Notes for Experimental Design

<u>Terminology</u>:

Experiment – A procedure for investigating the effect of an experimental condition on a response variable.

Experimental units – Individuals on whom an experiment is performed (usually called subjects or participants).

Factor – A variable whose levels are controlled by the experimenter.

Extraneous factor – A variable that is not of interest in the current study but is thought to affect the response variable.

Treatment – The process, intervention, or other controlled circumstance applied to randomly assigned experimental units. Treatments are the different levels of a single factor or are made up of combinations of levels of two or more factors.

Confounding variable - associated in a noncausal way with a factor and affects the response (usually found in experiments). Because of the confounding, we find that we can't tell whether any effect we see was caused by our factor or by the confounding variable – or even both working together.

For example, a professor of Cornell University performed an experiment to investigate the effect of a teacher's classroom style on student evaluations. He taught a class in developmental psychology during two successive terms to students in two very similar classes. He kept everything about his teaching identical (same text, syllabus, office hours, etc.) and modified only his style in class. During the fall term he maintained a subdued demeanor. During the spring term, he used expansive gestures and lectured with more enthusiasm, varying his vocal pitch and using more hand gestures. He administered a standard student evaluation form at the end of each term. The students in the fall term class rated him only an average teacher. Those in the spring term class rated him an excellent teacher, praising his knowledge and accessibility, and even the quality of the textbook.

How much of the difference he observed was due to his difference in manner, and how much might have been due to the season of the year? Fall term starts out warm and sunny and ends cold and black. Spring term starts out cold and black and ends with blooming flowers and singing birds. Might students' overall happiness have been affected by the season and reflected in their evaluations?

Direct control – Holding extraneous factors constant so that their effects are not confounded with those of the experimental conditions.

For example, suppose we test two laundry detergents and carefully control the water temperature at 180°F. This would reduce the variation in our results due to water temperature, but what could we say about the detergents' performance in cold water? Not much.

Blocking – When groups of experimental units are similar (have common extraneous factors), it is often a good idea to gather them together into blocks. By blocking we isolate the variability attributable to the differences between the blocks so that we can see the differences caused by the treatments more clearly. Blocking is an important compromise between randomization and control, but not required in an experimental design.

For example, say we were allocating players to two 5-player basketball teams from a pool of 10 children. We might do so at random to equalize the talent. But what if there were two 14-year-olds and eight 11-year-olds in the group? Randomizing may place both 14-year-olds on the same team. So wouldn't it be better to assign one 14-year-old to each team (at random) and four of the 11-year-olds to each team (at random)? By doing this we improve the fairness in the short run. When we do this the variable *age* is called a **blocking variable**. The levels of *age* are called **blocks**.

Randomized block design – When randomization occurs only within blocks.

Placebo – A treatment known to have no effect, administered so that all groups experience the same conditions. Only by comparing with a placebo can we be sure that the observed effect of a treatment is not due simply to the placebo effect.

Placebo effect – the tendency of many human subjects to show a response even when administered a placebo. In fact, it's not unusual for 20% or more of subjects given a placebo to report an improvement.

Control group – The experimental units assigned to a baseline treatment level, typically either the default treatment, no treatment, or a placebo treatment.

Blind(ing) – Any individual associated with an experiment who is not aware of how subjects have been allocated to treatment groups.

Double-blind – When every individual who can either *influence* the results (i.e. subjects) or *evaluate* the results (i.e. judges, treating physicians) is blinded.

Principals of experimental design –

- **Randomize** subjects to treatments (or of treatments to trials) to even out effects we cannot control and/or ensure that the experiment does not systematically favor one experimental condition over another.
- **Control** aspects of the experiment that we know may have an effect on the response, but that are not the factors being studied.
- **Replicate** over as many subjects as possible. Results for a single subject are just anecdotes.
- **Block** to reduce the effects of identifiable attributes of the subjects that cannot be controlled.

The best experiments are usually randomized, comparative, double-blind, and placebo-controlled.

Before proceeding with an experiment, you should be able to give a satisfactory answer to each of the following questions:

1. What is the research question that data from the experiment will be sued to answer?
2. What is the response variable?
3. How are the values of the response variable to be determined?
4. What are the factors (explanatory variables) for the experiment?
5. For each factor, how many different values are there, and what are these values?
6. What are the treatments for the experiment?
7. What extraneous variables might influence the response?
8. How does the design incorporate random assignment of subjects to treatments (or treatments to subjects) or random assignment of treatments to trials?
9. For each extraneous variable, how does the design protect against its potential influence on the response through blocking, direct control, or randomization?
10. Will you be able to answer the research question using the data collected in this experiment?

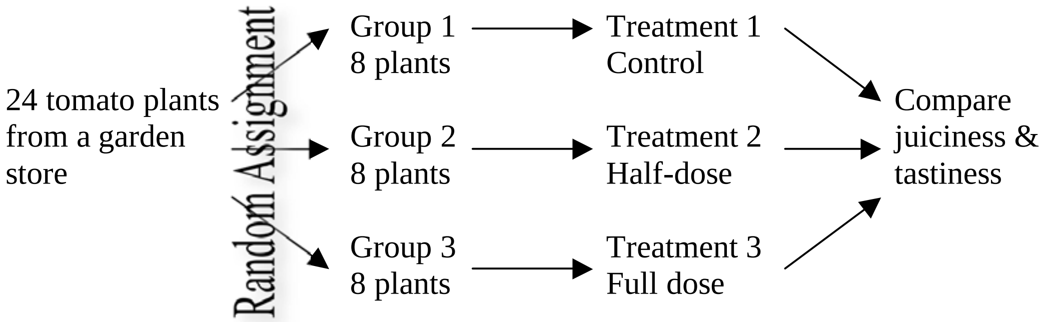
Designing an Experiment *Step-by-Step*

<u>*One factor example*</u>: An ad for OptiGro plant fertilizer claims that with this product we will grow “juicier, tastier” tomatoes. We’d like to test this claim, and wonder whether we might be able to get by with half the specified dose. How can we set up an experiment to check out the claim?

