Module 7a

Introduction to Evaluation Design:
Planning and Managing Performance Evaluations
Module Objectives

By the end of this module you will be able to:

1. Identify three major types of evaluation designs along a continuum.

2. Identify the kinds of questions that performance evaluations address.

3. Identify 7 main designs for performance evaluations, including their strengths and weaknesses.

4. Select an optimal performance evaluation design for a given context.
## ADS References for this Module

<table>
<thead>
<tr>
<th>ADS Reference</th>
<th>Topic or Issue</th>
</tr>
</thead>
<tbody>
<tr>
<td>ADS 203.6 and 203.3.1.3</td>
<td>Definitions and use of performance evaluations</td>
</tr>
<tr>
<td>ADS 203.3.1</td>
<td>Learning and accountability purposes of evaluation can be achieved simultaneously</td>
</tr>
<tr>
<td>ADS 203.3.4</td>
<td>Integrating PE and IE into design of projects</td>
</tr>
<tr>
<td>ADS 203.3.11</td>
<td>Evaluations will be unbiased in measurement and reporting and most will be external</td>
</tr>
<tr>
<td>ADS 203.3.14</td>
<td>Operating units required to conduct at least one PE for each large program that it implements. Large = exceeds average project size for that operating unit</td>
</tr>
<tr>
<td>ADS 203.3.1.6</td>
<td>Evaluations will use the best methods possible to generate the highest quality and most credible evidence about the questions being asked</td>
</tr>
<tr>
<td>ADS 203.3.1.2.; 3.1.4</td>
<td>Planning and managing performance evaluations</td>
</tr>
</tbody>
</table>
Definitions

• What is Evaluation?

• What is a Performance Evaluation?

• What is a design?
The Right Design is Key to Planning a Successful PE

• When is an evaluation design the “right design”?
• When it is directly linked to the evaluation questions
• When its selection is based on sound information regarding the what’s, why’s and how’s of the intervention to evaluate
• When it takes validity into account
# The Timing of Different Evaluations

<table>
<thead>
<tr>
<th>Evaluation Type</th>
<th>Timing</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Impact Evaluation</strong></td>
<td>Produced at or after project completion and used to plan or scale-up future projects.</td>
</tr>
<tr>
<td></td>
<td>• IEs should be designed during the earliest stages of program design.</td>
</tr>
<tr>
<td></td>
<td>• Some IEs can be designed later in the program though many more technical challenges are involved.</td>
</tr>
<tr>
<td><strong>Performance Evaluation</strong></td>
<td>A range of evaluation options that can be produced at any phase of the program cycle to improve performance.</td>
</tr>
<tr>
<td></td>
<td>• All projects will consider the design of PEs during the project design stage. This is part of the preparation of a performance management plan.</td>
</tr>
<tr>
<td></td>
<td>• All PEs should be commissioned well in advance of the time frame for program/strategy decisions.</td>
</tr>
</tbody>
</table>
Focus on causality 

Performance Evaluation Designs

Non experimental designs

Designs without comparison groups or randomized assignment

Example: Performance Evaluations that focus on implementation processes and efficiency (descriptive and normative or comparative questions)

Impact Evaluation Designs

Weaker QEDs

Designs with comparison groups – but not randomized assignment

Example: Performance Evaluation that marginally focus on results, by answering questions such as “Did the program achieve its objectives?”

The strength of this design depends on the technical rigor of the technique used to identify the comparison group

Stronger QEDs

Experimental designs

Designs with randomized assignment (inclusion of a control group) to definitively answer cause-effect questions

Focus on implementation issues

QED = quasi-experimental design

Increasing Statistical Rigor
When to design a Performance Evaluation?

- At any phase of the program cycle as long as it will help to improve the performance of the project/program to evaluate
- Often during the project design stage (e.g., as part of the M & E Plan)
- In any case, well in advance of the program/strategy decisions informed by the evaluation findings

C-12: Common Tools and Methods used in Performance Evaluations
What Questions Can PE Address?

Questions on program relevance to the needs of the target population
✓ What was the situation of the target population before the project began?
✓ To what extent does the program respond to their identified needs?

Question on program implementation mechanisms
✓ Is the program operating according to the program design?
✓ How effectively and efficiently is it being implemented?
✓ To what extent was the project management structure effective?
✓ Is the program reaching/benefiting all sectors of the target population?
✓ What is the program actually doing and delivering, and how does this compare with the program design?
✓ What has the program delivered (at different milestones including end-of-project) and how does this compare with the program design?
Questions on program participatory and consultative processes?
✓ To what extent is the program being implemented in a participatory way?
✓ Are all appropriate stakeholders being properly consulted?
✓ Are all stakeholders involved according to the program design?
✓ How satisfied are different stakeholders with the program?
✓ What role did gender play in the implementation and results of the program? Did the project design take into consideration gender issues such as gender division of labor, time use, and control of resources?

Questions on lessons learned
✓ How could implementation be improved?
✓ What lessons can be identified that apply to future programs?
How Many PE Designs Exist?

