

What is program evaluation?

A beginners guide

Produced by

Gene Shackman, Ph.D.
The Global Social Change Research Project
<http://gsociology.icaap.org/>
Free Resources for Methods in Program Evaluation
<http://gsociology.icaap.org/methods>

What is evaluation

This is the sixth in a set of handouts on the key program evaluation questions. This is about whether programs caused outcomes.

Did the program have an effect?

How do you know whether the program improves people's lives?

One commonly used way to find out is to ask whether the program **caused an outcome**. If the program caused the outcome, then one could argue that the program improved people's lives.

On the other hand, if the program **did not cause** the outcome, then one would argue that, since the program did not cause the outcome then the program did not improve people's lives.

How to figure this out?

One way to figure out whether the program caused the outcome is to use **comparisons**:

- Compare people on the program to people not on the program.
- First, **randomly assign** people to either be in the program (the 'treatment' group) or to not be in the program (the 'comparison' group).
- Measure the treatment group after they have been on the program and compare them to people in the comparison group.
- If the 'treatment' group people are better off than are the comparison group people, then it is reasonable to argue that the program caused that outcome.

What is evaluation

Comparisons and cause:

The idea is this:

- people were randomly assigned to be in the program or not in the program.
- Since they were randomly assigned, then the two groups of people should be pretty much the same except that one group was in the program and the other group wasn't.
- Therefore, after one of the groups has been in the program, if there are any differences between the two groups of people, then the difference must be from being in the program.
- That is, the program caused the outcome.

Additional Resources:

Why do social experiments? Chapter 7 in The Magenta Book, from Policy Hub, National School of Government.
http://www.nationalschool.gov.uk/policyhub/evaluating_policy/

What is evaluation

Advantages and disadvantages of random assignment to treatment and comparison groups.

Advantages:

- Results provide clearest demonstration of whether a program causes an outcome.
- Provides results that are easiest to explain.

Disadvantages:

- Can tell whether a program caused outcome, but doesn't give much in depth information about **why** or **how**.
- If people drop out of the program or evaluation unequally between treatment and comparison group, the results of the study will be biased, or not representative of all clients.
- When randomly assigning people to treatment or comparison groups, the people in the comparison group are denied program participation, at least temporarily, and so may be denied treatment known to be beneficial.

Summary of points from:

Why do social experiments? Chapter 7 in The Magenta Book, from Policy Hub, National School of Government.
http://www.nationalschool.gov.uk/policyhub/evaluating_policy/

In addition, randomly assigning people to be in the program is not how programs really work, so results of the evaluation may not apply to the program as it really exists.

What is evaluation

Did the program have an effect: Alternative methods.

Random assignment is often not workable, so there are other ways to do evaluation. These methods include

- non-random assignment,
- use information from multiple sources
- a very clear model of the program.

All of these methods face strong difficulties in addressing whether a program *caused* an outcome, but are useful in making reasonable arguments.

What is evaluation

Non-random assignment In general, in this method, called quasi-experimental design, people “are not randomly assigned to groups but statistical controls are used instead.”

Quasi-experimental designs. In Statnotes: Topics in Multivariate Analysis, by G. David Garson <http://www2.chass.ncsu.edu/garson/pa765/design.htm#quasi>

There are several versions of this approach:

- Comparing people already on the program to those who are not on the program. One example is to observe (O) people before they join the program or there is an intervention (X), then observe both groups after :

Pretest-posttest design

-Intervention group	O _{before}	X	O _{after}
-Comparison group	O _{before}		O _{after}

- Measuring the client many times before they join the program (or before a new intervention) and many times afterward, then compare before to after. One example is:

Time series design

-Intervention group	O ₁	O ₂	X	O ₃	O ₄
---------------------	----------------	----------------	---	----------------	----------------

- Combination of the two above

Time series design

-Intervention group	O ₁	O ₂	X	O ₃	O ₄
-Control group	O ₁	O ₂		O ₃	O ₄

What is evaluation

A major challenge to this approach is that people **on** the program may start off being very different from the people **not on** the program.

