Module 6: Descriptive, Normative, and Impact Evaluation Designs
Introduction

- What Is Evaluation Design?
- Connecting Questions to Design
- Design Elements
- Types of Designs for Impact Evaluation
- Key Points about Design
Evaluation Design

• The total process of specifying a plan for:
  - collecting data
  - analyzing data
  - reporting results
  - getting the results used
Design Process

Questions ➔ Designs ➔ Methods ➔ Analysis ➔ Reporting
Approach to Development Evaluation

Focus the Evaluation
- Purpose
- Terms of Reference
- Program logic model
- Program outcome model
- Specification of evaluation questions
- Identification of stakeholders

Use Evaluation
- Communicate Findings
- Feed-back
- Decision-making
- Action Plan

Design & Methodology
- Evaluation questions
- Data collection design
- Measurement strategy
- Sampling strategy
- Data Collection strategy
- Develop data collection instruments
- Involve stakeholders

Gather & Analyze Data
- Gather data according to protocols
- Prepare data for analysis
- Analyze data

Report Findings
- Interpret the data
- Write report
- Make recommendations

Use Evaluation
- Communicate Findings
- Feed-back
- Decision-making
- Action Plan
Design Matrix

• Another organizing tool to help plan an evaluation
• Organizes questions and the plans for collecting information to answer questions
Matrix Elements

- Design Matrix Planning Instrument for:
- Major Issues Being Addressed
- Major Assumptions Being Made
- Questions
- Sub-questions
- Type of Question
- Design
- Measures or Indicators
- Criteria for Normative Questions
<table>
<thead>
<tr>
<th>Questions</th>
<th>Sub-Questions</th>
<th>Type of Question</th>
<th>Design</th>
<th>Measures or Indicators</th>
<th>Criteria for Normative Questions</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Matrix Elements (page 2)

- Data Sources
- Sample
- Data Collection Instrument
- Data Analysis
- Comments
<table>
<thead>
<tr>
<th>Data Sources</th>
<th>Sample Collection Instrument</th>
<th>Data Analysis</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Answering Descriptive Questions

- Descriptive questions generally use non-experimental designs.
- Common designs for descriptive questions:
  - one-shot
  - cross-sectional
  - before-and-after
  - time series
  - longitudinal
  - case studies
One-shot Designs

• A look at a group receiving an intervention at one point in time, following the intervention
• Use to answer questions such as:
  - How many women were trained?
  - How many participants received job counseling as well as vocational training?
  - How did you like the training?
  - How did you find out about the training?
One-shot

- Represented as: 
  \[- X \quad O_1\]
- There is one group receiving the treatment “X” and one observation “O”
- There is no before treatment/intervention measure
Cross-sectional Designs

- Also show a snapshot at one point in time
- Also interested in sub-group responses
- Often used with survey method
- Subgroups may be:
  - age
  - gender
  - income
  - education
  - ethnicity
  - amount of intervention received
Cross-sectional

- Evaluation question may focus on
  - participant satisfaction of services
  - why they did not use services
  - find out current status of people from an intervention a few years ago

- Evaluation questions might be:
  - Do participants with different levels of education have different views on the value of training?
  - Did women receive different training services than their male counterparts?
Cross-sectional

• Represented as:
  - X \[O_1\]
  \[O_2\]
  \[O_3\]

• The observation is made after the intervention “X” and responses of subgroups (“O\(_1\), O\(_2\), O\(_3\)” and so on) receiving the interventions are examined
Before-and-after Designs

• Also called pre- and post-designs
• Ask about group characteristics
• There is no comparison group
• Evaluation questions:
  - Did program participants increase their knowledge of parenting techniques?
  - What was the change in wages earned, two years after the training intervention?
Before-and-after Designs

- Represented as: $O_1 \times O_2$
- Observation, intervention, observation
Time Series Designs

• Look for change over time
• Purpose is to explore and describe changes over time - either after, or before and after the intervention
• Can be used to discern trends
• Often there are existing data that can be used
Time Series

• Evaluation questions:
  - What are the trends in child mortality rates before and after and over time for an intervention?
  - What are the changes in participant attitudes over time towards women entrepreneurs?
Time Series

- Represented as:
  \[ O_1 \ O_2 \ O_3 \ X \ O_4 \ O_5 \ O_6 \]
- Several (three shown above) observations are made prior to the intervention and again three more times after the intervention
Longitudinal Study

- A type of time series design
- Repeated measures of the same variable are taken from the study population
- Panel design is one type of longitudinal study where a small group of people is tracked at multiple points over time
  - almost always use qualitative questions (open-ended survey questions, in-depth interviews, and observation)
  - can give a more in-depth perspective
Longitudinal

- Evaluation question:
  - How did the allocation of social benefits effect families’ transition into and out of poverty?
  - a study looking at Poland’s family allowance from 1993 to 1996
Case Study Design

• Descriptive case study
• In-depth information is collected over time to better understand the particular case or cases
• Useful for describing what implementation of the intervention looked like - and why things happened the way they did
• May be used to examine program extremes, or a typical intervention
Case Study

• Represented as:
  \( O_1 \)
  \( O_2 \)
  \( O_3 \)
Answering Normative Questions

