Evaluation Design: Topics to Consider

NASEM: Action Collaborative on Preventing Sexual Harassment in Higher Education – 2019 Public Summit

Ann Kearns Davoren, PhD
NORC at the University of Chicago

November 20, 2019
Guiding Evaluation Questions

There is no one size fits all evaluation framework.

- What is the intervention?
- What is the goal of the intervention?
- What is the objective of the evaluation? (In other words…Why am I collecting data?)
  - What evaluation methods are available to me?
  - What data is available / collectable?
  - Do I have a counterfactual?
What is the intervention? & What is the goal of the intervention?

- Logic model and model of change identify the intervention and the goal(s) of the intervention.

Resources / Inputs → Activities → Outputs → Outcomes → Impact

- Multi-prong / multi-topic programs
  - How do you define the intervention?
  - What are the goals of those interventions?
  - How do you evaluate the results?
### What is the objective of the evaluation?

- **Two primary evaluation types:**

<table>
<thead>
<tr>
<th></th>
<th>Formative Evaluation</th>
<th>Summative Evaluation</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Objective</strong></td>
<td>Understand &amp; Improve</td>
<td>Estimate Impact</td>
</tr>
<tr>
<td><strong>RQs</strong></td>
<td>Who, what, where, when, why, how</td>
<td>Changes, effects, impacts</td>
</tr>
<tr>
<td><strong>Methods</strong></td>
<td>Case study</td>
<td>Randomized control trial (RCT)</td>
</tr>
<tr>
<td></td>
<td>Grounded Theory</td>
<td>Quasi-experimental (QED)</td>
</tr>
<tr>
<td></td>
<td>Ethnography… (etc.)</td>
<td>Non-Experimental</td>
</tr>
<tr>
<td><strong>Data</strong></td>
<td>Qualitative</td>
<td>Quantitative</td>
</tr>
</tbody>
</table>

- Process evaluation (implementation assessment) essential to both
- Mixing the methods and purpose of the evaluation is a common and useful practice
Formative evaluations are designed to answer questions about how a program works, usually with the intent to improve.

<table>
<thead>
<tr>
<th>Method</th>
<th>Data</th>
<th>Description / Consideration</th>
</tr>
</thead>
<tbody>
<tr>
<td>Qualitative</td>
<td>Interviews</td>
<td>Semi-structured conversations one-on-one</td>
</tr>
<tr>
<td>(e.g., ethnography, grounded theory, case study, etc.)</td>
<td>Focus groups</td>
<td>Semi-structured conversations in a group setting. Might consider needing to separate genders, students/faculty/admin.</td>
</tr>
<tr>
<td>Observation</td>
<td>Observation</td>
<td>Observe intervention in natural habitat (can be particularly useful to assess implementation)</td>
</tr>
</tbody>
</table>

**Pros**
- Can do a deep dive into the subject / context
- Better understand the “why”

**Cons**
- Not representative of your population
- Sensitive subject to discuss

**Quantitative data (surveys, admin) is a useful addition here.**
What evaluation methods are available to me?

- **Summative evaluations** - program's effectiveness

<table>
<thead>
<tr>
<th>Method</th>
<th>Data</th>
<th>Description / Consideration</th>
</tr>
</thead>
<tbody>
<tr>
<td>Randomized control trial (RCT)</td>
<td>- Random assignment to treatment / control &lt;br&gt;- Gold standard in intervention evaluation &lt;br&gt;- Strong counterfactual &lt;br&gt;- Resource heavy &lt;br&gt;- Ethical considerations</td>
<td>Peer matching - Match on institution or within institution &lt;br&gt;- Provides counterfactual &lt;br&gt;- Can be paired with another method like ITS &lt;br&gt;- Within institution matching will be susceptible to diffusion effects</td>
</tr>
<tr>
<td>Quasi-experimental (QED)</td>
<td>- Surveys &lt;br&gt;- Admin Data</td>
<td>Interrupted Time Series - Baseline data + post-intervention data &lt;br&gt;- Uses full population (no assignment needed) &lt;br&gt;- Difficult to isolate effect of intervention without a comparison group (like peer matching)</td>
</tr>
</tbody>
</table>
What data is available / collectable?

Key questions to consider when choosing evaluation measures:

1. What secondary data sources are available for the evaluation?
2. Are the secondary data sources sufficient or is primary data collection necessary to answer the research questions?
3. Are there resources and time to collect primary data?
4. What is an appropriate timeframe for tracking performance?
5. Are there standardized or validated measures available to address research questions?
6. If using secondary data, is there a standard way of calculating outcome measures from the raw data (e.g., 12 month time frame for victimization)?
Resources

- Robert Wood Johnson Foundation Logic Model Guide
- W.K. Kellogg Foundation Logic Model Development Guide
- W.K. Kellogg Foundation Guide to Evaluation