Safe Schools Healthy Students
National Center Brief

Evaluation: Designs and Approaches

July 2004

The choice of a design for an outcome evaluation is often influenced by the need to compromise between cost and certainty. Generally, the more certain you want to be about your program's outcomes and impact, the more costly the evaluation. It is part of an evaluator's job to help you make an informed decision about your evaluation design.

This publication describes two of the basic choices that must be made when designing an evaluation. The first is about the choice of an evaluation's design. The second concerns choosing a quantitative, qualitative, or mixed approach.

Issues to Consider When Selecting a Design

There are a number of important issues to consider before selecting an evaluation design. These include the following:

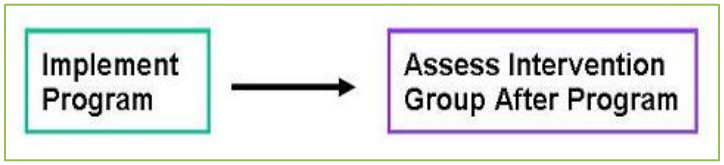
- Complex evaluations cost more, but allow for greater confidence in their findings.
- Complex evaluation designs are more difficult to implement and thus require greater expertise in research methods and analysis.
- Complex designs are not without their own problems. Increasing the complexity of an evaluation can also increase the chances that something will go wrong – especially if there is not enough evaluation expertise involved in the effort.
- No evaluation design is immune to threats to its validity. There is a long list of possible complications associated with any evaluation study. However, your evaluator will help you maximize the quality of your evaluation study.
- Don't assume that an intervention that works in one setting will always work in others. It's unlikely that any intervention will work equally well with all types of people and in all settings.

Some evaluation is better than none. Though you may not have the money or resources to conduct the evaluation of your dreams, start somewhere—even if that means using the least rigorous design.

Four Commonly Used Designs for Outcome Evaluations

There are four commonly used designs for outcome evaluations. They are described below starting with the least expensive design – which also provides the least certain results – and ends with the most costly – but also most rigorous – design.

One-Group, Post-Test Only Design (least expensive, least rigorous)

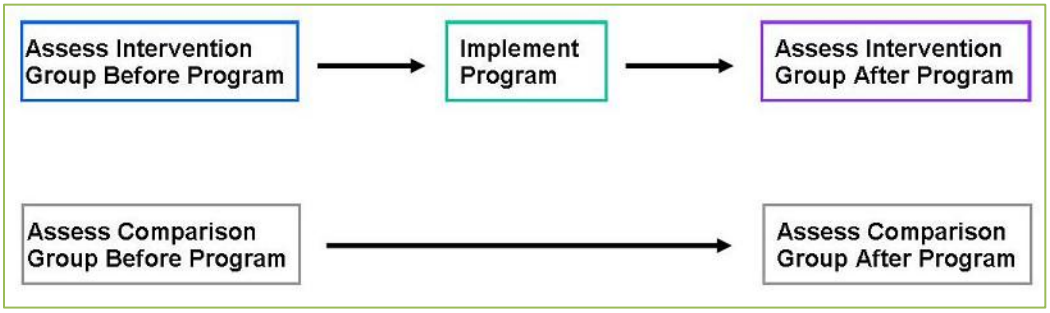


In the one-group, post-only Design, a post-test, such as survey, is administered to the program or study participants after the intervention has been administered. Although relatively inexpensive, this design does not measure changes in knowledge, attitudes, or behavior. Nor does it allow comparisons between people taking part in the program to people who did not participate.

One-Group, Pre-Test and Post-Test Design

The one group, pre-test and post-test design is more informative than the pre-test only design because it provides information on changes in knowledge, attitudes, or behavior of the program participants or study subjects that occurred during the time in which the intervention took place. All things being equal, this design can provide some evidence that the intervention produced these changes. However, it cannot conclusively demonstrate this. The changes may have occurred because of other reasons. For example, the changes may be a consequence of the fact that participants were older and more mature during the pre-test. Or they may have been exposed to another intervention (such as a national public education campaign) during the period in which they participated in the program.

Pre-Test and Post-Test with Comparison Group Design

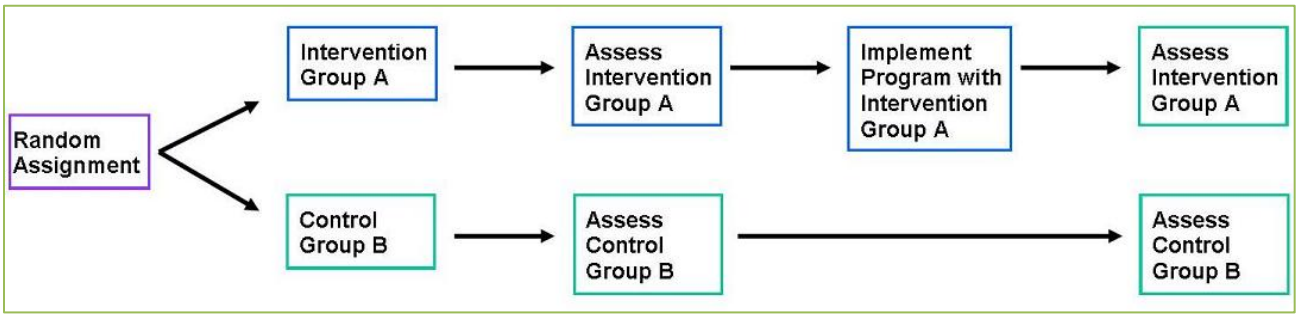


In evaluations using this design, the pre-tests and post-tests are administered to both the intervention group – that is, the people participating in the program – and another similar group that does not participate (and, consequently, does not receive the intervention). The addition of a comparison group helps determine whether any changes in knowledge, attitudes, or behavior can be attributed to the intervention.

