
HOW CAN THE EFFECTS OR IMPACTS OF OUR STRATEGIES BE MEASURED?

Most groups who want to move toward antiracism and inclusion goals cannot do all of the things that it would take to make the change. For example, perhaps their strategies will change the way banks decide who is credit-worthy, while other groups may be working on changing the income of people who apply for loans. And no group may be working on the interest rates that banks charge in different parts of the community or to different groups of people.

In almost all cases, the strategies of a community change group are not comprehensive — that is, they do not address all the factors that affect the community outcome. Many other factors are at work — some contributing to the changes the group wishes to see happen, others countering those changes. Therefore, when changes in the community are observed, it is both important and difficult to tell whether it was a specific strategy or something else that led to those changes.

Control Group Design

Social science has modeled some of its methods of measuring effects or impacts of specific activities or strategies (often called “interventions”) on those of the natural sciences. In particular, there is one evaluation method that is well accepted as providing evidence of the effectiveness of a given strategy. This method closely approximates an experiment in which “subjects” are randomly assigned to receive the “intervention” or to be the target of a strategy and others are not. Conceptually, this method creates two groups who have virtually identical backgrounds and experiences, so that differences observed between the groups after implementation of the intervention or strategy can be ascribed to the intervention itself. This method – that of using randomly assigned control groups – is the “gold standard” for social science evaluations.

However, there are a number of reasons why it is often difficult to use the control group evaluation design. Sometimes there are not enough potential “subjects” interested in participating, or there are ethical concerns about withholding the intervention from some who might benefit. In other instances, the potential “subjects” are in naturally occurring groups so that the intervention or strategy cannot be used with some and not with others. An example might be students in classrooms where the strategy is teacher training in new instructional strategies. In still other instances, the strategy is intended to affect a group or institution as a whole. An example might be strategies to change the lending policies of banks with the goal of increasing the ease with which residents in specific neighborhoods can apply for and receive mortgages. In all of these situations it is not possible to randomly assign individual subjects to receive the intervention or be in the control group.

Two other evaluation designs are comparison groups and pre-post observations. Both have important limitations that groups should understand.

Comparison Group Design

When control groups are not feasible, evaluators sometimes try to create comparison groups. Comparison groups are usually formed by trying to match subjects who receive the intervention or are the target of a strategy with individuals, groups, institutions or communities that share many of the same characteristics but have not had the opportunity or interest to be involved in the intervention or strategy.

However, there are often differences between the groups that are difficult to measure but could be related to whether or not the strategy or intervention is likely to work.

HOW CAN THE EFFECTS OR IMPACTS OF OUR STRATEGIES BE MEASURED?

For example, two communities with similar demographics, housing stock and access to financial institutions may still be very different in terms of their history of working together, the kinds of leaders who have emerged or their patterns of making decisions. These kinds of differences (and many others) could influence how well a particular strategy works for that group.

If the outcomes for the groups are different, it will be difficult to know what caused that difference – whether it was that particular strategy, or these other kinds of things. And if the outcomes are the same between the two groups, you still have the same dilemma. Your strategy might have been effective but the things you couldn't measure about the groups may have affected the comparison group in ways that made it behave like the group with which you were working.

Comparison group designs can be helpful, when accompanied by an in- depth examination of differences between the groups, in understanding what factors may contribute to the strategy's effects.

Pre-Post Design

Pre-post measures of change in the individuals, groups, institutions or communities have the advantage of overcoming the problems of unknown differences between comparison groups. However, this design is not able to eliminate the possibility that the observed changes would have occurred for the same subjects without the intervention or strategy.

Even though these designs cannot rigorously assess whether and by how much a strategy is making a difference, each has value as long as its limitations are understood.

Pre-post change measures can establish whether a strategy meets a minimum threshold of effect. For example, think about measures of knowledge or skills conducted before and after a training program. In this case, regardless of whether program participants could have gained the same knowledge and skills in other ways, at a minimum training would have to have resulted in increased knowledge and skill to be at all worthwhile.

In other situations the data show that ground was lost on the outcomes of interest. In this case, when the pre-post measures show no change or even a reversal, the strategy may have been effective in slowing this process, but without independent evidence of the trend this conclusion would be difficult to defend.

Inferential Designs

There are also approaches to drawing conclusions about the success of a strategy or intervention based on finding consistent patterns of results. These might be called “inferential” designs in that they make inferences about a strategy from evidence that relates to an underlying theory or set of assumptions.

Developing a theory of change, and using it to guide evaluation, lends itself well to this approach. An expected chain of influence for a strategy or set of strategies is laid out in advance, linking activities to further actions or benefits, and to immediate, intermediate and longer-term outcomes. If the results you see follow the pattern you laid out in advance, that gives you some confidence that your strategies influenced the observed changes.

For example, suppose you had a theory of change that said: if we organize a community to demand increased access to capital, train bank leaders and loan officers about institutional racism and the bias built into current loan granting procedures, engage a few key power brokers in a personal transformation effort designed to open themselves up to change and offer technical support to banks and community groups on progressive lending strategies, then more people of color in our community will receive mortgage loans based on their current credit-worthiness.

HOW CAN THE EFFECTS OR IMPACTS OF OUR STRATEGIES BE MEASURED?

And then, suppose your evaluation showed that each of these strategies was implemented well and engaged the people you had hoped, that people's knowledge and skills changed in the ways intended, that people's behavior changed in the ways predicted, and that more mortgage loans were granted in the targeted neighborhood. Absent some other major influence on mortgage lending in your community, you would be in a good position to say that your strategies very likely contributed in substantial ways to the improved outcome.

A related approach is called "results mapping" in which in-depth data collection and analysis of a number of specific instances of the observed pattern of change are conducted. The focus of this approach is to create a "map" of strategies and actions and their results in individual cases, accumulating evidence across a series of cases to understand the extent to which planned activities lead to the expected changes in outcomes. This approach is especially helpful when the outcomes are a result of a complex set of events and/or involve healing, transformation or prevention.