# ELEMENTS OF EXPERIMENTAL DESIGN

# Elements of Experimental Design

In the previous lecture, we recognized many potential challenges we need to address when designing an experiment for comparing different treatments. We are now ready to introduce the five basic elements of experimental design:

0. A research question.
1. A set of treatments to compare.
2. A set of experimental units.
3. **Assignment of treatments to the experimental units.**
4. Measurements of experimental outcomes (on experimental units).

# The set of treatment to be studied

Suppose we now have an interesting research question and we will investigate the question using a comparative experiment.

The **treatments** are a set of conditions and/or procedures we will subject each experimental unit to:

- In the contact lens example, the two types of lenses are the treatments.
- In a clinical trial, treatments may include different doses of a drug administered to patients.
- In a crop field experiment, treatments can be different combinations of irrigation methods, fertilizers.

# The importance of a control treatment

Often times, we will include a control treatment to establish a baseline response level:

- A control treatment can be “no treatment”.
- In plant studies, a control treatment sometimes means treating the plants with water as opposed to some chemicals.
- In clinical trials, the control treatment can be a placebo drug or a sham procedure (more on this point later).

# Example: inadequate control treatment

In some experiments, we make measurements overtime. Some researchers will think that measurements at time 0 can be used as a the control. This may not be adequate.

Suppose we want to study pathogen effect on the leaves of plant: we make a measure at time 0, apply the pathogen, make measurements at 10 minutes, 1 hr, 10 hrs, and so on. What is the problem with this set up?

# Example: inadequate control treatment

In some experiments, we make measurements overtime. Some researchers will think that measurements at time 0 can be used as a the control. This may not be adequate.

Suppose we want to study pathogen effect on the leaves of plant: we make a measure at time 0, apply the pathogen, make measurements at 10 minutes, 1 hr, 10 hrs, and so on. What is the problem with this set up?

**Problem:** Under this set up, we can see how things change over time—but only under the pathogen treatment. We don't know what will happen over time if no pathogen is applied.

# Placebo

In clinical trials, a **control** treatment often means prescribing a “placebo” to patients.

A placebo (a dummy pill) is made to look like the real drug, but does not have the active substance/ingredient. A placebo can also be a procedure.

**A placebo is different from “no treatment”.**

A placebo can actually have real effect: many patients indeed get better after receiving the placebo treatment. For a drug to be approved by FDA, it has to perform better than a placebo.

It is not always possible/ethical to use a “placebo” control: e.g., in AIDS trials, the ‘controls’ will receive the best existing therapy.

# Double-blind study

In clinical trials, to avoid bias, we often use a “double-blind” study.

In a “double-blind” study, neither the patients, nor the doctors administering the drugs or researchers monitoring the outcome, know whether the patient receive the real pill or the placebo.

“Double-blind methods can be applied to any experimental situation in which there is a possibility that the results will be affected by conscious or unconscious bias on the part of researchers, participants, or both.” — Wikipedia

# The set of experimental units in the study

An experimental unit is the unit to which the treatment is applied.

An experiment unit is the smallest sub-unit of the experimental material such that any two different units may potentially receive different treatments.

An experimental unit may not be the smallest observational unit in the sample.

To correctly identify the experimental units, we need to understand how the treatments are assigned to the units.

# A note on terminology

Recall that, in statistics, when we say a **sample**, we mean a subset of individuals from a population.

In plant science, sometimes a sample means a single plant. To avoid confusion, in this class, we will try to use “units” to refer to individual plants, and reserve the use of “sample” to mean a subset of individuals from a population.

# Choosing experimental units

When choosing the experimental units, the keyword is **representative**: the units used in the experiment should be representative of units that we would like to generalize our experimental conclusions to.

Experimental units should be representative of some population, but similar enough that random (experimental) errors are small. The greater the similarity of the experimental units, the more precise the inference—but, also, the less generalizable the results. (For example, race and gender in medical trials.)

The efficiency of the experiments with heterogeneous experimental units can be increased by **blocking** (comparing treatments with blocks of similar units).

# Assignment of treatments to experimental units

In a designed experiment, we have to decide how to assign the treatments to the individuals in the sample. This is the most important element of experiment design.

In fact, only after we see how treatments are assigned, can we clearly identify the experimental units.

Randomization is the key in this step: random assignment of treatments protects against bias, by balancing potentially **confounding factors** between treatment groups.

# Randomization

Problem with observational studies: many factors differ between "treatments", so there is the potential for confounding (effects of a factor that is related to both the treatment and the response, e.g., socioeconomic status in the cancer/smoking example).

In a designed experiment, randomization—introduced by Fisher—ensures that confounding factors will be balanced between treatments, if the sample size $n$ is large enough.

Before Fisher, treatment had been assigned on a subjective or systematic basis.

Random assignment of treatments protects against bias, by balancing potentially confounding factors between groups.

Throughout this course, we will emphasize **follow the randomization**.

# Example 1. Identify experimental unit

Example 1.1. from Kuehl. A simple animal diet ration study has one cage or pen of six animals assigned to ration A and a second cage or pen of six animals assigned to ration B. Weight gain or some other appropriate data are collected to test the efficacy of the rations. The necessary measurements are made on each of the animals in the pens at the end of the study.

What are the experimental units in this study?

# Example 1 continued

The animals are not the experimental units, even though they are the smallest observational units.

The treatments, rations A and B, are applied at the pen level, so the two pens are the experimental units.

As a consequence, in this experiment, **the treatments are not replicated**. As we will see later, that means, we will not be able to evaluate the error variation from such an unreplicated experiment.

# Example 1 continued

One possible solution is to use more cages/pens: we can four pens with three animals in each pen.

