Biology 116 Lab Manual

Dr. Tristyn Hay and Dr. Robin Young

2021-08-16

Contents

Welcome

Welcome to the BIOL116 lab manual!

The material in this manual is designed to be used as part of the BIOL116 labs. It integrates heavily with additional material that can be accessed only through the BIOL116 Canvas page, including assignments, quizzes, and other material that is specific to this year's offering of the course. As such you will be expected to access BOTH this website and Canvas regularly.

Your first lab, Lab 1, is asynchronous (meaning that you complete it online, in your own time during that week) and the material is posted exclusively on Canvas. **Your second lab**, Lab 2, will be your first chance to physically enter the lab, and meet your TA and your fellow lab mates. The material for Lab 2 can be found in this text, under the Lab 2A and Lab2B headings (with links to this content also on Canvas).

The rest of the labs will continue to alternate between asynchronous, online labs, and synchronous, in-person labs. **For in-person labs it is vital that you attend the lab in which you are registered** (Check your registration online to be sure of where you are supposed to be!).

There's much more information about what to expect from labs, the grading policies, and the weekly schedule in the Canvas Module for Lab 1. For now, we will say that we're looking forward to the start of term, and to meeting you (both virtually and in-person). Science is amazing and exciting and cool, and we can't wait to share our love of science and biology with you!

Copyright

Please use the following for citing this document

Hay, T., Young, R. (2021). *Biology 116 Lab Manual*. https://ubco-biology.github.io/BIOL-116-Lab-Manual/

All source files are available from https://github.com/ubco-biology/BIOL-116-Lab-Manual.

Lab 1: Introduction

Last updated 2021-08-16

Content for the first lab will all be on Canvas, the learning management system used at UBC. Enrolled students can access this content at https://canvas.ubc.ca/courses/90147

Lab 2A: The Process of Science

Last updated 2021-08-16

Note on Copyright

Images and some lessons used in this lab have been adapted from www.understandingscience.org and used with permission from Deb Farkas, Stan Hitomi, and Judy Scotchmoor.

Understanding Science is a National Science Foundation grant funded project produced by the UC Museum of Paleontology of the University of California at Berkeley.

Introduction to the Process of Science

Objectives

At the end of this lab students should be able to:

- List the main components involved in the ”process of science”
- Provide examples of scientific activities and / or list where these activities would generally fit in the ”process of science” chart.
- Explain - in your own words or through examples - what is meant by each of the following statements regarding the process of science:
  - The process of science involves testing ideas with data
  - Though they can be supported by evidence, it is often difficult for scientific hypotheses to be proven or directly observed
  - Scientific knowledge is subject to change as new evidence and perspectives emerge
  - The process of science is often non-linear
  - The process of science involves observation, exploration, discovery, testing, communication, and application

- Scientists can use multiple methods - experiments, observations, models - to test their ideas
- Science involves curiosity and creativity
- Scientists work together to share their ideas
- Scientists aim for their studies to be replicable
- Over time, scientific ideas or hypotheses that are supported by multiple methods and replicated in various studies become established theories that are widely accepted by the scientific community

Discussion Questions

1. What is science?
2. What do scientists do and how - in general - do they do it? Discuss this with a partner.

Mystery Cylinders

Mystery Cylinders An exercise in the scientific process

Imagine a container, a cylinder that you are unable to open, and you cannot see inside. How would you manage to figure out what is inside the cylinder? What would your thinking process be?

This activity is broken down into 3 steps:

1. Observe
2. Share and Collaborate
3. Submit

Phase 1 – "Observe" your cylinder

Outside of opening the container, you have the option to hold it, shake it etc. What types of observations would you like to make? During this activity, you and your group will be able to interact with the cylinder. Think about the question you are addressing and how you would go about answering it.

Work in groups to come up with three suggestions of ways you can manipulate your cylinders to try and determine what is inside. Remember that not all of your initial ideas will be plausible - you don't always have all the ideal tools available to you. You do not have an X-ray machine in the lab . Offer three possible tests that could likely be done in lab.

Figure 1: Figure 1. Mystery cylinders.Image by UBCO, licensed under CC BY 4.0

Designate one person in your group to write down all observations and information you attain during your experiment.Following this, you and your group now must try to reach a consensus about what is inside the interior of each cylinder.

