Data Collection

Experiments

Experiments are the gold standard since they allow us to make causal conclusions. For example, they can answer questions such as

- "If I observe fewer infections in a group that wasn't vaccinated versus a group that was, can I conclude it was the vaccine that caused the lower infection rate, or could it be something else?"
- "Suppose we suspect that a new teaching method results in better learning outcomes. Can I definitely attribute the improvement to the new method?"

Some of the terms that are used in experiments include:

- **Response variable** (dependent variable): The outcome of interest, measured on each subject or entity participating in the study.
- **Explanatory variable** (predictor or independent variable): A variable we think might help to explain the value of the response variable.
- **Experiment**: The researcher manipulates the explanatory variables to see the effect on the response.
- **Factor**: a categorical explanatory variable.
- **Levels**: values of a factor.
- **Treatment**: a particular combination of values for the factors.
- **Experimental units**: smallest unit to which a treatment is applied.

Example 1
When there is only one factor, the treatments are the levels of the factor. It is also possible that an investigator might want to simultaneously manipulate two explanatory variables to see the effect on the response. Imagine a study comparing two drugs, call them drug A and B, each at two different doses: high and low.

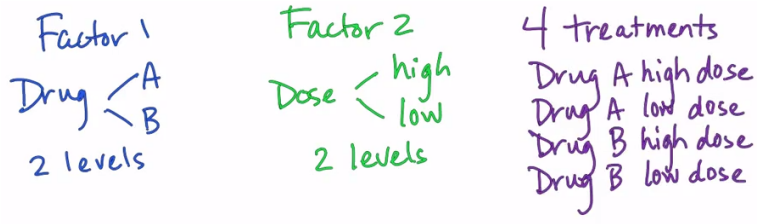

Figure 1: Factor-level combinations of Example 1

In this situation we have two factors each with two levels, but four different combinations and therefore four different possible treatments.

Treatments can be given to animate or inanimate objects, e.g., students of a class trying a new teaching method or plants in one pot receiving a new type of fertilizer. The experimental unit is then not the individual student or plant, but the whole class, or pot.

Extraneous factors are factors that are not of interest in the current study, but are thought to affect the response. They need to be controlled to rule out the possibility that they are causing any observed differences in the response. To control for an extraneous factor we could:

1. Hold it **constant** (we will explore this more below).
2. Use **blocking**: Blocks are groups of experimental units that are similar in the extraneous factor. All treatments are then randomly assigned to experimental units within each block.

Example 2
If we think a vaccine might work differently on males and females, we can make the decision to hold sex constant and study only females. This limits the generalizability of the study, but it also eliminates the potential of having the extraneous variable confound the results. If we wanted to include different age groups which might have differential response, we can treat age group as a block and randomly assign treatments within each age group. So in each age group we have subjects who both receive the vaccine and do not receive the vaccine.

If there are extraneous factors that cannot be controlled for, or that have not been identified, or that are unknown to the researcher, randomisation must be used. **Randomisation** is used to assign experimental units to treatment groups. By randomly assigning individuals to treatment groups, we can ensure that any differences that exist between the groups, in any possible extraneous variables, are just due to chance; the fact that we expect to see some differences due to this chance variation is part of our statistical model. After randomisation, when we average out this chance variation, the treatment groups are essentially the same.

Once we have eliminated other differences between the treatment groups, if the response

variable is different among the groups, the only explanation is the treatment; **cause and effect** conclusions can then be made.

In summary, there are three fundamental principles of experimental design:

1. **Control** the identified extraneous variables by blocking or holding them constant.
2. Use **randomisation** to randomly assigned experimental units to treatment groups.
3. Use **replication** by applying each treatment to more than one experimental unit

The word control also has another meaning in experiments. Experiments often have what we call a **control group**, which either does not receive a treatment or receives the current standard treatment. If the goal of an experiment is to show that a treatment affects the response, you need to have at least one other group for comparison since the simple act of studying an experimental unit may cause it to change. Studying at least two groups undergoing the same experiment allows us to compare them under the same circumstances. The control group is itself a treatment group, and if there is only one factor, control is considered a level of the factor.