The more similarity there is between the people who receive the intervention and those who do not, the more confidence there is that changes detected in the intervention group, but not the control group, were actually a result of the program or intervention. Thus, the comparison group should be as similar to the intervention group in terms of gender, race or ethnicity, socioeconomic status, and education as possible.

This design is both more expensive and complex than the two designs described previously. And it still leaves room for alternative explanations, because the intervention and control groups may differ in some undetected but important ways.

Pre-Test and Post-Test with Control Group Design (most expensive, most rigorous)



This design offers the greatest possibility of attributing evaluation outcomes to program activities. By randomly assigning individuals from the same target population to either an intervention or control group, all members of that target population have an equal chance of becoming a member of either group. This should ensure that members of the intervention and control groups are similar with respect to the key variables that might affect their performance on the pre-test and post-test. This type of evaluation is more complex and expensive than the other three, but provides the highest degree of certainty that it was the intervention that caused any changes detected by the evaluation.

Quantitative, Qualitative, and Mixed Approaches to Evaluation

Another choice to make when designing an evaluation is whether to gather quantitative data, qualitative data, or both.

Quantitative Data

Quantitative data can be counted, measured, and reported in numerical form and answer questions such as who, what, where, and how much. For example, a quantitative evaluation of a school-based violence prevention program might use disciplinary reports to discover that the intervention resulted in a 10 percent decrease in incidents of physical fighting on the campus.

The quantitative approach is useful for describing concrete phenomena and for statistically analyzing results, such as calculating the percentage decrease of cigarette use among 8th-grade students. Some examples of quantitative data include test scores, attendance rates, drop-out rates, and survey rating scales.

Advantages of collecting quantitative data include the following:

- Data collection instruments can be used with large numbers of study participants.
- Data collection instruments can be standardized, allowing for easy comparison within and across studies.
- Data are easily compiled for analysis.
- Findings can be presented succinctly.
- Findings are more widely accepted as being scientific and applicable than those from qualitative evaluations.

Qualitative Data

Qualitative data are reported in narrative form. Examples of qualitative data include written descriptions of program activities, testimonials of program effects, comments about how a program was or was not helpful, case studies, analyses of existing files, focus groups, key informant interviews, and observational studies. Qualitative evaluations might not yield results that are accepted as scientifically rigorous as those from quantitative evaluations. But the qualitative approach can provide important insights into how well a program is working and what can be done to increase its impact.

Qualitative data can also provide information about how participants – including the people responsible for operating the program as well as the target audience – feel about the program. For example, a qualitative evaluation of a school-based violence prevention program might use teacher interviews to find representative comments about what they thought about the programs impact, such as “I really think I learned a lot in the program. And I’ve been able to intervene in several situations that might otherwise have resulted in fights among students.”

Benefits of collecting qualitative data include the following:

- It promotes understanding of diverse stakeholder perspectives (e.g., what the program means to different people).
- It may shed light on unanticipated outcomes.
- Stakeholders, funders, policymakers, and the public may find quotes and anecdotes easier to understand and more appealing than statistical data.
- It can generate new ideas about how to make the program work better.

Mixed-Method Evaluations

The ideal evaluation combines quantitative and qualitative methods. A mixed-method approach offers a range of perspectives on a program's processes and outcomes. Benefits of this type of approach include the following:

- It increases the validity of your findings by allowing you to examine the same phenomenon in different ways.
- It can result in better data collection instruments. For example, focus groups can be invaluable in the development or selection of a questionnaire used to gather quantitative data.
- It promotes greater understanding of the findings. Quantitative data can show that change occurred and how much change took place, while qualitative data can help you and others understand what happened and why.
- It offers something for everyone. Some stakeholders may respond more favorably to a presentation featuring charts and graphs. Others may prefer anecdotes and stories.

By using different sources and methods at various points in the evaluation process, your evaluation team can build on the strengths of each type of data collection and minimize the weaknesses of any single approach.

References:

Additional evaluation resources can be found in the Evaluation Toolkit.

This publication is based on material from Locating, Hiring, and Managing an Evaluator, a web-based course designed and implemented by the Northeast Center for the Application of Prevention Technologies (CAPT), and from Are You Making Progress? Increasing Accountability Through Evaluation, a web-based course developed and implemented by CAPT and the National Coordinator Training and Technical Assistance Center, Health and Human Development Programs, Education Development Center, Inc.