To show that an explanatory variable *causes* change, we need an experiment.

General Education Statistics
Class Notes
The Design of Experiments (Section 1.6)

A well-designed observational study can give us good information but we can never be sure that the explanatory variable is actually *causing* the change in the response variable. Perhaps the effect is simply a coincidence. An experiment will allow us to determine **causality**.

Definitions: An **experiment** is a controlled study conducted to determine the effect of varying one or more explanatory variables (or **factors**) has on a response variable. Any combination of the values of the factors is called a **treatment**.

The **experimental unit** (or **subject**) is a person, object or some other well-defined item upon which a treatment is applied.

A **completely randomized design** is one in which each experimental unit is randomly assigned to a treatment.

Consider this experiment.

expl 1: The English Department of a community college is considering adopting an online version of the freshman English course. To compare the new online course to the traditional course, an English Department faculty member randomly splits a section of her course. Half of the students receive the traditional course and the other half is given an online version. At the end of the semester, both groups will be given a test to determine which performed better.

a.) What makes this an experiment and not an observational study?

b.) Who are the experimental units?

c.) What is the population for which this study applies?

d.) What are the treatments?

e.) What is the response variable?

Since we randomly assign individuals to each treatment, we can say more definitively that the explanatory variable (type of course) did in fact cause any change we notice in the response variable (test score). We still have to be careful in the design of our experiment. Let's delve into some details.

expl 2: Lipitor is a cholesterol lowering drug made by Pfizer. In the Collaborative Atorvastatin Diabetes Study (CARDS), the effect of Lipitor on cardiovascular disease was assessed in 2838 subjects, ages 40 to 75, with type 2 diabetes but without prior history of cardiovascular disease. In this **placebo-controlled**, **double-blind experiment**, subjects were split roughly in half and randomly assigned to either Lipitor 10 mg daily or a placebo. The subjects were followed for four years. The researchers counted the subjects in each group who experienced a major cardiovascular event, such as a stroke or heart attack. This experiment found that Lipitor did reduce the number of cardiovascular events (83 events in the Lipitor group versus 127 events in the placebo group) and deaths (61 in Lipitor group versus 82 in placebo group).

a.) What are the treatments?

b.) The group that gets the placebo is called the **control group**. What is a placebo? Why would we want half the subjects to take a placebo?

c.) What is the response variable? Is it a qualitative or quantitative variable?

For *each* subject, what are they measuring at the end?

d.) A **blind study** is one where the subjects do not know which treatment (Lipitor or placebo) they are receiving. What do you think **double-blind** means? Why would we want this?

We will *not* be designing our own experiments.

Steps of Designing an Experiment:

Step 1: Identify the explicit problem to be solved. This is often referred to as the **claim**. Identify the response variable and the population to be studied.

Step 2: Determine the factors that affect the response variable. Once the factors are identified, it must be determined which factors are to be fixed at some predetermined level (the control), which factors will be manipulated, and which factors will be uncontrolled.

Step 3: Determine the number of experimental units. As a general rule, choose as many experimental units as time and money allow. Techniques exist for determining sample size, provided certain information is available.

Step 4: Determine the level(s) of the factors (explanatory variables). There are two ways to deal with the factors: control and randomize.

1. **Control:** There are two ways to control the factors.

   a) Set the level of a factor at one value throughout the experiment (if you are *not* interested in its effect on the response variable).

   b) Set the level of a factor at various levels (if you are interested in its effect on the response variable). The combinations of the levels of all varied factors constitute the treatments in the experiment.

2. **Randomize:** Randomize the experimental units to various treatment groups so that the effects of variables whose level cannot be controlled is minimized. The idea is that randomization "averages out" the effect of uncontrolled explanatory variables.

Step 5: Conduct the Experiment.

a) **Replication** occurs when each treatment is applied to more than one experimental unit. This helps to assure that the effect of a treatment is not due to some characteristic of a single experimental unit. It is recommended that each treatment group have the same number of experimental units.

b) Collect and process the data by measuring the response variable. Any difference in the value of the response variable is assumed to be a result of differences in the level of the treatment.

Step 6: Test the claim. This is the subject of inferential statistics. Inferential statistics is a process in which generalizations about a population are made on the basis of results obtained from a sample. Provide a statement regarding the level of confidence in the generalization. Methods of inferential statistics are presented later in the text.

Definition: A **matched-pairs design** is an experimental design in which the experimental units are paired up. The pairs are matched up so that they are somehow related (that is, the same person before and after a treatment, twins, husband and wife, same geographical location, and so on). There are only two levels of treatment in a matched-pairs design.

Experiments on twins in World War II Germany:
Experiments on twin children in concentration camps were created to show the similarities and differences in the genetics of twins, as well as to see if the human body can be unnaturally manipulated. The central leader of the experiments was Josef Mengele, who from 1943 to 1944 performed experiments on nearly 1,500 sets of imprisoned twins at Auschwitz. About 200 people survived these studies. The twins were arranged by age and sex and kept in barracks between experiments, which ranged from injection of different dyes into the eyes of twins to see whether it would change their color to sewing twins together in attempts to create conjoined twins.
(source: https://en.wikipedia.org/wiki/Nazi_human_experimentation)

Coke versus Pepsi Taste Tests:
Taste tests are a good example of a matched-pairs design. When you are asked to sample both Coke and Pepsi and say which you like best, you are a "matched pair". But instead of two different people, they use one person with two different products.

Pre-tests paired with Post-tests:
Sometimes in a class, you will be given a test at the very beginning of the semester and the same test at the end of the course. This is done to determine if learning took place. Again, you are a "matched pair". Comparing the same person before and after the teaching gives more information than if they gave the tests to two different people.

Definition: Placebo effect: To know you are receiving any treatment can be powerful medicine. The **placebo effect** refers to how people will improve even though the treatment they received has no actual efficacy. It is truly mind over matter. Using placebos in experiments will help offset this effect in the data.

Definition: Hawthorne effect: This refers to when a person changes their behavior because they know they are being observed. Imagine trying to figure the percentage of people who wash their hands after using a public restroom by sitting next to the sink with a clipboard. How would you account for that?