Tracy Scott and Brandy Simula
Emory University

Building Assignment-Based Strategies for Assessing Transdisciplinary Learning Outcomes
Emory University Quality Enhancement Plan: 
*The Nature of Evidence: How Do You Know?*

- **QEP Goal**: to enhance learning about the nature of evidence
  - Focused on the first year undergraduate experience

- **Curricular Component**: First Year Seminars (FSEM)
  - Develop designated Evidence-Focused First Year Seminars (E-FSEM)
  - Aim is to make teaching about evidence explicit
  - Evidence-Focused Criteria pertain to the focus & structure of the course
    - Applicable across disciplines and content areas (to date, 37 departments have participated)
Evidence Learning Outcomes

1. Distinguish uses of evidence in a discipline and/or between disciplines.
2. Identify, select, and gather evidence.
3. Evaluate and analyze evidence.
4. Build arguments based on evidence and assess the arguments of others.

Evidence seminars must address 3 of the 4 Learning Outcomes
A Perspective on Assessment

• Regardless of theoretical orientation, assessment is a research process and should follow the best research methodology.

• Our goal: use best social science research methods for the QEP assessment.

  **Research Question:** *Does the QEP enhance student learning about evidence?*

  **Outcome** (dependent variable): *Student learning about evidence*

  **Independent variable:** *QEP Curricular Element: Evidence-Focused First Year Seminars*

• Two Research Designs (Assessment Strategies) using different operationalizations of the Outcome: “Student Learning about Evidence”

  **2015-2017:** Standardized Test of Critical Thinking/Evidence Skills

  **2017-2020:** Evidence Assignments from First Year Seminar Courses
E-FSEM Assessment: 2015-2017

• Research Design: 2015-2016 and 2016-2017 academic years
  • Experimental (E-FSEM) vs Control (non-evidence FSEM)
  • Pre-and-Post-Test

• Assessment Measure
  • Quantitative, Standardized Test: Watson-Glaser II

• E-FSEMs (67 courses; all of the E-FSEMs)
  • Courses included humanities, social sciences, natural sciences

• Non-evidence FSEMs (44 courses; self-selected)
  • Courses included humanities, social sciences, natural sciences

• Results: No statistically significant differences pre-post for either group, either year.
Change in Mean WG-II Scores (Pre-Post)

- Experimental: 26.17 (Pre), 25.87 (Post)
- Control: 25.61 (Pre), 25.31 (Post)

N=1,301
Other Analyses: 2015-2017

2015-2016: analyzed WG-II by SAT/ACT scores
• Statistically significant correlation between SAT/ACT scores and WG-II scores (at both pre and post assessment, $p < 0.01$ for both)

2016-2017: analyzed WG-II by self-reported effort & motivation
• Statistically significant correlation between changes in effort and motivation and changes in WG-II scores
  • Students aren’t motivated to do tests that “don’t count for anything”
  • Better to assess something associated with course learning and evaluation
E-FSEM Assessment: 2015-2017
Lessons Learned

• **WG-II** did not capture any differences in student ability resulting from evidence first year seminars
  • Not entirely surprising: “critical thinking ability” (which WG-II measures) is related to but not identical to QEP learning outcomes
  • As with SAT/ACT scores, very difficult to change or move “up”

• **WG-II** scores were related to Motivation and Effort
  • Students aren’t motivated to do tests that don’t “count for anything”
  • Would be better to evaluate something associated with course learning and evaluation (grades)
E-FSEM Assessment: 2017-2020

• Research Design: Experimental (E-FSEM) vs Control (non-evidence FSEM)

• Assessment Measure
  • Student coursework: **1 course assignment completed by all students**
  • Qualitative, Standard Rubrics (based on AAC&U’s VALUE Rubrics)
  • (NOTE: not all assignments will address all learning outcomes)

• Procedure
  • Adapted VALUE Rubrics for QEP Evidence Learning Outcomes
  • Trained team of (7) graduate student scorers, led by Post-Doctoral Fellow
  • Tested for Inter-coder reliability