Note If you do not take extensive notes throughout the experimentation process - electronically or on paper - you will not have the information that you need to complete this assignment.

Phase 2 - Share & Collaborate

Your instructor will then help your group to discuss findings with other groups.

- Discuss whether you are using similar techniques to analyze your cylinders.
- Collaborate with others to see if you can arrive at a consensus guess for the interior of your cylinder(s).

Phase 3 – Submit for marks

Once you feel confident with your predictions, you will need to submit them to Canvas for marks. There is a separate file posted with the assignment instructions.

Discussion Question

1. With your original group, make a list of the **specific things** that you did during this activity and how / why they are a part of doing science - e.g. twisting the cylinders = ??

NOTE Wait for your instructor before proceeding to the next section *Science Flowchart*.

Science Flowchart

A Simple Flowchart

Study the simple science flowchart with your instructor:

A Complex Flowchart

Now let's view the complex science flowchart:

Class Discussion

Follow your instructor's directions as a class to place examples of the types of things you did in the mystery box activity into each circle. See the list generated for the previous discussion question.

This discussion marks the end of the first part of this lab (Lab 2). If you have not already done so, make sure you have the information you need to complete your submission of the Mystery cylinder assignment before the end of the day. Once you leave your lab you will not have the opportunity to come back if there's something you forgot or missed.

When you're ready, you can proceed to Part 2 of Lab 2, where the research project is discussed.

Mystery Cylinder Assignment

Purpose To help you learn how to keep good lab notes and make predictions.

Rationale A key piece of science is to record the work you do so that you can learn from what worked and what didn't. Keeping accurate records of the work that you do can make all of the difference. Patents and other scientific

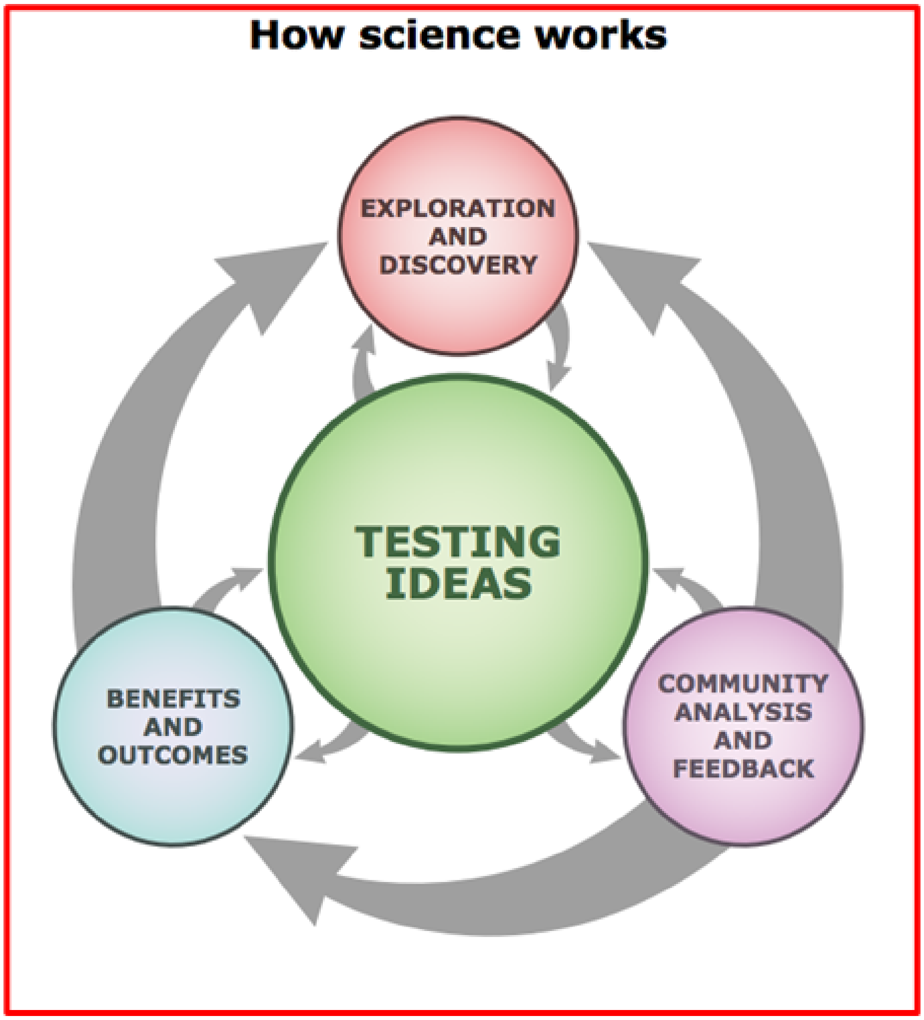

Figure 2: Figure 2. Simple science flowchart.Image used with permission from The University of California Museum of Paleontology, Berkeley, and the Regents of the University of California.

