Chapter 2

Fundamentals of Experimental Design

One of the aims of most scientific investigations is to establish a link between cause and effect. For instance, a researcher might wish to determine what effect than an increase in atmospheric carbon dioxide (CO_2) will have on plant photosynthesis. Given the researcher's background knowledge, he or she comes up the hypothesis that an increase in CO_2 will result in an increase in leaf photosynthetic rate. The next step would be for the researcher to design an experiment to test the hypothesis.

So how does one design an experiment? This may seem like a trivial question and most biology students would probably quickly answer that you need a control and a treatment group. However, there is more to designing an effective experiment than this, and, in fact, not all experiments need a control group.

An effective experimental design is one that reduces the level of variation in the data that is not due to the imposed treatment (**error variation**) as well as minimizes the influence of extraneous variables. Error variation could result from the improper calibration or use of instruments, inconsistent procedures or random causes (not necessarily controllable). An experiment that eliminates sources of bias and systematic errors coupled with a high degree of accuracy and precision in measurements will result in minimal error variation. Thus, the remaining variation in the data can be attributed to either to random, natural variation or to the imposed treatment or condition. Because biological systems are by nature integrated systems it is important to *control* all other factors that may influence the variable of interest, the **response variable**, so that we can isolate treatment effects. While a good experimental design might not necessarily have a "control" group, it is necessary that an experiment be controlled so that the influences of factors other than the treatment are minimized.

These goals can be accomplished by addressing the two of the fundamental principles of experimental design: replication and randomization. While most students have heard those terms before, or at least variations of them, it is important to really understand them in order to design an effective experiment.

Replication allows us to control and quantify random variation and obtain reliable estimates of population parameters from a sample. A **replicate** is simply a repeated measure of the same variable on a different unit. While this may seem straight forward, in practice it isn't always an easy task to obtain *true* replicates. There is a potential trap called **pseudoreplication**. Pseudoreplication is making repeated measurements on non-independent units. As an example, let's consider our previous scenario investigating the effects of elevated CO_2 on photosynthesis. The researcher took six plants, and divided them equally between two environments that differed in CO_2 concentrations. Then photosynthetic rates measured were measured on three leaves per plant. What would be the level of replication, also referred to as the sample size, for each group (n)? If the researcher analyzed the data as n=9 (3 measurements x 3 plants), he or she would be guilty of pseudoreplication because the 3 measurements on each plant actually represents a sub-sampling of individuals; thus, there would really only be 3 replicates in this experiment (n=3).

Sub-sampling itself isn't a bad thing; in fact, it can be a very useful way to assess the **precision** of measurements and improve **accuracy**. Precision is a measure of how repeatable each measurement is on a single individual, and the precision determines the number of significant digits that must be reported. Another reason for making multiple measurements per individual is to ensure that each measurement accurately represents the value for that individual. Sub-samples can be averaged and then the average be used as the value for that individual. This provides greater accuracy if measurements tend to fluctuate over short time periods. Plus, if there is an error in measurements a single measurement can be discarded rather than loosing a replicate.

Maximizing the number of true replicates increases the accuracy of population parameter estimate and allows the magnitude of natural variation to be quantified. Choosing the appropriate sample size for a study is not any easy task because the greater the level of random variation, the higher the number of replicates that are needed to obtain a reliable estimate of population parameters (a necessary requisite to distinguish treatment effects). There are techniques that one can use to estimate the necessary sample size for a given level of error variation; however, these are often impractical for student research projects. If time permits, one can run a trial investigation that is identical to the "real" experiment in all aspects except using a very small sample size (e.g. n=3). This is useful for several reasons. First, you get an idea of any problems with the experimental design and can make modifications before investing a lot of resources. Secondly, by analyzing the trial data you can quantify the level of error variation and increase the number of replicates accordingly. In most cases the choice of an appropriate sample size is also going to be influenced by practical concerns such as availability of resources, time, and money.

One criterion of statistical analyses is that samples consist of individuals chosen from the parent population at random. Failure to do so could mean sampling from a population other than the one of biological interest. For example, if a researcher did a study on a species of desert rat collected from the field using bait traps set near cacti. They may only be learning about the population of the species that lives or forages near cacti and not about the species as a whole. Randomization also eliminates the problems of inducing systematic effects and bias into the experimental design all of which would lead to data that is essentially worthless.

Systematic effects are influences (often environmental) to which only a subset of the individuals being studied are subjected. They can often be eliminated by randomization in the spatial organization of treatment groups. For instance, while it might be convenient to lump all of the individuals in a particular treatment group together this may introduce error because adjacent individuals tend to be more similar than those that are spatially separated. A classic within a greenhouse can result in dramatic differences that are simply due to the subtle environmental gradients that exist (i.e. variation in light intensity) rather than due to the treatment being applied (Figure 2.1). Organizing the plants and then randomly assigning individuals to treatment groups would evenly distribute any environmental influences, thus minimizing systematic effects (Figure 2.1).

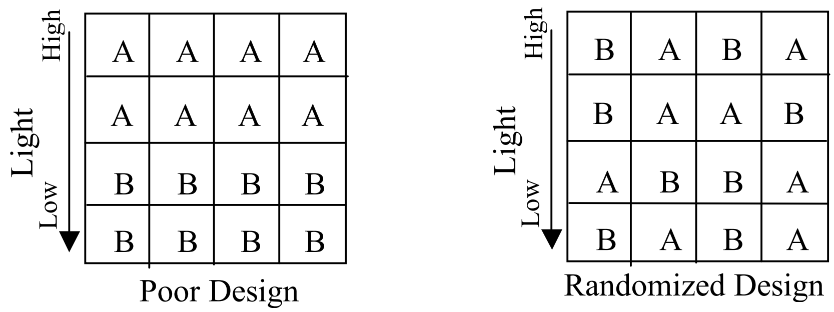

Figure 2.1. Examples of a poor treatment design (left) that is not randomized and one that is randomized (right) to minimize systematic effects. Letters A and B represent different treatments.

Randomization can also eliminate sources of bias, which can be completely unintentional but lead to ambiguous or incorrect results. For instance, if one were to assign select 10 seedlings from a group of 25 and designate them into one treatment group and then designate another 10 as the second group this could lead to bias since it is human nature to select the healthiest looking individuals first. A better approach would be to number all of the 25 plants and then use a random number generator to create two groups of 10 numbers. Then the treatments would be assigned to individuals randomly rather than by the experimenter.

Attention to experimental design before initiating any scientific investigation is important because no statistical procedure can fix a bad experiment. In designing an experiment you need to ensure adequate replication and randomization both in sample collection and treatment assignment. A focus on these two aspects of experimental design will greatly improve your experiments and eliminate time wasted collecting useless data.