For example, some people choose to be on the program while others choose not to be on the program. This choice alone makes these two groups different.

One way to deal with this is to collect as much information as possible on characteristics of the people and program that relate to the program outcome (what the program is supposed to do), and use this information in statistical analysis to “control” for the differences between people on the program vs people not on the program.

The problem is that there may be differences, some critical, that are not observed, and for which the evaluator has no data.

Additional Resources:

AllPsych On Line. By Dr. Christopher L. Heffner
Section 5.3 Quasi-Experimental Design
<http://allpsych.com/researchmethods/quasiexperimentaldesign.html>

Quasi-experimental designs. In Statnotes: Topics in Multivariate Analysis, by G. David Garson
<http://www2.chass.ncsu.edu/garson/pa765/design.htm#quasi>

Diagrams on previous page from:
Measuring the Difference: Guide to Planning and Evaluating Health Information Outreach. Stage 4, Planning Evaluation. National Network of Libraries of Medicine
<http://nnlm.gov/evaluation/guide/>

What is evaluation

Collecting information from multiple sources could support the idea that the program causes the outcome if different sources agree.

For example, collect information from:

- Program participants
- Program staff
- Community members
- Subject experts
- Published research and reports

Collect data through many methods, for example:

- Surveys
- Interviews
- Observations
- Program data

If data from different sources don't agree, it doesn't **necessarily** mean the results from any of the sources are not valid. However, the more agreement there is from different sources, the more confident you can be about your conclusions.

Additional Resources:

An Introduction to Mixed Method Research. By Jennifer Byrne and Áine Humble. Atlantic Research Centre for Family-Work Issues. Mount Saint Vincent University.
<http://www.msvu.ca/ARCFamilyWork/publications.asp>

What is evaluation

A very clear model of the program can show how you expect the program to lead to the outcome.

If the model is very clear, and shows how the program operates to produce the outcome, then the evaluation can 'test' the model.

If research then shows that the data collected supports the model, then an argument can be made that the model is true, and the program fits the model.

Conclusion:

If the evaluation can get random assignment, or quasi experimental design, or information from multiple sources and methods that all agree, or a clear program model that can be tested, then it may be possible to conclude that the program caused the outcome.

If the above methods cannot be met, then evaluations can not say the program caused the outcome, but can describe what happened.

For example, the evaluation could say, "After the program, people were better off." But this doesn't mean it was the program that made the people better off. It could have been something else, unrelated to the program.

Thus, if there is a need to know whether the program caused the outcome, there needs to be random assignment, or quasi experimental design, or information from multiple sources and methods that all agree, or a clear program model that can be tested.

What is evaluation

Copyright by Gene Shackman, July 2008.

Some usual legal disclaimers:

This handout can be freely distributed without need for permission, provided it is distributed as is. Distribution for any commercial purpose is strictly forbidden. This handout cannot be sold under any circumstances.

This handout is only for education purposes. It does not represent any guidelines, recommendations or requirements about how to do program evaluation. The only purpose is to provide the general public, consumers, students, and evaluators with information about things that may go into evaluations, so that evaluation may be better understood, and evaluators and clients might work better together to get more out of their evaluation.

In my work on this handout, I do not represent or speak for any organization. I prepared this on my own time, at home, and was not supported by any organization.

Materials on web sites listed in this handout do not necessarily reflect my opinions, nor do I assume any responsibility for the content provided at these web sites. This handout only lists web sites with legal content. Listing a website is not necessarily endorsement of any services or organization. The sites are only listed because they have some freely available information. I also do not have any financial relationships with any site or organization listed on this handout.

I also benefited greatly from feedback from folks on various email lists, and I thank them all!

The most recent version of this handout is always at
<http://gsociology.icaap.org/methods/>