Beyond Anecdote: Challenges in Bringing Rigor to Service-Learning Research

Sherril Gelmon, Andrew Furco, Barbara Holland and Robert Bringle
Session Overview

- Background and context
- Criteria for evaluating research and methods
- Consideration of several large-scale, role model projects
- Recommendations for enhanced research
- Discussion
Role of Research

- Building intellectual foundations
- Integrating theory and practice
- Reporting to funders and partners
- Building organizational support
- Persuading others to adopt service-learning
- Documenting practices/strategies
- Program improvement
Clarifying Terms

- **Research** – theoretic frame, scientific design, control for causality, robust analysis, validity/reliability, generalizes
- **Program Evaluation** – funder-driven, biased sample, narrow focus on specific program and anticipated outcomes, lack of controls for causality
- **Evaluation Research** – coherent program design and outcomes, procedures allow for causal inferences, clear implications beyond idiosyncratic program that was evaluated
Types of Assessment
(Bringle & Hatcher, MJCSL, 2000)

◆ Reflection: Activities producing information oriented toward the self-assessment of persons who are engaged in an experience

◆ Process Evaluation or Monitoring: Activities producing information about how a class, course, or program was implemented

◆ Outcome Evaluation: Activities producing information about what outcomes occurred as a result of a class, course, or program
Types of Assessment (continued)

- Correlational: Activities producing information about what relationship exists between aspects of a class, course, or program

- Experimental Research: Activities producing information about why a specific outcome occurred
Research Limitations

- Lack of common definition for service-learning and related research
- Variation in programmatic practices and purposes
- Studies conducted as self-studies by advocates of service-learning
- Studies mostly commissioned by funders with narrow, specific questions
More Research Limitations

- Few experimental studies
- Limited number of longitudinal studies
- Small sample sizes
- Many studies based on participant self-report
- Data collection often dictated by reporting requirements and expectations
Implications

- Limited generalizability or predictive value of most studies
- Weak causal connections
- Predisposing factors unknown
- Results subject to alternate explanations
- Skeptical reaction from scholars and policymakers
- Limited evidence for building support
Criteria and Methods to Assess Research

- Framework developed to assess articles
- Five core topics in critical analysis:
  - Background
  - Literature review/theoretical background
  - Methods and research design
  - Results, findings, and discussion
  - Interpretations and conclusions
### What Counts As “Evidence”?

<table>
<thead>
<tr>
<th></th>
<th>Evidence</th>
<th>Possible Evidence</th>
<th>No Evidence</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Comparison</strong></td>
<td>2 or more groups compared (treatment and control group)</td>
<td>2 or more groups compared treatment and control group or matched comparison groups</td>
<td>Only one group is studied (Lack of control or comparison group)</td>
</tr>
<tr>
<td><strong>Random Assignment</strong></td>
<td>Subjects are assigned randomly to groups</td>
<td>Subjects are assigned randomly to groups or randomization is approximated.</td>
<td>Lack of randomization</td>
</tr>
<tr>
<td><strong>Consistent Treatment</strong></td>
<td>Treatment is implemented uniformly across the group</td>
<td>Minimal variation in treatment across members of the treatment group</td>
<td>Gross variation in treatment across members of the treatment group</td>
</tr>
</tbody>
</table>
What Counts As “Evidence”?

<table>
<thead>
<tr>
<th>Multiple Sites</th>
<th>Evidence</th>
<th>Possible Evidence</th>
<th>No Evidence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Treatment is offered at more than one “site”</td>
<td>Treatment is offered at one or more sites</td>
<td>Treatment is limited to one site or is site specific</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Replication</th>
<th>Evidence</th>
<th>Possible Evidence</th>
<th>No Evidence</th>
</tr>
</thead>
<tbody>
<tr>
<td>The study is replicated (multiple cases) using the same procedures</td>
<td>The study is replicated but not in the exact same manner</td>
<td>The study is not replicated; one set of data are used to draw conclusions</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Controls</th>
<th>Evidence</th>
<th>Possible Evidence</th>
<th>No Evidence</th>
</tr>
</thead>
<tbody>
<tr>
<td>External influences are controlled for in both treatment and comparison groups</td>
<td>Most or all external influences are controlled for in both treatment and comparison groups</td>
<td>There are few or no controls for influences that might contaminate the findings</td>
<td></td>
</tr>
</tbody>
</table>

Principles of High Quality Research

Research that:

◆ Poses significant questions
◆ Is linked with relevant theory
◆ Uses methods that permit direct investigation of the question
◆ Provides coherent, explicit chain of reasoning
◆ Is replicable and generalizable
◆ Is scrutinized and critiqued by qualified professionals
<table>
<thead>
<tr>
<th>PARADIGM</th>
<th>DESIGN</th>
<th>TYPE</th>
<th>FORM</th>
</tr>
</thead>
<tbody>
<tr>
<td>Positivistic</td>
<td>Experimental</td>
<td>Basic</td>
<td>Quantitative</td>
</tr>
<tr>
<td>Interpretive</td>
<td>Non-experimental</td>
<td>Applied</td>
<td>Qualitative</td>
</tr>
<tr>
<td>Critical</td>
<td>Quasi-experimental</td>
<td>Action</td>
<td>Mixed Methods</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Evaluation</td>
<td></td>
</tr>
</tbody>
</table>
Convincing Research

- Guided by theory
- Clear constructs
- Control for differences among groups
- Multiple indicators
- Multiple methods
- Converging results across different methods
- Confidence in conclusions
- Implications for teaching and learning in general
Is Eyler & Giles Convincing Research?