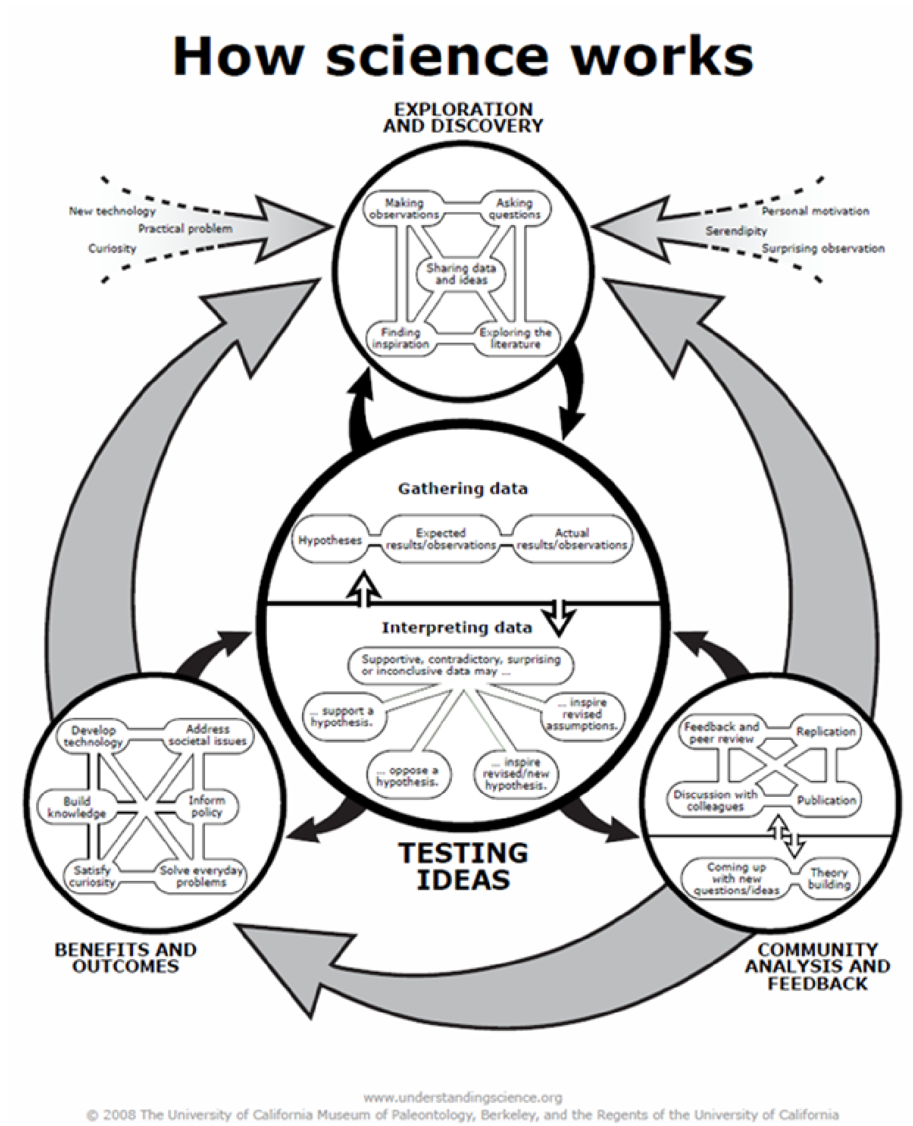

Figure 3: Figure 3. Complex science flowchart.Image used with permission from The University of California Museum of Paleontology, Berkeley, and the Regents of the University of California.

discoveries can be worth a lot of money to those that can prove that the work they did was theirs. This requires accurate and honest note keeping.