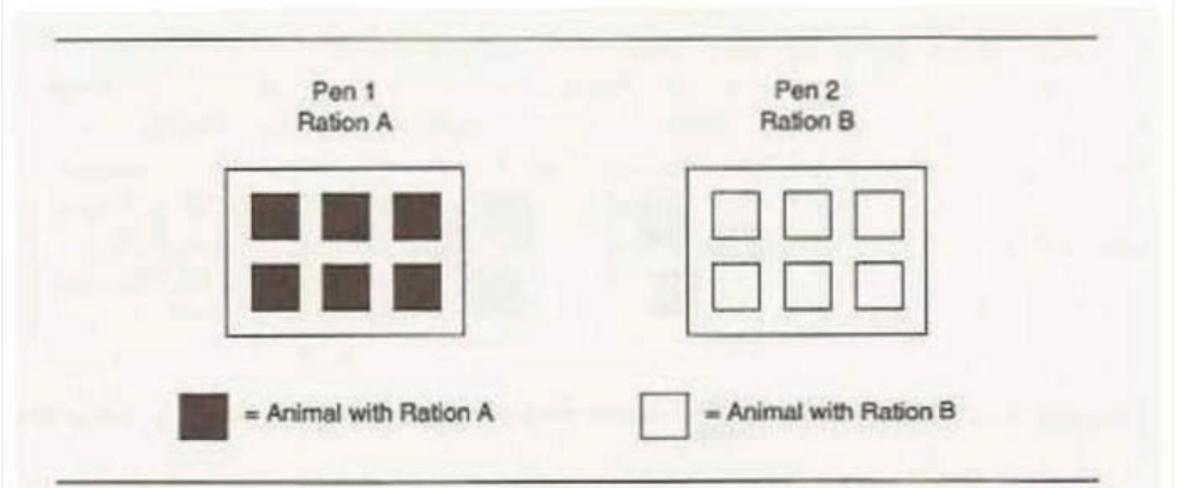


**Figure 1.4** Illustration of an unreplicated experiment

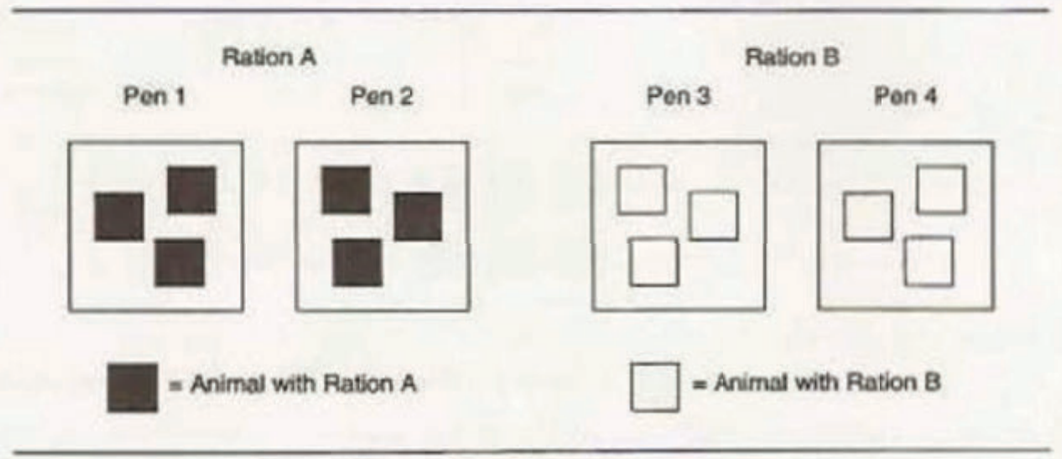


**Figure 1.5** Illustration of a replicated experiment

# Measurements of experimental outcomes

Finally, we need to measure experimental outcomes (on experimental units).

Examples: changes in symptoms, time to death, strength of a material, production of crops, and so on.

Ideally, the measurer should be “blind” to the treatment to avoid bias.

Multiple measurements can be made on each experimental unit. We can make multiple measurements over time (time-series responses).

# Multiple measurements are not he same as replicates

In experimental design, when we say replicates, we emphasize it is **the replicates of the treatments**. Later, we will explain that increase the number of replicates is crucial to improve the power to detect true difference between treatments: increase the number of measurements on each individual will usually not increase the power of a study.

Example: suppose we raise 10 fish in each of the two tanks with different temperatures and want to see how the temperature affects the growth of the fish. In this example, the intended treatment is the temperature. Like the ration example earlier, we do not have real replicates for the temperature treatment: each temperature is only applied on one experimental unit (the tank). Therefore, we will not be able to evaluate the uncertainty associated with the treatment. If we measure the body weights of the ten fish in each tank, they should be considered as multiple measurements on the same experimental unit, not replicates.

# Factors affecting the measurements

If there are factors that can affect the measurements of the experimental outcome, they should be considered together with other treatment factors in the design of the experiment,

For example, in a plant study, multiple technicians (or PhD students) may carry out the experiments and measure experimental outcome, then the different technicians/students should be considered as a factor in the experiment.

If we measure the gene expression using a sequencing machine, but the machine can only measure 8 samples in a batch. If we have more than 8 experimental units and have to measure the outcome in different batches, then the potential batch effects have to be considered in the design and analysis of the experiment.

# Summary

We introduced the five basic elements of experimental design:

0. A scientific question.
1. A set of treatments to compare.
   - We talked about control, placebo, and double-blind experiments
2. A set of experimental units.
   - Experiment units should be representative of a target population
   - Blocking can be used to increase the efficiency of a design
3. **Assignment of treatments to the experimental units.**
   - This is the key step in the experimental design
4. Measurements of experimental outcomes (on experimental units).
   - Multiple measurements are not the same as replicates.

In practice, we have to clarify these five elements when design an experiment.