• Similar to descriptive questions
• Normative always assessed against a criterion:
  - a specified desired or mandatory goal, target, or standard to be reached
• Generally the same designs work for normative questions as descriptive questions
Answering Cause-Effect Questions

- Pose the greatest challenge
- Need a well thought out design
- Design attempts to rule out feasible explanations other than the intervention
- Internal validity: a design’s ability to rule out other explanations
Common Threats to Internal Validity

- History (events occurring at the same time)
- Maturation of subjects (getting older changes the results)
- Testing (learning how to take the test)
- Instrumentation (changes in data collection instruments or procedures)
- Selection bias (participants may be different to begin with)
- Attrition (a specific group of people may drop out)
Impact Designs

• Can use experimental and quasi-experimental designs

• Experimental sometimes called the “medical model”
  - randomly assign participants to a group, group does not know who is in the treatment or placebo group (“blind studies”)
Controlling

• To reduce the possibility of believing we know something as true which is really not, need to control everything but the intervention, including:
  - the implementation of an intervention
  - who receives it
  - the environment in which it is delivered
Impact and Multi-site and Cluster Evaluations

• Each site and the nature of the interventions may vary in different locations
• Complexity may limit options for design
Design Elements for Impact Questions

- For evaluators doing *traditional* experimental evaluation:
  - before-and-after measures
  - comparison groups
  - random assignment to the comparison groups
- For *newer* approaches (i.e. cluster, multi-site, and rapid assessment)
  - use of control variables
  - use of natural variation
  - causal tracing strategies
Before-and-After Measures

• Change is measured by comparing key measures after the intervention began against the measures taken before the intervention began
• Before measure might be called the baseline
• Collecting baseline data might be called a baseline study
• Change alone does not prove causality
Comparison Groups

- Treatment group: group that received treatment
- Control group: group that does not receive treatment
- If the intervention causes change those in treatment group show more change than the control group
- Again, alternative explanations *must* be ruled out before drawing conclusions
Random Assignment

• Random: people, or things are placed in groups by chance
• Random assignment makes groups comparable
• Not always an option but it is possible more often than you think
  - when not all participants can receive the intervention at once
Use of Control Variables

• Random assignment impossible?
  - Rule out alternative explanations by statistically controlling for them:
    • prior performance or prevalence levels
    • socioeconomic status
    • prior soil quality
    • weather / climate
Use of Natural Variation

• Inconsistent implementation? Turn it into an advantage

• Useful evidence includes:
  - less extensive implementation
    • smaller (or no) impact
  - better quality implementation
    • more positive results and/or fewer negative impacts
Causal Tracing Strategies

- Based on the general principles used in traditional experimental and quasi-experimental designs, but:
  - can be used for rapid assessments
  - can be used without high-level statistical expertise
  - can be used on small scale interventions where numbers preclude statistical analysis
  - can be used for evaluations with a qualitative component
  - involves the evaluator doing some detective work
Causal Tracing Strategies

• Ask yourself:
  - What decisions are likely to be based on the evidence from this evaluation?
  - How certain do I need to be about my conclusions?
  - What information can I feasibly collect?
  - What combination of information will give me the certainty I need?

• Remember: this list is a menu of possible sources of evidence, not a strict checklist of requirements
9 Causal Tracing Evidence Sources

- Causal list inference
- Modus operandi
- Temporal precedence
- Constant conjunction

- Contiguity of influence
- Strength of association
- Biological gradient
- Coherence
- Analogy
Types

- Experimental design
- Quasi-experimental design
- Correlational design
- Case study design
- Non-experimental design
Experimental Design

• Called the “true experiment”
  - involves random assignment
  - uses comparison groups
  - often includes before-and-after measures
• Considered the optimum approach but can be difficult to implement
• Drawback:
  - often small scale, less generalizable
Randomized Control Trials (RCTs)

- Consider when:
  - you have a discrete, concrete intervention - singular, well-defined
  - implementation can be standardized
  - valid and reliable measures exist for the outcome to be tested
  - random assignment is possible
  - random assignment is ethical
Randomized Control Trials (RCTs)

• NOT appropriate for:
  - complex, multi-dimensional and highly context-specific community interventions
  - ethical constraints

• NOT needed if:
  - face validity is high
  - observed changes are dramatic
  - link between treatment and outcome is direct
Quasi-experimental Design

• The design is similar to true experimental design but:
  - no random assignment
  - uses naturally-occurring comparison groups
  - requires more data to rule out alternative explanations
Examples of Quasi-experimental Design

• Before-and-after: good for descriptive questions
• Matched and non-equivalent comparison design
• Time series and interrupted time series design
• Correlational design using statistical controls
• Longitudinal design
• Panel design
Correlational Design

- Often used when seeking to answer questions about relationships and associations
- Often used with already available data
Case Study Design

- Used when the researcher wants to gain an in-depth understanding of a process, event, or situation
- Good to learn how something works or why something happens
- Are often more practical than a national study
- Can consist of a single case or multiple cases
- Can use qualitative or quantitative methods to collect data
Key Points about Design

• There is no perfect design
• Each design has strengths and weaknesses
• There are always trade-offs - time, costs, practicality
• Acknowledge trade-offs and potential weaknesses
• Provide some assessment of their likely impact on your results and conclusions
To continue on to the Next Module click here

To return to the Table of Contents click here