Instructions For this assignment, you will be working with your teaching assistants to determine the contents of a ”mystery cylinder”. You will have the opportunity to see, touch, and manipulate each cylinder in the lab and you are expected to work within your lab group to come up with 3 experiments that your group will conduct. Once you have completed the experiments, you will be expected to make a prediction regarding the contents of your ”mystery cylinder”.

If you do not take extensive notes in your lab notebook throughout the experimentation process, you will not have the information that you need to complete this assignment.

Submission For this assignment, you will submit your predictions for the contents of your mystery cylinder, along with a rationale that explains how the experiments you ran, which were documented in your lab notes, support your prediction.

You have two options for submission of this part of your work:

1. You can submit a 1-page written prediction, with rationale. It should be no more than one page (double-spaced, Arial 12pt font, 1cm margins)
2. You can submit a video (no more than 2 min)

You will upload your mystery cylinder assignment on Canvas via the dedicated assignment entitled ”Mystery Cylinder Submission”. This assignment is **due before 11:59 pm the day of your lab**.

While it is encouraged that you work together prior to submission, your lab submission must be entirely your own work.

This means:

- No copying and pasting from other sources - even if you plan to ’tweak’. This is not your own words.
- Do not work it out together, and then alter the final draft to make it look less similar. Figure it out together, and then go away and write your final submission yourself.

Mystery Cylinder Assignment Grading Rubric

Criteria

Ratings

Pts

Predictions - All prediction(s) are present - Prediction(s) are clear and consise - Prediction(s) are reasonable

5pts – Full Marks: All criteria are met

3pts – Proficient: Two of the three criteria are met

2pts – Unsatisfactory: One fo the criteria are met

0pts – No Marks: No criteria are met

5 pts

Rationale - A rationale is provided for the prediction(s) - Each rationale explains how the experiments conducted supported each prediction - The rationale(s) provided are clear and easily understood

5pts – Full Marks: All criteria are met

3pts – Proficient: Two of the three criteria are met

2pts – Unsatisfactory: One fo the criteria are met

0pts – No Marks: No criteria are met

5 pts

Total points: 10

Lab 2B: Research Project

Last updated 2021-08-16

Over the term, you and your partner will be designing and conducting an experiment, analysing the data you generate and then presenting your results. In order to prepare you for this task, you will need to have read through the material below in advance of lab.

Learning Outcomes

Students will participate in the process of science and demonstrate scientific thinking. Students will be introduced to the characteristics of the experimental method of research and learn:

- How to form a hypothesis.
- How to design an experiment to test your hypothesis.
- How to analyse and display your results.
- How to interpret your results so as to support or reject your hypothesis.
- How to present your results as an oral presentation.

In addition, you will begin to explore the principles and practices of Open Science, as a way to help ensure reproducibility and transparency in the work that you do as scientists.

Science & The Scientific Method

Science is just one of many different lenses through which we, as humans, attempt to better understand the world around us. Science as a discipline is built on the foundation of observable, measurable evidence that allows us to make testable predictions about the world around us. There are other ways of learning about our world - faith, intuition, memory, imagination, etc. Each of these "ways of knowing" uses different kinds of evidence to help us make sense of our world. Science is also a highly collaborative endeavor, where each scientist builds on the work of their colleagues and mentors, to collectively construct scientific knowledge. This collaborative aspect is central to the ideas of Open Science, which is a key part of what you'll be learning about throughout your degree at UBCO.

The Scientific Method

The way we do science is to follow a set of guidelines called "The Scientific Method". This method helps us make sure that we do our best to test our ideas in such a way that we've done our best to remove our own opinions and personal biases from the process. Removing personal opinions and biases, and focusing on what the evidence is telling us is absolutely vital to science. It is the heart of everything that we, as scientists try to do every day.

Deductive Logic

Some scientists are interested in measuring and recording observations in nature. This is sometimes called "descriptive science". Scientific sampling and collection methods allow researchers to describe nature accurately. We can then use these observations to make generalizations about the world using "deductive logic".

Inductive Logic

Other scientists wonder why things are the way they are. They form ideas or hypotheses using "inductive logic" about how things must work and then test these ideas in experiments that they design.

