Lesson 28: Designing Experiments

TERMS:

Control Group - A Control Group is treated identically in all respects to the group receiving the treatment except that the members of the control group do not receive the treatment.

Placebo – An inactive treatment. There is a proven phenomenon called the placebo effect. Patients receiving Placebo tablets which have no active drug ingredient (e.g. a sugar tablet) may experience a certain beneficial effect.

Three critical elements of experiment design:

1.) Control

Assign a placebo to control group so that you know if the treatment actually caused the effect.

Plan to control for possible lurking variables using blocking, matched pairs, etc.

Once data is collected and the statistical analysis appears to show a link between the treatment and a resulting effect, the researchers will apply controls to the data to see if the effect holds. Examples of these controls include separating subjects by gender, race, education, socio-economic status, health, age, etc. By showing that the other variables are NOT changing the effect, the evidence that the treatment caused the effect strengthens.

2.) Randomization

Randomly draw the sample from the population (SRS is ideal). If subjects are not drawn from the target population, then the results cannot be generalized to the population and are limited to the subjects in the study (No volunteers).

Randomly assign the treatments to the subjects.

An experiment that randomly assigns treatments to subjects selected randomly from the population is called a **Completely Randomized Design**.

3.) Replication

Use a large sample size so the experiment is repeated many times.

Ideally, studies are repeated in multiple settings by different research teams. When a study is published with a surprising or profound finding, it is often replicated by researchers that are independent of the original team to see if the effect holds.

When the link between the treatment and result holds despite controlling for lurking variables and can be replicated, the finding is called **Robust**.

Daily Data Collection

All students will then flip a coin and be assigned to the heads or tails group. All students will perform the ruler drop test to test their reaction time. This will be performed 5 times and the averages recorded. If heads, they will drink 4 oz. of soda from the red cup. If tails, they will drink 4 oz. of soda from the blue cup. One cup is regular coke and one is caffeine free coke. After 20 minutes the ruler drop test will be repeated.

Question: does caffeine affect reaction time?

	Heads		Tails	
	Baseline	After Treatment	Baseline	After Treatment
1				
2				
3				
4				
5				
6				
7				
8				
9				
10				
11				
12				
13				
14				
15				
16				
17				
18				
19				
20				
Mean				
St Dev				

Examples:

When Will I Ever Use This Stuff?

Vocabulary of experiments

Researchers at the University of North Carolina were concerned about the increasing dropout rate in the state's high schools, especially for low-income students. Surveys of recent dropouts revealed that many of these students had started to lose interest during middle school. They said they saw little connection between what they were studying in school and their future plans. To change this perception, researchers developed a program called CareerStart. The big idea of the program is that teachers show students how the topics they learn get used in specific careers.

To test the effectiveness of CareerStart, the researchers recruited 14 middle schools in Forsyth County to participate in an experiment. Seven of the schools, chosen at random, used CareerStart along with the district's standard curriculum. The other seven schools just followed the standard curriculum. Researchers followed both groups of students for several years, collecting data on students' attendance, behavior, standardized test scores, level of engagement in school, and whether the students graduated from high school. *Results:* Students at schools that used CareerStart generally had better attendance and fewer discipline problems, earned higher test scores, reported greater engagement in their classes, and were more likely to graduate.[20]

PROBLEM: **Identify the experimental units, explanatory and response variables, and the treatments in the CareerStart experiment.**

TV Advertising

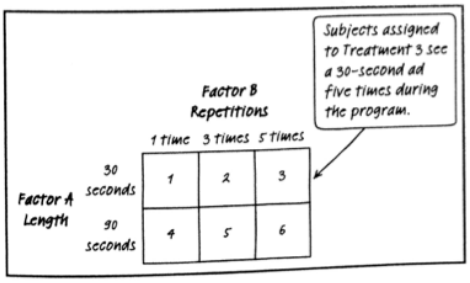
Experiments with multiple explanatory variables

What are the effects of repeated exposure to an advertising message? The answer may depend on both the length of the ad and on how often it is repeated. An experiment investigated this question using undergraduate students as subjects. All subjects viewed a 40-minute television program that included ads for a digital camera. Some subjects saw a 30-second commercial; others, a 90-second version. The same commercial was shown either 1, 3, or 5 times during the program. After viewing, all the subjects answered questions about their recall of the ad, their attitude toward the camera, and their intention to purchase it.[21]

PROBLEM: **For the advertising study,**

(a) identify the explanatory and response variables, and

(b) list all the treatments.



Which Works Better: Online or In-Class SAT Preparation?

A bad experiment

A high school regularly offers a review course to prepare students for the SAT. This year, budget cuts will allow the school to offer only an online version of the course. Over the past 10 years, the average SAT score of students in the classroom course was 1620. The online group gets an average score of 1780. That's roughly 10% higher than the long-time average for those who took the classroom review course. Is the online course more effective?

This experiment has a very simple design. A group of subjects (the students) were exposed to a treatment (the online course), and the outcome (SAT scores) was observed. Here is the design:

Students → Online course → SAT scores

A closer look showed that the students in the online review course were quite different from the students who took the classroom course in past years. They had higher GPAs and were taking more AP classes.

The effect of online versus in-class instruction is mixed up with the effect of these lurking variables. Maybe the online students earned higher SAT scores because they were smarter to begin with, not because the online course prepared them better. This confounding prevents us from concluding that the online course is more effective than classroom instruction.

SAT Prep: Online versus Classroom

How random assignment works

This year, the high school has enough budget money to compare the online SAT course with traditional classroom instruction. Fifty students have agreed to participate in an experiment comparing the two instructional methods.

PROBLEM: **Describe how you would randomly assign 25 students to each of the two methods:**

(a) Using 50 identical slips of paper

(b) Using Table D

(c) Using technology

Conserving Energy

A completely randomized design

Many utility companies have introduced programs to encourage energy conservation among their customers. An electric company considers placing small digital displays in households to show current electricity use and what the cost would be if this use continued for a month. Will the displays reduce electricity use? One cheaper approach is to give customers a chart and information about monitoring their electricity use from their outside meter. Would this method work almost as well? The company decides to conduct an experiment to compare these two approaches (display, chart) with a *control group* of customers who receive information about energy consumption but no help in monitoring electricity use.

PROBLEM: **Outline a completely randomized design involving 60 single-family residences in the same city who are willing to participate in such an experiment. Write a few sentences describing how you would implement your design.**

CHECK YOUR UNDERSTANDING

Music students often don't evaluate their own performances accurately. Can small-group discussions help? The subjects were 29 students preparing for the end-of-semester performance that is an important part of their grade. Assign 15 students to the treatment: videotape a practice performance, ask the student to evaluate it, then have the student discuss the tape with a small group of other students. The remaining 14 students form a control group who watch and evaluate their tapes alone. At the end of the semester, the discussion-group students evaluated their final performance more accurately.[22]

1. Outline a completely randomized design for this experiment. Follow the model of Figure 4.4.
2. Describe how you would carry out the random assignment. Provide enough detail that a classmate could implement your procedure.
3. What is the purpose of the control group in this experiment?

HW 28 - Section 4-2: 57, 63, 65, 67, 69, 71