
<th>BSCI380</th>
<th>BSCI424</th>
<th>BSCI412</th>
<th>BSCI417</th>
<th>BSCI422</th>
<th>BSCI423</th>
<th>MC GM**</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>12</td>
<td>N/N (2/1)</td>
<td>No (2)</td>
<td>N/N (0/1)</td>
<td>No (0)</td>
<td>No (2)</td>
<td>No (3)</td>
<td>No (2)</td>
<td>No (3)</td>
</tr>
<tr>
<td>2</td>
<td>3, 4, 10</td>
<td>N/Y (3)</td>
<td>Yes (2)</td>
<td>N/N (2/3)</td>
<td>Yes (3)</td>
<td>Yes (3)</td>
<td>Yes (0)</td>
<td>Yes (0)</td>
<td>No (2)</td>
</tr>
<tr>
<td>3</td>
<td>2, 3</td>
<td>N/Y (2/3)</td>
<td>Yes (2)</td>
<td>N/N (2/3)</td>
<td>No (1)</td>
<td>Yes (3)</td>
<td>Yes (0)</td>
<td>Yes (0)</td>
<td>No (3)</td>
</tr>
<tr>
<td>4</td>
<td>3, 4, 10</td>
<td>N/N (3/2)</td>
<td>Yes (2)</td>
<td>Y/N (3/3)</td>
<td>No (3)</td>
<td>Yes (3)</td>
<td>Yes (0)</td>
<td>Yes (0)</td>
<td>No (3)</td>
</tr>
<tr>
<td>5</td>
<td>6</td>
<td>N/N (0/1)</td>
<td>No (2)</td>
<td>N/Y (0/2)</td>
<td>No (1)</td>
<td>Yes (3)</td>
<td>Yes (2)</td>
<td>Yes (1)</td>
<td>No (2)</td>
</tr>
<tr>
<td>6</td>
<td>13</td>
<td>N/N (2/1)</td>
<td>Yes (0)</td>
<td>?/N (0/3)</td>
<td>No (0)</td>
<td>No (2)</td>
<td>No (2)</td>
<td>No (2)</td>
<td>No (3)</td>
</tr>
<tr>
<td>7</td>
<td>10</td>
<td>N/N (3/1)</td>
<td>Yes (0)</td>
<td>Y/N (3/0)</td>
<td>No (3)</td>
<td>Yes (3)</td>
<td>Yes (0)</td>
<td>Yes (0)</td>
<td>No (3)</td>
</tr>
<tr>
<td>8</td>
<td>12</td>
<td>N/N (1/1)</td>
<td>No (2)</td>
<td>N/N (0/2)</td>
<td>No (0)</td>
<td>No (2)</td>
<td>No (3)</td>
<td>No (3)</td>
<td>No (2)</td>
</tr>
<tr>
<td>9</td>
<td>7, 1, 12</td>
<td>N/N (2/1)</td>
<td>No (2)</td>
<td>N/N (0/2)</td>
<td>No (0)</td>
<td>Yes (3)</td>
<td>No (3)</td>
<td>No (2)</td>
<td>No (3)</td>
</tr>
<tr>
<td>10</td>
<td>4, 5, 9, 12</td>
<td>N/N (3/2)</td>
<td>Yes (3)</td>
<td>?/N (2/3)</td>
<td>No (1)</td>
<td>Yes (3)</td>
<td>Yes (2)</td>
<td>Yes (0)</td>
<td>No (1)</td>
</tr>
<tr>
<td>11</td>
<td>8</td>
<td>N/N (2/1)</td>
<td>Yes (3)</td>
<td>?/N (2/3)</td>
<td>No (0)</td>
<td>Yes (3)</td>
<td>Yes (1)</td>
<td>Yes (0)</td>
<td>No (0)</td>
</tr>
<tr>
<td>12</td>
<td>3</td>
<td>N/N (2/2)</td>
<td>Yes (1)</td>
<td>?/N (2/2)</td>
<td>No (3)</td>
<td>Yes (3)</td>
<td>Yes (1)</td>
<td>Yes (0)</td>
<td>No (3)</td>
</tr>
<tr>
<td>13</td>
<td>7, 9</td>
<td>N/Y (2/2)</td>
<td>Yes (2)</td>
<td>Y/N (3/3)</td>
<td>Yes (2)</td>
<td>Yes (3)</td>
<td>Yes (0)</td>
<td>Yes (0)</td>
<td>No (3)</td>
</tr>
<tr>
<td>14</td>
<td>10</td>
<td>N/N (2/1)</td>
<td>Yes (1)</td>
<td>?/N (2/0)</td>
<td>No (1)</td>
<td>Yes (3)</td>
<td>Yes (0)</td>
<td>Yes (0)</td>
<td>No (3)</td>
</tr>
<tr>
<td>15</td>
<td>13</td>
<td>N/N (2/2)</td>
<td>Yes (0)</td>
<td>N/N (0/2)</td>
<td>No (0)</td>
<td>No (2)</td>
<td>N/Y (3)</td>
<td>No (3)</td>
<td>No (3)</td>
</tr>
<tr>
<td>16</td>
<td>9, 10</td>
<td>N/N (2/1)</td>
<td>Yes (1)</td>
<td>N/N (3/2)</td>
<td>No (3)</td>
<td>Yes (3)</td>
<td>Yes (0)</td>
<td>Yes (0)</td>
<td>No (0)</td>
</tr>
</tbody>
</table>

For each question, instructors reported: 1) Their assumptions about student prior knowledge (Yes, No, or (?) for don’t know); and 2) The level of topic coverage in their classes (0 = not at all; 1 = briefly; 2 = moderately; 3 = detailed). Two numbers or letters in one box indicates feedback from two instructors.