• 2017-2018 Assessment of Assignment Sets from:
  • **26 E-FSEM**s (humanities, social science, and natural science courses)
  • **10 control FSEM**s (humanities, social science, and natural science courses)
VALUE Rubrics (AAC&U): Adapted for QEP Learning Outcomes

• QEP Assessment Committee created modified rubrics drawing on relevant VALUE rubrics:
  • Critical Thinking
  • Informational Literacy
  • Inquiry and Analysis

• QEP Rubric 1: Identify and evaluate evidence
  • Assesses our QEP Learning Outcomes 2 and 3

• QEP Rubric 2: Build and assess arguments using evidence
  • Assesses our QEP Learning Outcome 4
<table>
<thead>
<tr>
<th>A. Access the Needed Information</th>
<th>Milestone 1</th>
<th>Milestone 2</th>
<th>Benchmark 3</th>
<th>Not present in assignment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Accesses information using effective, well-designed search strategies and most appropriate information sources. Demonstrates ability to refine search.</td>
<td>Accesses information using variety of search strategies and some relevant information sources.</td>
<td>Accesses information using simple search strategies, retrieves information from limited and similar sources.</td>
<td>Accesses information randomly, retrieves information that lacks relevance and quality.</td>
<td>Not present in assignment</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>B. Evaluate Information and Its Sources Critically</th>
<th>Milestone 1</th>
<th>Milestone 2</th>
<th>Benchmark 3</th>
<th>Not present in assignment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chooses a variety of information sources appropriate to the scope and discipline of the research question. Selects sources after considering the importance (to the researched topic) of the multiple criteria used (such as relevance to the research question, currency, authority, audience, and bias or point of view).</td>
<td>Chooses a variety of information sources. Selects sources using multiple criteria (such as relevance to the research question, currency, and authority).</td>
<td>Chooses a variety of information sources. Selects sources using basic criteria (such as relevance to the research question and currency).</td>
<td>Chooses few information sources. Selects sources using limited criteria (such as relevance to the research question).</td>
<td>Not present in assignment</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>C. Evidence (Selecting and using information to investigate a point of view or conclusion)</th>
<th>Milestone 1</th>
<th>Milestone 2</th>
<th>Benchmark 3</th>
<th>Not present in assignment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Information is taken from source(s) with enough interpretation/evaluation to develop a comprehensive analysis or synthesis. Viewpoints of experts are questioned thoroughly.</td>
<td>Information is taken from source(s) with some interpretation/evaluation, but not enough to develop a coherent analysis or synthesis. Viewpoints of experts are subject to questioning.</td>
<td>Information is taken from source(s) without any interpretation/evaluation. Viewpoints of experts are taken as fact, without question.</td>
<td>Information is taken from source(s) without any interpretation/evaluation. Viewpoints of experts are taken as fact, without question.</td>
<td>Not present in assignment</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>D. Limitations and Implications</th>
<th>Milestone 1</th>
<th>Milestone 2</th>
<th>Benchmark 3</th>
<th>Not present in assignment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Insightfully discusses in detail relevant and supported limitations and implications.</td>
<td>Discusses relevant and supported limitations and implications.</td>
<td>Presents relevant and supported limitations and implications.</td>
<td>Presents limitations and implications, but they are possibly irrelevant and unsupported.</td>
<td>Not present in assignment</td>
</tr>
</tbody>
</table>
Evidence Courses: Rubric Criteria Scores (in percentages)*
*(due to rounding error, percentages do not all sum to 100)

Rubric 1 Performance by Evidence Courses
(n= 344 assignments)