The five steps below represent only the steps involved in explaining where you’re headed and why. You must also show the mechanics of calculating statistics and tell what you’ve learned about your experiment. Statements in the right column are the kinds of things you would need to say in *proposing* an experiment and are included in the *methods* section of a report once the experiment is run.

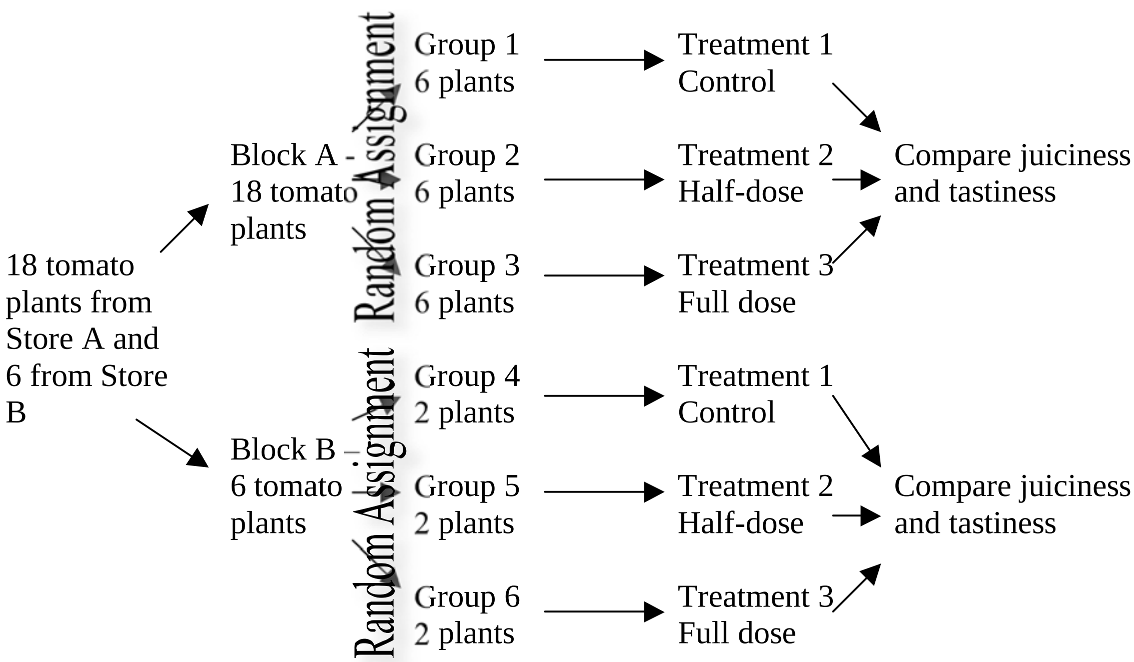
1. **Plan –** state what you want to know.
   We want to know whether tomato plants grown with OptiGro yield juicer, tastier tomatoes than plants raised in otherwise similar circumstances but without the fertilizer.
2. **Response** – specify the response variable
   We’ll evaluate the juiciness and taste of the tomatoes by asking a panel of judges to rate them on a scale of 1 to 7 in juiciness and in taste.
3. **Treatments –** specify the factor levels and the treatments
   The factor is OptiGro’s fertilizer. We’ll grow tomatoes at three different factor levels: some with no fertilizer, some with half the specified amount, and some with the full dose of OptiGro. These are the three treatments.
4. **Experimental units** – specify the experimental units
   We’ll obtain 24 tomato plants of the same variety from a local garden store.
5. **Experimental design** – observe the principles of design:
   a. **Control** any sources of variability you know of and can control
      We’ll locate the farm plots near each other so that the plants get similar amounts of sun and rain and experience similar temperatures.
   b. **Randomly assign** experimental units to treatments to equalize the effects of unknown or uncontrollable sources of variation. Specify how the random numbers needed for randomization will be obtained
      We’ll randomly divide the plants into three groups. We’ll use random numbers from a table to determine the assignment.
   c. **Replicate** results by placing more than one plant in each treatment group
      There are 8 plants in each treatment group.
   d. **Make a picture –** a diagram of your design can help you think about it clearly

Diagram



<u>*Blocking example*</u>: Say we wanted to plant 24 tomato plants but the garden store only had 18 plants left. So we drive down to another nursery and buy 6 more plants of that variety. We may worry that the tomato plants from the two stores are different somehow. We can design an experiment by blocking.

Diagram



What Can Go Wrong?

Don't give up just because you can't run an experiment. Sometimes we can't run an experiment because we can't identify or control the factors. If we can't perform an experiment, often an observational study is a good choice.

Beware of confounding. Use randomization whenever possible to ensure that the factors not in your experiment are not confounded with your treatment levels. Be alert to confounding that cannot be avoided, and report it along with your results.

Bad things can happen even to good experiments. It's generally a good idea to collect as much information as possible about your experimental units and the circumstances of the experiment. For example, in the tomato experiment, it would be wise to record details of the weather that might affect the plants and any facts available about their growing situation. Sometimes we can use this extra information during the analysis to reduce bias.

Don't spend your entire budget on the first run. It's a good idea to try a small pilot experiment before running the full-scale experiment. You may learn how to choose factor levels more effectively, about effects you forgot to control, and/or about unanticipated confoundings.

This packet of information about experimental design has been adapted from "Stats: Modeling Our World" by Bock, Velleman, and DeVeaux.