- Guided by theory
- Clear constructs
- Control for differences among groups
- Multiple indicators
- Multiple methods
- Converging results
- Confidence in conclusions
- Implications for teaching and learning in general
Consideration of Recent Projects

- If we could do these over, what would it take to make them stronger as examples of good service-learning research?
- Health Professions School in Service to the Nation
- Engaging Campuses and Communities (CAPHE)
- California Campus Compact Study on Partner Perceptions
- Indiana Campus Compact Retention Project
Health Professions Schools in Service to the Nation

- National demonstration project on service-learning in health professions education
- Three years (1995-1998)
- 17 schools/programs completed project
- Goal: To identify role model practices in integrating SL into health professions
- Research: Large scale evaluation with individual campus activities and common national approaches
If We Could Redesign Research: HPSISN

- Greater control over monitoring and testing various strategies
- Use of common strategies, data collection tools, analyses across sites
- More sites for stronger comparisons
- Rigorous comparisons with non-SL sites
- More research funding to enable this work
Engaging Campuses and Communities

- Consortium for the Advancement of Private Higher Education (CAPHE)
- 13 campuses involved in institutionalizing civic engagement through experiential learning
- Service-learning as common strategy
- Three year process
- Research focus: Process of developing effective practice and institutionalization
- Goal: Identify best practices
If We Could Redesign Research: CAPHE

- Balance the “etic” (outsider) and the “emic” (insider)
- Balance the deductive with the inductive
- Clarify terms and develop a common language
- Constantly connect data to the context
- Triangulate data from multiple data sources
California Campus Compact
Study of Partner Perceptions

Question: How do SL community partners characterize their collaboration with academic institutions? Motivations, expectations, barriers, facilitators

Focus groups in 8 communities; diverse types of institutional partners

Common protocol, three facilitators, same notetaker/recorder for all

Emphasis on control/consistency/objectivity in data collection; pure community voice
Define eligible participants more specifically; eliminate any campus connections (alumni, retiree, adjunct)

Use one facilitator for all; style differences

Use software for data analysis – save time!!

Organize partners in cohorts (school, youth dev, health, social service, etc.)

Follow with a survey?
Objective: Is the presence or absence of a service-learning class during freshman students’ first semester related to re-enrollment at the same institution the following year?

- Multiple first-year service learning and non-service learning classes at 8 institutions
- Pre-test survey to statistically control for pre-existing differences
- Measured differences among SL classes (e.g., reflection, service)
Non-service learning classes were not always close comparables

Did not know what happened in students’ second semester (e.g., More service learning?)

Did not ask about other forms of student engagement that might have contributed to next-year retention

Scale of intervention (i.e., single course) might not have been sufficient for the retention measure (i.e., enrollment next year)
Use of Mixed Methods

- Mixed methods are not necessarily better, but they can help when they:
  - Are selected based upon theory and constructs
  - Provide complementary types of information
  - Produce converging results
  - Are all based on multiple indicators
Use of Multiple Indicators

◆ Multiple indicators are superior (quantitative and qualitative) whether focusing on:
  ■ Items on a scale
  ■ Time samples
  ■ Journal entries
  ■ Courses or campuses
  ■ Skills
  ■ Intentions

◆ See Bringle, Phillips, & Hudson, 2004 for a collection of scales
Clarity of Constructs

- Applies to both quantitative and qualitative strategies
- Review literature for existing theory and research
- Conduct focus groups and structured interviews
- Establish good theoretical statement(s)
Domain Sampling

- Need clear definition of the domain
- Need to obtain a representative sample
- Need to evaluate the sample of the domain
Recommendations for Enhanced Research

- Strong theoretical base
- Multiple comparable units with large samples
- Adequate funding to support approach
- More analytic, less descriptive
- More rigor, less idiosyncratic
- Expert researchers leading work
- Program leaders supportive of research
- Replicate high quality studies
Recommendations for Enhanced Research

- Enhance support of graduate students and their research
- Engage researchers from other disciplines to study service-learning as an intervention
- Support junior faculty to develop expertise in this domain of research
- Create faculty development opportunities to enhance research capabilities
- Look for allies -- such as institutional research
• Questions about perspectives?

• Questions about specific research projects?

• Thoughts on recommendations?
For Further Information

◆ Sherril Gelmon, Portland State University, gelmons@pdx.edu
◆ Andy Furco, University of California at Berkeley, afurco@berkeley.edu
◆ Barbara Holland, National Service-learning Clearinghouse, barbarah@etr.org
◆ Bob Bringle, Indiana University Purdue University Indianapolis, rbringle@iupui.edu