Whether a scientist is conducting descriptive science or inductive, experimental science, they will be using evidence to support the claims that they make, and they will be using the framework of the scientific method to help them do this in a way that avoids bias and focusses on the evidence they have. **In this Research Project we will be focusing on experimental research, rather than descriptive science.**

Reproducibility

One of the most important concepts in science is the idea of **reproducibility.** This means that if our methods are carefully recorded, and shared with others, they will be able to reproduce the work that we did, and get the same results. It also means that scientists never base their conclusions on only one experiment. We always attempt to **replicate** our results, by running the same experiment, following our proposed method, at least 2-3 more times - and sometimes more! - to ensure that the results we are getting are consistent, and not some kind of random fluke. Replication is very much about keeping records and adhering to standards and conventions when possible. We've already talked about the importance of note taking in the "Mystery Container" activity and we'll continue to emphasize this; keeping track of your experiment with notes is foundational to science.

Understand that scientists do these things because they do not accept anything on faith. Science is not a belief system; scientists are convinced only by evidence and data. Any idea is up for debate and everything can be criticized. It takes many years and many experiments to convince scientists that something is true or not true. For example, the idea that the continental plates are moving slowly over the surface of the Earth took decades to take root. It was debated and tested and tested again by many people before it was fully accepted as fact. This is as it should be... scientists don't much like to be wrong about the big stuff. We don't accept any explanation as "true" before every other possible explanation has been tested and rejected. It is only when we have done our best to disprove an idea, and it has stood up to everything that we've thrown at it, that we then begin to accept that maybe... just maybe... it might be true... at least for now, until new evidence is found that throws everything into doubt again.

Open Science

Open Science is a movement to make scientific research transparent and accessible to everyone. This not only gives the best opportunity for research to be critically examined and ensure reproducibility, it also makes it easier for scientists to share their work with others and build on the work that has been done before. You will be learning more about the principles and practices of Open Science in Lab 3. For now, know that as you work on your research project, you will follow the stages of a typical registered report and implement Open Science practices including:

- Throughout the experiment, using appropriate **version control** on electronic documents and proper file and data management practices.
- Performing a literature review on your research topic and documenting a list of consulted studies, how they were found, and the strengths, limitations, and weaknesses of each.
- Submitting a written proposal with an established **a priori hypothesis** and experiment design outlined for your experiment. This will be marked before the experiment implementation phase and TA feedback incorporated into the project as needed.
- Implementing the study according to your plan, and noting any deviations from that plan. These reflections will be submitted for marks.
- Submitting and presenting a poster that details your experiences implementing the research plan - including any changes recorded, justification for changes, analysis of the data, and your interpretation and conclusion.

- Conducting a peer review of other students' poster presentations using the poster presentation rubric as a guideline.

Experimental Research

Now that we've briefly introduced you to Open Science, and what it means in the context of this project, let's talk more about what you're going to be doing. In a nutshell, you are going to use the scientific method to learn something about an organism, in the controlled setting of a lab experiment.

We expect that you have been exposed to the principles of the Scientific Method before this, while you were still in high school. As such, we are assuming that you have a little bit of prior knowledge to draw from. The diagram below shows an overview of the steps involved, and in the following sections, we will discuss each step of the scientific method, and how it applies in the context of this research project.

Observations & Questions

Observation of Phenomena & Formation of a Question

The first step in any research is to decide on a "problem" or focus for the investigation. This often happens naturally when you observe something that interests you, but that you can't entirely explain. In your head, a question forms, such as "why did it do that?"

Deciding on a problem to explore and formulating a question is your very first task.

As all good scientists, the question you ask should be informed by existing knowledge and relevant research. A review of the published scientific literature is a good way to do this.

For this research project you will be using mealworms to conduct your study. Your research question must include the effect of "x" on some physiological parameter such as survival, some aspect of behaviour, some aspect of vision or colour vision, or some aspect of hearing. Your TA will be available to help you consider your options as you work on this.

Statement of Hypothesis

A hypothesis is an unproven explanation for the observed phenomena. In its simplest form, a hypothesis is an "educated guess" or intuitive hunch that is