*The final version of the concept inventory includes additional question that covers concept 11.

**MC GM, Montgomery College General Microbiology course.

---

Assess both Course and Program Level

Scores from each course

<table>
<thead>
<tr>
<th>Course</th>
<th>Pre</th>
<th>Post</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>General Microbiology - BSCL 223/Fall 2006</td>
<td>31.1 ± 15.6</td>
<td>48.1 ± 16.9</td>
<td>&lt;0.001***</td>
</tr>
<tr>
<td>General Microbiology - BSCL 223/Spring 2007</td>
<td>31.9 ± 18</td>
<td>44.2 ± 19.4</td>
<td>&lt;0.001***</td>
</tr>
<tr>
<td>General Microbiology - BSCL 223/Fall 2008</td>
<td>26.1 ± 15.6</td>
<td>47.1 ± 16.2</td>
<td>&lt;0.001***</td>
</tr>
<tr>
<td>General Microbiology - BSCL 223/Spring 2009</td>
<td>30.1 ± 14.8</td>
<td>49.1 ± 14.1</td>
<td>&lt;0.001***</td>
</tr>
<tr>
<td>Bioinformatics - BSCL 380/Fall 2008</td>
<td>46.3 ± 18.2</td>
<td>49.1 ± 14.9</td>
<td>0.252</td>
</tr>
<tr>
<td>Pathogenic Microbiology - BSCL 424/Fall 2006</td>
<td>43.9 ± 16.9</td>
<td>51.1 ± 18.1</td>
<td>&lt;0.001***</td>
</tr>
<tr>
<td>Pathogenic Microbiology - BSCL 424/Spring 2008</td>
<td>45.2 ± 15.0</td>
<td>51.0 ± 13.6</td>
<td>&lt;0.001***</td>
</tr>
<tr>
<td>Epidemiology - BSCL 425/Spring 2007</td>
<td>44.1 ± 20.6</td>
<td>42.8 ± 22.5</td>
<td>0.606</td>
</tr>
<tr>
<td>Microbial Genetics - BSCL 412/Spring 2007</td>
<td>51.4 ± 16.25</td>
<td>49.0 ± 20.8</td>
<td>0.266</td>
</tr>
<tr>
<td>Bioinformatics - BSCL 412/Fall 2008</td>
<td>53.8 ± 17.4</td>
<td>52.8 ± 16.6</td>
<td>0.72</td>
</tr>
<tr>
<td>Microbial Genetics - BSCL 412/Spring 2009</td>
<td>54.0 ± 14.2</td>
<td>56.2 ± 18.1</td>
<td>0.407</td>
</tr>
<tr>
<td>Microbial Pathogenesis - BSCL 417/Spring 2008</td>
<td>58.9 ± 20.4</td>
<td>65.0 ± 18.3</td>
<td>0.24</td>
</tr>
<tr>
<td>Microbial Pathogenesis - BSCL 417/Spring 2009</td>
<td>51.0 ± 15.4</td>
<td>52.5 ± 14.3</td>
<td>&lt;0.05*</td>
</tr>
<tr>
<td>Immunology Lecture - BSCL 422/Spring 2007</td>
<td>59.9 ± 19.4</td>
<td>64.3 ± 21.9</td>
<td>&lt;0.05*</td>
</tr>
<tr>
<td>Immunology Lecture - BSCL 422/Fall 2008</td>
<td>53.7 ± 16.9</td>
<td>56.6 ± 17.8</td>
<td>0.132</td>
</tr>
<tr>
<td>Immunology Lecture - BSCL 422/Spring 2009</td>
<td>58.6 ± 15.7</td>
<td>61.6 ± 16.4</td>
<td>0.572</td>
</tr>
</tbody>
</table>

Note: The number of students in this table reflects students who responded to the pre- and post-survey and gave their permission to use this data for research purposes. *p < .05, **p < .001.

Figure 1. Concept Inventory scores according to the number of HPI course taken by the students. Based upon the Concept Inventory.

Common Features

- Focus on Broad Learning Goals, then specific measures
  - Framed by meaningful questions:
    - What is junior E&M I about? How is it different from the introductory E&M course?
    - What do we want our students to truly understand and remember 5 years after they have completed our set of our courses?
- Involved both course and program level goals
- Faculty Ownership and Direction
- Regular meetings (but not too intensive)
- Support (Post doc or grad student)
In These Examples

<table>
<thead>
<tr>
<th>Improve</th>
<th>Judge</th>
</tr>
</thead>
<tbody>
<tr>
<td>Course</td>
<td>Program Review: Accountability and Allocation of Resources (Department, Institution)</td>
</tr>
<tr>
<td>- Assessment data (Interaction w/ Colleagues)</td>
<td>- Summaries of % of students meeting goals.</td>
</tr>
<tr>
<td>Program</td>
<td></td>
</tr>
<tr>
<td>- Course data from different program stages</td>
<td>- Summaries of student performance (over time and instructor)</td>
</tr>
</tbody>
</table>
Thank You

Wieman Course Transformation Model


Marbach-Ad Research Group Model

- Marbach-Ad, G., Briken, V., Frauwirth, K., Gao, L.-Y., Hutcheson, S. W., Joseph, S. W., … Smith, A. C. (2007). A faculty team works to create content linkages among various courses to increase meaningful learning of targeted concepts of microbiology. *CBE Life Sciences Education, 6*(2), 155–62. doi:10.1187/cbe.06-12-0212