

Module 4: Evaluation Designs



© 2007. The World Bank Group. All rights reserved.



Learning Objectives

At the end of this module, participants will have increased their knowledge of:

- the importance of design in answering impact questions
- different design options
- the strengths and weaknesses of each design
- how to select appropriate designs for a given situation



2

Europe and Central Asia Region and World Bank Institute Evaluation Group





Evaluation Design - Definition

Evaluation Design:

- the technical part of the evaluation plan, the clarification of the links between evaluation questions, arrangements for data collection, analysis, and how evaluative judgments will be made
- represents the overall strategy for how you will systematically collect and analyze data to answer your research question
- is most critical for cause-and-effect questions due to the need for attribution

3

Europe and Central Asia Region and World Bank Institute Evaluation Group



Attribution

A key question for all evaluations:

- How do we know if the observed changes in the project participants or communities

Income, health, attitudes, school attendance, etc.

- are due to the project or intervention

An intervention seeks to change attitudes, beliefs, or behaviors: new teaching materials, a marketing campaign to prevent HIV/AIDS, a project to increase agricultural production

- or to other unrelated factors?

Changes in the economy, demographic movements, other development projects, etc.

4

Europe and Central Asia Region and World Bank Institute Evaluation Group



Challenges of Impact Questions in Evaluation

Attribution Issue

- To Demonstrate Impact
- To Eliminate Other Possible Explanations
- To Establish Internal Validity

**Internal Validity:
Establish “attribution”
Eliminates other
plausible explanations**

5

Europe and Central Asia Region and World Bank Institute Evaluation Group



Internal Validity

Can we be certain that the project or intervention produced the outcomes or consequences we observe?

If we are certain that these outcomes or consequences are due **only** to the intervention, then we have achieved internal validity.

If we are uncertain or there may be other possible explanations for the outcomes, then we have not achieved internal validity and **we cannot say with certainty that the intervention produced the outcomes that we observe.**

6

Europe and Central Asia Region and World Bank Institute Evaluation Group



Threats to Internal Validity

- **History** Event that took place during the project or intervention.
- **Maturation** Life cycle change (e.g., skills increase because people get older).
- **Testing** Risk is that people “learn” how to test.
- **Instrumentation** Changes in data collection, instrument, procedures, or measures in pre/post or comparative designs.

7

Europe and Central Asia Region and World Bank Institute Evaluation Group



Threats to Internal Validity -Cont'd

- **Regression to the Mean** Extreme scores on one measure of test are likely to be less extreme on the next measure of test.
- **Selection** The groups from which data are collected may be different in ways that effect the results.
- **Attrition** Different rates of dropping from groups may affect results.

These threats are possible. However, it does not mean they actually exist. You must consider the probability that they do.

8

Europe and Central Asia Region and World Bank Institute Evaluation Group



Counterfactual

The goal of a good impact evaluation design is to define a logically sound counterfactual that can test for and eliminate alternative explanations.

This is achieved by identifying a group (*called a control or comparison group*) that can match as closely as possible the characteristics of the project group.

- If the changes observed in the project group do not occur in the comparison group we can have greater confidence that the project or intervention contributed to the changes.

9

Europe and Central Asia Region and World Bank Institute Evaluation Group



Discussion

A government provided farmers with new irrigation canals. After three years, the government declared the project to be a success because there was a 20 percent increase in yields per hectare in the areas with access to the irrigation canals.

What is the counterfactual argument?



10

Europe and Central Asia Region and World Bank Institute Evaluation Group



Elements of a Strong Design

Can you

- control the **implementation of the project or intervention?**
- obtain measures **before and after** the intervention was implemented?
- obtain measures **between groups** that received the intervention and those that did not?
- **control** who receives the intervention?
- employ **random assignment?**

11

Europe and Central Asia Region and World Bank Institute Evaluation Group



Types of Design

- Experimental
- Quasi-Experimental
- Non-Experimental

12

Europe and Central Asia Region and World Bank Institute Evaluation Group



Summary of Design Characteristics

		Features		
		Random assignment	Before/After	Comparison group
T y p e	Experimental	√	√	√
	Quasi-experimental		√	√
	Non-experimental		?	?

