Does the Difference Make a Difference?		
Term	**Before reading, I thought this meant…**	**After reading I learned…**
statistically significant	When an observed difference is too large to be believed to have occurred by chance.	

Blinding, Placebos, Blocking, Confounding		
Term	**Before reading, I thought this meant…**	**After reading I learned…**
Treatment blinding	When the person being treated does not know if they are getting the actual treatment or if they are getting a placebo.	
Single and double blind	When one or both of the groups associated with the experiment (subjects/treaters/etc. and evaluators) do not know who has received treatment	
placebo	A "fake" treatment that looks just like the actual treatment	
placebo effect	Subjects treated with placebos sometimes show a response just because they believe they are getting some treatment. (20% or more)	
block design	When there are pre-existing difference between groups of experimental units, we randomize treatments within the groups (blocks)	
matched design	Subjects who are similar in ways not being studied are paired and then compared with each other on the variables of interest.	
confounding factors	When the levels of one factor are entangled with the levels of another factor such that their effects cannot be separated.	

Describe some ethical considerations in experiments using humans or animals.

harmful to animals/people

does the good outweigh the bad

Experiments and Samples

obs. studies

- Both experiments and sample surveys use randomization to get unbiased data.
- But they do so in different ways and for different purposes:
  - Sample surveys attempt to randomly select participants from the population in order to estimate population parameters. allow for generalization to overall population
  - Experiments try to assess the effects of treatments are usually done with a pool of volunteers (which can create its own set of difficulties). But we randomly assign the treatments to our volunteers to reduce bias. spreads the effect of uncontrollable factors

 Chapter 1, Slide 12

The best experiments are usually

- randomized.
- comparative.
- double-blind.
- placebo-controlled.

Step-by-Step Example DESIGNING AN EXPERIMENT

An ad for OptiGro plant fertilizer claims that with this product you will grow "juicier, tastier" tomatoes. You'd like to test this claim, and wonder whether you might be able to get by with half the specified dose. How can you set up an experiment to check out the claim?

Of course, you'll have to get some tomatoes, try growing some plants with the product and some without, and see what happens. But you'll need a clearer plan than that. How should you design your experiment?

Let's work through the design, step by step. We'll design the simplest kind of experiment, in which each subject is equally likely to end up in any treatment group. This is called a **completely randomized experiment**. Since this is a *design* for an experiment, most of the steps are part of the *Think* stage. The statements in the right column are the kinds of things you would need to say in *proposing* an experiment. You'd need to include them in the "methods" section of a report once the experiment is run.

A **completely randomized experiment** is the ideal simple design, just as a *simple random sample* is the ideal simple sample—and for many of the same reasons.

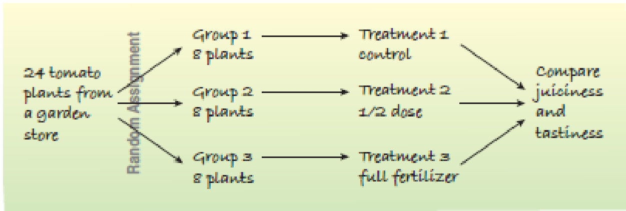
QUESTION: How would you design an experiment to test OptiGro fertilizer?

THINK ➡ Plan State what you want to know.	I want to know whether tomato plants grown with OptiGro yield juicier, tastier tomatoes than plants raised in otherwise similar circumstances but without the fertilizer. If so, I wonder if half as much fertilizer would also work.

(*continued*)

SHOW ➡ Response Specify the response variable.	I'll evaluate the juiciness and taste of the tomatoes by asking a panel of judges to rate them on a scale from 1 to 7 in juiciness and in taste.
Treatments Specify the factor levels and the treatments.	The factor is fertilizer, specifically OptiGro fertilizer. I'll grow tomatoes at three different factor levels: some with no fertilizer, some with half the specified amount of OptiGro, and some with the full dose of OptiGro. These are the three treatments.
Experimental Units Specify the experimental units.	I'll buy 24 tomato plants of the same variety from a local garden store.
Experimental Design Observe the principles of design:	
■ **Control** any sources of variability you know of and can control.	I'll locate the garden plots near each other so that the plants get similar amounts of sun and rain and experience similar temperatures. I will weed the plots equally and otherwise treat the plants alike.
■ **Replicate** results by placing more than one plant in each treatment group.	I'll use 8 plants in each treatment group.
■ **Randomly assign** experimental units to treatments, to equalize the effects of unknown or uncontrollable sources of variation. Describe how the randomization will be accomplished.	I'll take 24 cards—8 hearts, 8 diamonds, and 8 spades—from a deck and shuffle them thoroughly. Then I'll deal out one card in front of each tomato plant. I'll put the 8 plants with hearts in Group 1, the 8 with diamonds in Group 2, and the remaining 8 with spades in Group 3.