In our use of the word replication, we are not talking about replicating an experiment to check whether a result found in one study also holds when you do another study. Having replicates allows the researcher to estimate variability in the measurement of the response which we can't do if we only have one observation for each treatment. It also ensures that the treatment groups are more comparable in values of the extraneous factors, by having the opportunity to have different values of the extraneous factors within each treatment group.

If the only way to establish a causal relationship is to carry out a randomised, controlled experiment, why would anyone carry out a study that is not an experiment? The answer is simple: it is not always possible for ethical or practical reasons. If we want to study the effect of smoking on human mortality we can't randomly assign some people to smoke a pack a day and some people not to. That would be unethical!

Example 3
A good example of an experiment is a story that is slightly adapted from a project presented by two Montreal students at the 2011 Canada Wide Science Fair. Their project was the winner of the Statistical Society of Canada Award for excellence in the use of statistical methodology, which they won for their excellent example of a well designed experiment.

The students were interested in ischemic preconditioning, which is a technique to create resistance to loss of oxygen through loss of blood supply to tissues. Ischemic preconditioning works by applying brief episodes of restricted blood flow in order to protect against damage from a subsequent longer term episode. Their research question was: "Can ischemic preconditioning improve athletic performance"

Study design:

- 2 factors:
    1. amount of pressure applied (using a blood pressure cuff), 2 levels:
        - 20 pounds
        - 0 pounds
    2. length of time pressure, 2 levels:
        - 10 minutes
        - 20 minutes
- Treatments: 4 pressure x time combinations
- Replication: each treatment was applied to 10 teenage males.
- Experimental units: 40 participating teenagers
- Response variable: length of time a wall squat position can be held

Note the use of a control group here. The students needed to know if ischemic conditioning worked so they gave a sham treatment of 0 pounds to some of their experimental units, some of whom got zero pounds for ten minutes and some for 20 minutes. Moreover, to control for extraneous factors, male teenagers of similar athletic ability were chosen because they participate in sports at school. Although this was not done in this study, another way they could have controlled for extraneous factors was to use blocking. For example, they could have enrolled both athletes and non-athletes in the study and assigned the four treatments to subjects within each of the athletic groups. Another principle of experimental design is randomisation, and the 40 experimental units were randomly assigned to which of the four treatments they received.

This ischemic preconditioning experiment also illustrates some other characteristics of excellent design which should be used when appropriate and possible. The students used single-blinding in their study. **Blinding**, if it can be used in an experiment, reduces the potential for bias since people don't know if a treatment is in place or not:

- Experimental units are blinded if they do not know which treatment they received.
- The person measuring the response is blinded if he/she does not know which treatment was given to which experimental unit.
- Experiments can be **single-blind** (if the subject is blinded) or **double-blind** (if both the subject and experimenter are blinded).

When people show change when participating in an experiment, whether or not they receive a treatment, it is known as the **placebo effect**. For this reason, the control group is often a **placebo** which is something that is identical to the treatment received by the treatment

groups, except that it contains no active ingredients. In our ischemic preconditioning experiment, the placebo was the application of zero pounds pressure to some of the experimental units.

Randomisation plays two key roles in data collection. The table in Figure **??** summarizes those two roles:

		Assignment of treatments	
		Random	Not random
Selection of experimental units	Random	Can make causal conclusions that can be generalised to the population	Cannot make causal conclusions but results can be generalised to the population
	Not random	Can make causal conclusions about only the participating experimental units	Cannot interpret observed relationships as causal and cannot generalise beyond the participating experimental units

Figure 2: Randomisation in Experiments

By using random sampling, we get a representative sample and ensure that we can generalize our results to a larger group or population, free of any selection bias. Later on in our inferential procedures, we will use a statistical model that accounts for the random variation that causes our sample to differ from other samples that we could have possibly chosen in another random selection.

Randomly assigning experimental units to the treatments they receive eliminates the effects of extraneous factors, allowing us to make causal conclusions. While the extraneous factors may vary from treatment group to treatment group somewhat, randomisation ensures that these differences are due to chance variation only, and not any systematic difference that could confound our interpretation of the effect of the treatment. So if we want to make causal conclusions that we can generalize to an entire population, we need to randomize both in our selection of our experimental units and in assigning treatments to experimental units.