

 Lagunita is retiring and will shut down at 12 noon Pacific Time on March 31, 2020. A few courses may be open for self-enrollment for a limited time. We will continue to offer courses on other online learning platforms; visit <http://online.stanford.edu>.

Course > Producing Data: Designing Studies > Experiments with One Explanatory Variable > Pitfalls in Experimentation

 Bookmark this page

Pitfalls in Experimentation

Learning Objective: Identify the design of a study (controlled experiment vs. observational study) and other features of the study design (randomized, blind etc.).

Learning Objective: Explain how the study design impacts the types of conclusions that can be drawn.

Pitfalls in Experimentation

Some of the inherent difficulties that may be encountered in experimentation are the Hawthorne effect, lack of realism, noncompliance, and treatments that are unethical, impossible, or impractical to impose.

We already introduced a hypothetical experiment to determine if people tend to snack more while they watch TV: Recruit participants for the study. While they are presumably waiting to be interviewed, half of the individuals sit in a waiting room with snacks available and a TV on. The other half sit in a waiting room with snacks available and no TV, just magazines. Researchers determine whether people consume more snacks in the TV setting.

Suppose that, in fact, the subjects who sat in the waiting room with the TV consumed more snacks than those who sat in the room without the TV. Could we conclude that in their everyday lives, and in their own homes, people eat more snacks when the TV is on? Not necessarily, because people's behavior in this very controlled setting may be quite different from their ordinary behavior. If they suspect their snacking behavior is being observed, they may alter their behavior, either consciously or subconsciously. This phenomenon, whereby people in an experiment behave differently from how

they would normally behave, is called the **Hawthorne effect**. Even if they don't suspect they are being observed in the waiting room, the relationship between TV and snacking there might not be representative of what it is in real life. One of the greatest advantages of an experiment—that researchers take control of the explanatory variable—can also be a disadvantage in that it may result in a rather unrealistic setting. **Lack of realism** (also called **lack of ecological validity**) is a possible drawback to the use of an experiment rather than an observational study to explore a relationship. Depending on the explanatory variable of interest, it may be quite easy or it may be virtually impossible to take control of the variable's values and still maintain a fairly natural setting.

In our hypothetical smoking cessation example, both the observational study and the experiment were carried out on a random sample of 1,000 smokers with intentions to quit. In the case of the observational study, it would be reasonably feasible to locate 1,000 such people in the population at large, identify their intended method, and contact them again a year later to establish whether they succeeded or not. In the case of the experiment, it is not so easy to take control of the explanatory variable (cessation method) merely by telling all 1,000 subjects what method they must use.

Noncompliance (failure to submit to the assigned treatment) could enter in on such a large scale as to render the results invalid. In order to ensure that the subjects in each treatment group actually undergo the assigned treatment, researchers would need to pay for the treatment and make it easily available. The cost of doing that for a group of 1,000 people would go beyond the budget of most researchers. Even if the drugs or therapy were paid for, it is very unlikely that most of the subjects contacted at random would be willing to use a method not of their own choosing, but dictated by the researchers. From a practical standpoint, such a study would most likely be carried out on a smaller group of volunteers, recruited via flyers or some other sort of advertisement. The fact that they are volunteers might make them somewhat different from the larger population of smokers with intentions to quit, but it would reduce the more worrisome problem of noncompliance. Volunteers may have a better overall chance of success, but if researchers are primarily concerned with which method is most successful, then the relative success of the various methods should be roughly the same for the volunteer sample as it would be for the general population, as long as the methods are randomly assigned. Thus, the most vital stage for randomization in an experiment is during the assignment of treatments, rather than the selection of subjects.

There are other, more serious drawbacks to experimentation, as illustrated in the following hypothetical examples:

Example

Suppose researchers want to determine if the drug Ecstasy causes memory loss. One possible design would be to take a group of volunteers and randomly assign some to take Ecstasy on a regular basis, while the others are given a placebo. Test them periodically to see if the Ecstasy group experiences more memory problems than the placebo group.

The obvious flaw in this experiment is that it is unethical (and actually also illegal) to administer a dangerous drug like Ecstasy, even if the subjects are volunteers. The only feasible design to seek answers to this particular research question would be an observational study.

Example

Suppose researchers want to determine whether females wash their hair more frequently than males.

It is impossible to assign some subjects to be female and others male, and so an experiment is not an option here. Again, an observational study would be the only way to proceed.

Example

Suppose researchers want to determine whether being in a lower income bracket may be responsible for obesity in women, at least to some extent, because they can't afford more nutritious meals and don't have the means to participate in fitness activities.

The socioeconomic status of the study subject is a variable that cannot be controlled by the researchers, so an experiment is impossible. (Even if the researchers could somehow raise the money to provide a random sample of women with substantial salaries, the effects of their eating habits during their lives before the study began would still be present, and would affect the study's outcome.)

These examples should convince you that, depending on the variables of interest, researching their relationship via an experiment may be too unrealistic, unethical, or impractical. Observational studies are subject to flaws, but often they are the only recourse.

Let's Summarize

1. Observational studies:

- * The explanatory variable's values are allowed to occur naturally.
- * Because of the possibility of lurking variables, it is difficult to establish causation.
- * If possible, control for suspected lurking variables by studying groups of similar individuals separately.
- * Some lurking variables are difficult to control for; others may not be identified.

2. Experiments

- * The explanatory variable's values are controlled by researchers (treatment is imposed).
- * Randomized assignment to treatments automatically controls for all lurking variables.
- * Making subjects blind avoids the placebo effect.
- * Making researchers blind avoids conscious or subconscious influences on their subjective assessment of responses.
- * A randomized controlled double-blind experiment is generally optimal for establishing causation.
- * A lack of realism may prevent researchers from generalizing experimental results to real-life situations.
- * Noncompliance may undermine an experiment. A volunteer sample might solve (at least partially) this problem.
- * It is impossible, impractical or unethical to impose some treatments.

Open Learning Initiative [🔗](#)



[🔗](#) Unless otherwise noted this work is licensed under a Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License [🔗](#).

© All Rights Reserved