Make a Picture A diagram of your design can help you think about it clearly.



Specify any other experiment details. You must give enough details so that another experimenter could exactly replicate your experiment. It's generally better to include details that might seem irrelevant than to leave out matters that could turn out to make a difference.

I will grow the plants until the tomatoes are mature, as judged by reaching a standard color.

I'll harvest the tomatoes when ripe and store them for evaluation.

Specify how to measure the response.

I'll set up a numerical scale of juiciness and one of tastiness for the taste testers. I'll ask several people to taste slices of tomato and rate them.

Once you collect the data, you'll need to display them and compare the results for the three treatment groups.

I will display the results with side-by-side boxplots to compare the tasters' evaluations of the three treatment groups.

TELL ➡ To answer the initial question, we ask whether the differences we observe in the means of the three groups are meaningful.

Because this is a randomized experiment, we can attribute significant differences to the treatments. To do this properly, we'll need methods from what is called "statistical inference," the subject of much of the rest of this book.

If the differences in taste and juiciness among the groups are greater than I would expect by knowing the usual variation among tomatoes, I may be able to conclude that these differences can be attributed to the fertilizer.

Just Checking

1. At one time, a method called "gastric freezing" was used to treat people with peptic ulcers. An inflatable bladder was inserted down the esophagus and into the stomach, and then a cold liquid was pumped into the bladder. Now you can find the following notice on the Internet site of a major insurance company:

   [Our company] does not cover gastric freezing (intragastric hypothermia) for chronic peptic ulcer disease. . . .

   Gastric freezing for chronic peptic ulcer disease is a non-surgical treatment which was popular about 20 years ago but now is seldom performed. It has been abandoned due to a high complication rate, only temporary improvement experienced by patients, and a lack of effectiveness when tested by double-blind, controlled clinical trials.

   What did that "controlled clinical trial" (experiment) probably look like? (Don't worry about "double-blind"; we'll get to that soon.)

   a) What was the factor in this experiment?
   b) What was the response variable?
   c) What were the treatments?
   d) How did researchers decide which subjects received which treatment?
   e) Were the results statistically significant?

(Check your answers on page 265.)

ANSWERS

1. a) The factor was type of treatment for peptic ulcer.
   b) The response variable could be a measure of relief from gastric ulcer pain or an evaluation by a physician of the state of the disease.
   c) Treatments would be gastric freezing and some alternative control treatment.
   d) Treatments should be assigned randomly.
   e) No. The website reports "lack of effectiveness," indicating that no large differences in patient healing were noted.

Just checking

2. Recall the experiment about gastric freezing, an old method for treating peptic ulcers that you read about in the first Just Checking. Doctors would insert an inflatable bladder down the patient's esophagus and into the stomach and then pump in a cold liquid. A major insurance company now states that it doesn't cover this treatment because "double-blind, controlled clinical trials" failed to demonstrate that gastric freezing was effective.

a) What does it mean that the experiment was double-blind?

b) Why would you recommend a placebo control?

c) Suppose that researchers suspected that the effectiveness of the gastric freezing treatment might depend on whether a patient had recently developed the peptic ulcer or had been suffering from the condition for a long time. How might the researchers have designed the experiment?

(Check your answers on page 265.)

ANSWERS

2. a) Neither the patients who received the treatment nor the doctor who evaluated them afterward knew what treatment they had received.

b) The placebo is needed to accomplish blinding. The best alternative would be using body-temperature liquid rather than the freezing liquid.

c) The researchers should block the subjects by the length of time they had had the ulcer, then randomly assign subjects in each block to the freezing and placebo groups.

Homework/Classwork:

Pg. 259-262 #1-10

Skip #3 and #5