13

Europe and Central Asia Region and World Bank Institute Evaluation Group



Experimental Design

Key features:

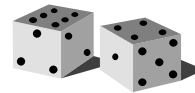
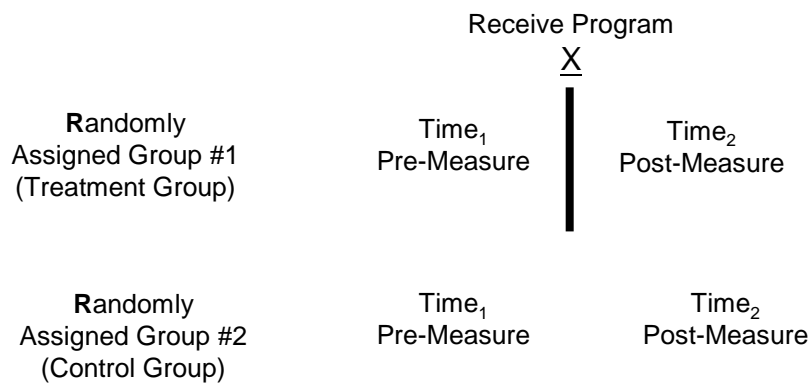
- Random assignment
- Before and After Measurement
- Comparison (with and without the program or intervention)

14

Europe and Central Asia Region and World Bank Institute Evaluation Group



Experimental Design



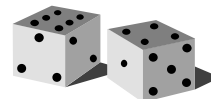
15

Europe and Central Asia Region and World Bank Institute Evaluation Group



Experimental Design

- Experimental design is the strongest design to establish cause-effect relationships:
 - It rules out most other plausible explanations for the results obtained.
 - It controls the treatment.
 - Random assignment assures the comparability of the two groups.
- If it is so great, why don't we use it all the time?
 - Ethics
 - Practicality
 - Often small scale



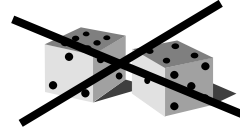
16

Europe and Central Asia Region and World Bank Institute Evaluation Group



Quasi-Experimental Design

Quasi means no random assignment!



Key features:

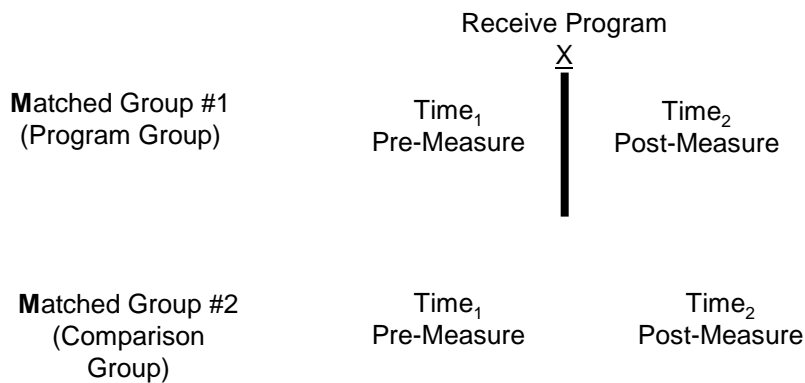
- Comparison (with and without the program)
- Might include before-and-after measures

17

Europe and Central Asia Region and World Bank Institute Evaluation Group



Quasi-Experimental Design



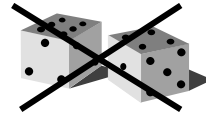
18

Europe and Central Asia Region and World Bank Institute Evaluation Group



Quasi-Experimental Designs

- Use when you cannot use random assignment to decide who gets the treatment.
- Weak because there may be selection bias.
- But this is often more practical in public sector research.



19

Europe and Central Asia Region and World Bank Institute Evaluation Group



Non-Experimental Designs

Key features:

- No random assignment
- Maybe no before program measures
- Maybe no comparison

20

Europe and Central Asia Region and World Bank Institute Evaluation Group



Summary: No Perfect Design

Quasi-experimental designs

- Useful in looking at differences
- Useful in giving context for measuring change.
- Threats: testing, instrumentation, regression to the mean, attrition, history, and maturation **may** be threats.
- Plausibility

23

Europe and Central Asia Region and World Bank Institute Evaluation Group



Summary: No Perfect Design

Non-experimental designs:

- useful for descriptive and normative questions
- very weak for cause/effect questions: many threats
- multiple design applications begin to build a set of cases

24

Europe and Central Asia Region and World Bank Institute Evaluation Group



Evaluation Plan

Case Projects

Which design would be most appropriate for each question or sub-question?

- Review the questions and sub-questions listed in the Evaluation Plan
- Complete the “design” column by entering what you believe is the best design for answering each question.
- Be prepared to justify your selection!

25

Europe and Central Asia Region and World Bank Institute Evaluation Group



Evaluation Plan

General Questions	Specific Sub-Questions	Type of Question		Indicators & Measures	Data Sources	Data Collection & Sampling	Data Analysis

26

Europe and Central Asia Region and World Bank Institute Evaluation Group



Exercise for your project

- Begin to develop a design for each of the impact questions selected for your project.

- Be sure to indicate if it is an experimental, quasi-experimental or a non-experimental design.

27

Europe and Central Asia Region and World Bank Institute Evaluation Group