?
Snapshot Design

- Looks at a group receiving an intervention at **one point in time** during or after the intervention
- Answers descriptive questions
- Example: examining how beneficiaries view the quality of the assistance they received
- Cost: Low

<table>
<thead>
<tr>
<th>Advantages</th>
<th>Disadvantages</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ease and practicality</td>
<td>Relatively weak design</td>
</tr>
<tr>
<td>Rapid and low cost</td>
<td>Many threats to validity</td>
</tr>
<tr>
<td>Best when combined with performance data</td>
<td>Weak ability to generalize findings to other subgroups</td>
</tr>
</tbody>
</table>
Simple Cross-Sectional Design

- Shows snapshot at one point in time across different subgroups
- Answers descriptive questions, often used with surveys or interviewing
- Example: comparison between subgroups on variables such as services received
- Cost: Low-medium

<table>
<thead>
<tr>
<th>Advantages</th>
<th>Disadvantages</th>
</tr>
</thead>
<tbody>
<tr>
<td>Clear picture at one point in time</td>
<td>No clear picture of what is happening over time</td>
</tr>
<tr>
<td>Better at discerning effects on subgroups</td>
<td>Little ability to generalize beyond immediate subgroups</td>
</tr>
</tbody>
</table>
Before and After Design (w/o comparison group)

- Aggregate measure of outcomes for a group **before and after the intervention**
- Answers descriptive and some normative (comparative) questions
- Example: baseline and end-of-project comparison of small farmer crop production in a particular region or district.
- Cost: Low-medium

<table>
<thead>
<tr>
<th>Advantages</th>
<th>Disadvantages</th>
</tr>
</thead>
<tbody>
<tr>
<td>Practical, context must be considered</td>
<td>Testing and instrumentation threats to validity</td>
</tr>
<tr>
<td>Flexible and can apply to a wide range of program types</td>
<td>Requires quality baseline info</td>
</tr>
</tbody>
</table>
Time Series

- Multiple measures of the outcome variable **for groups** both before and after the interventions
- Answers descriptive questions and normative questions
- Example: changes over time in HIV prevalence at the district or country level
- Cost: Low-high (depending on availability of secondary data)

<table>
<thead>
<tr>
<th>Advantages</th>
<th>Disadvantages</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rich perspective of change over time</td>
<td>Threat of testing bias</td>
</tr>
<tr>
<td>Maturation and history partially controlled</td>
<td>Costly, time consuming</td>
</tr>
<tr>
<td>Good when existing secondary data is available</td>
<td>Adequate secondary data often not available</td>
</tr>
</tbody>
</table>
Panel Design

- Repeated measurements from the same individuals over time
- Answers descriptive and some normative (comparative) questions
- Example: follow-up on graduates of a USAID training program – to learn what knowledge was retained; how it has been applied; and what long term effects it had on individuals and institutions.
- Cost: Low-medium

<table>
<thead>
<tr>
<th>Advantages</th>
<th>Disadvantages</th>
</tr>
</thead>
<tbody>
<tr>
<td>Practical, context must be considered</td>
<td>Testing and instrumentation threats to validity</td>
</tr>
<tr>
<td>In-depth info about KAPB</td>
<td>Requires quality baseline info</td>
</tr>
</tbody>
</table>
Case Study Design

- Used to gain in-depth understanding of a process event or situation and explain why results occurred – from the perspective of a case, a cultural context, the lived experiences of recipients.
- Answers descriptive questions
- Can be used to explore project extremes or typical cases
- Example: Case study to explore early and late adopters of new agricultural techniques.
- Cost: Medium

<table>
<thead>
<tr>
<th>Advantages</th>
<th>Disadvantages</th>
</tr>
</thead>
<tbody>
<tr>
<td>In-depth contextual information</td>
<td>Time consuming</td>
</tr>
<tr>
<td>In-depth understanding to aid survey design</td>
<td>Little external validity</td>
</tr>
<tr>
<td>and/or to interpret survey findings</td>
<td></td>
</tr>
<tr>
<td>Rich perspective on dynamics of change</td>
<td>Limited ability to generalize beyond cases</td>
</tr>
</tbody>
</table>
Developmental Evaluation Design

- Supports innovation and adaptation. Feedback is timely and rapid.
- Answers descriptive and sometimes normative questions
- Evaluator is a part of the team, working with implementers.
- Can be used in complex or pilot programs, where the intervention isn’t perfectly defined yet. Feedback is fed back into the project immediately.
- Cost: High

<table>
<thead>
<tr>
<th>Advantages</th>
<th>Disadvantages</th>
</tr>
</thead>
<tbody>
<tr>
<td>Highly adaptive and flexible</td>
<td>Time consuming</td>
</tr>
<tr>
<td>Can be used in volatile, emergent, and complex situations</td>
<td>Difficult to manage contractually, high cost</td>
</tr>
<tr>
<td>Rich perspective on dynamics of change</td>
<td>Limited external validity; perceptions of credibility</td>
</tr>
</tbody>
</table>
What Kind of Design is it? (Exercise 7-1a)

- Before and after
- Time series
- Snap shot
- Panel
- Case study
- Cross sectional

Which design?
Task 3: Case Study – Developing the Evaluation Design

1. In your case study teams develop a performance evaluation design that responds to your key normative and/or descriptive evaluation questions
   - Different designs may apply to different questions
   - Pay attention to threats to validity
   - Use your Evaluation Design Matrix

2. Identify some practical evaluation methods for your design
   - Use your Evaluation Design Matrix
   - Be prepared to present your evaluation design and methods to the large group
Review Questions

• What are the three main types of evaluation designs? What are their respective key features?

• What are the main differences between a performance evaluation and a rigorous impact evaluation?

• How many PE designs do we have and what are the main differences among them in terms of use and cost?