- Rubric 1A: 94% (Not Present: 4%, Benchmark 1: 2%, Milestone 2: 20%)
- Rubric 1B: 72% (Benchmark 1: 13%, Milestone 2: 13%, Milestone 3: 2%)
- Rubric 1C: 3% (Not Present: 9%, Benchmark 1: 28%, Milestone 2: 60%, Milestone 3: 9%)
- Rubric 1D: 55% (Not Present: 6%, Benchmark 1: 33%, Milestone 2: 6%)
<table>
<thead>
<tr>
<th>QEP RUBRIC for LO4: Learning Outcome 4: Build arguments based on evidence and assess the arguments of others</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Capstone 4</strong></td>
</tr>
<tr>
<td><strong>A. Existing Knowledge, Research, and/or Views</strong></td>
</tr>
<tr>
<td>Synthesizes in depth information from relevant sources representing various points of view/approaches.</td>
</tr>
<tr>
<td><strong>B. Determine the Extent of Information Needed</strong></td>
</tr>
<tr>
<td>Effectively defines the scope of the research question or thesis comprehensively. Can determine key concepts. Types of information (sources) selected directly relate to concepts or answer research question.</td>
</tr>
<tr>
<td><strong>C. Evidence (Selecting and using information to Investigate a point of view or conclusion)</strong></td>
</tr>
<tr>
<td>Information is taken from source(s) with enough interpretation/evaluation to develop a coherent analysis or synthesis. Viewpoints of experts are questioned thoroughly.</td>
</tr>
<tr>
<td><strong>D. Explanation of issues</strong></td>
</tr>
<tr>
<td>Issue/problem to be considered critically is stated clearly and described comprehensively, delivering all relevant information necessary for full understanding.</td>
</tr>
<tr>
<td><strong>E. Influence of Context and Assumptions</strong></td>
</tr>
<tr>
<td>Thoroughly (systematically and methodically) analyzes own and others' assumptions and carefully evaluates the relevance of contexts when presenting a position.</td>
</tr>
<tr>
<td><strong>F. Analysis</strong></td>
</tr>
<tr>
<td>Organizes and synthesizes evidence to reveal insightful patterns, differences, or similarities related to focus.</td>
</tr>
<tr>
<td><strong>G. Conclusions and Related Outcomes (Implications and consequences)</strong></td>
</tr>
<tr>
<td>Conclusions and related outcomes (consequences and implications) are logical and reflect student's informed evaluation and ability to place evidence and perspectives discussed in prior order.</td>
</tr>
</tbody>
</table>
Qualitative, Thematic Analysis

• Complements quantitative overview
• Enables identification of specific areas of strengths and weaknesses on rubrics criteria
• For instance:
  • On limitations criterion:
    • Strength: Most student work demonstrated an understanding that a single study does not “prove” an argument and additional work is needed
    • Weakness: Students struggled to articulate specific areas for future research beyond a general argument for “more” research
• Allows us to provide more course- and assignment- design useful information to faculty
E-FSEM Assessment: 2017-2020
Lessons Learned, so far...

• Using student coursework and transdisciplinary rubrics is an effective method of assessing evidence-specific abilities
• Rubrics applicable to all of the courses we assessed (across disciplines)
• Graduate students have excelled as scorers, with high inter-coder reliability
• Assignments not designed to rubrics nonetheless capture many of the evidence criteria on the rubrics
  • Underscoring point that faculty do not need to design assignments to assessment instruments for this method of assessing student learning
Conclusions

• Drawing on data from multiple years of implementation we discussed the challenges and successes of two different strategies for assessing transdisciplinary learning outcomes across first-year courses representing more than two dozen disciplines.

• We used a conceptual, research-informed process for selecting each of the assessment strategies utilized, as well as practical aspects of their implementation.

• Our results show that assignment-based assessment strategies are particularly effective for assessing students’ best, most highly-motivated efforts and for assessing specific learning outcomes that may not be adequately captured via instruments such as standardized tests of general skill sets.
Questions?

Contact:

• Tracy L. Scott, Director of the Quality Enhancement Plan, Emory University: tscott@emory.edu

• Brandy Simula, Post-Doctoral Fellow for the Quality Enhancement Plan, Emory University: bsimula@emory.edu

Website: www.evidence.emory.edu