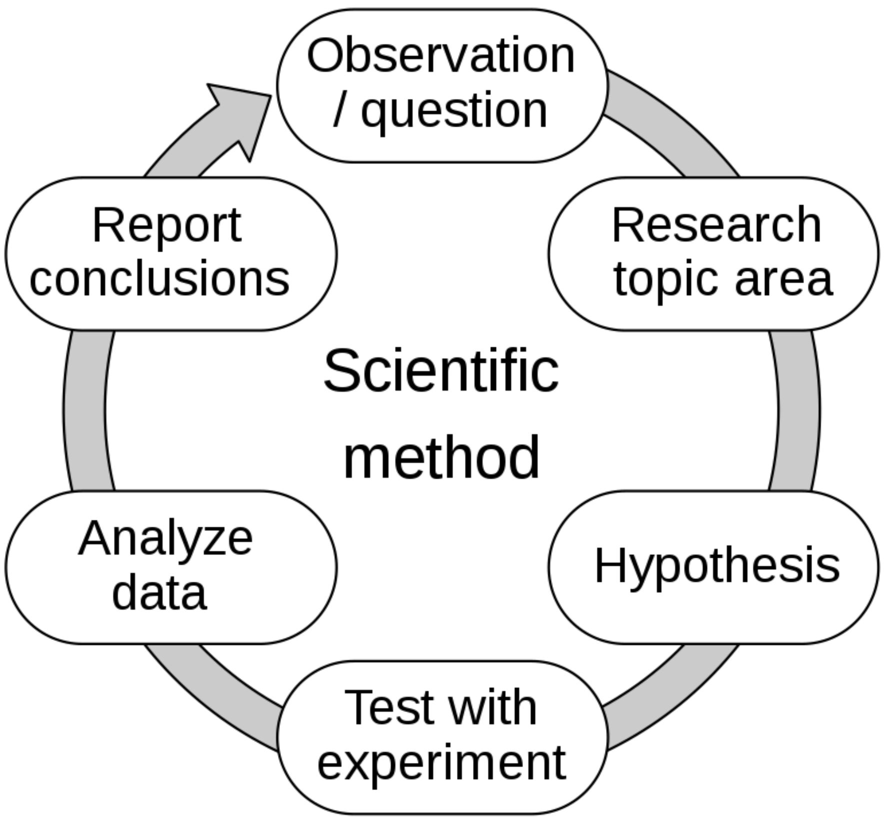

Figure 4: Figure 4. Visual representation of the steps of the Scientific Method.Image by Efbrazil, licensed under CC BY 4.0

proposed as a possible answer to the question you're interested in answering. There's a couple of things to know about hypothesis building before you get started:

A hypothesis is not a question, it is a statement

For example, "plants will grow faster at higher temperatures" is a hypothesis, whereas "will plants grow faster at a higher temperature?" is a question. They're generally related, but they're not the same.

A hypothesis must be testable

The hypothesis does not need to be "correct" - after all, there's really no way to know that at this point - but you do have to be able to test whether it is correct or not. In our example from above, we can test the hypothesis, by growing the plants at different temperatures, and observing the effect.

A testable hypothesis is a measurable hypothesis

Since your hypothesis must be testable, you must also have a way to **measure** the effect. In our example, we can measure the effect of temperature on growth by collecting information on plant height, or number of leaves, or some other parameter, when the plant is grown at the different temperatures.

The hypothesis comes before the experiment, and not the other way around.

We call this an a priori hypothesis, meaning that we made the hypothesis before we ran the experiment and learned the answer. Sometimes, because we want to show that we knew what we were doing, we feel the need to change the hypothesis we started with, so that it better reflects the results we got. This is known as **HARKing** (Hypothesizing After Results are Known), and is not considered to be good science. It's important to present your hypothesis as you originally developed it, and then discuss what you have learned about the topic based on your testing of that hypothesis.

Null & Alternate Hypotheses

For experimental research, we often write our hypothesis in the form of a **null hypothesis (Ho)** which states that the factor being tested will have no effect on the outcome - e.g., there will be no difference in the growth rate of plants at different temperatures - and an **alternate hypothesis (Ha)** which states that the factor being tested will have an effect on the outcome.

The data that you generate when you measure the differences will need to be quantitative - i.e.numerical - and analysed by some statistical test which will ultimately help you decide which of your two hypotheses are supported by the experiment that you conducted.

Remember your "hypothesis being supported" is not the same thing as either hypothesis being true. There are many factors that influence how robust your experiment is, and how likely it is to reflect the "truth". We call this the **power** of your study. Your methods, sample size, and the number of times you repeat the experiment and get the same result all play into the power of the work that you do, and how likely your results represent the underlying laws of nature.

Design & Test

Designing your experiment to test your hypothesis

Once your null and alternate hypotheses have been properly stated, you then design an experiment that will help you collect the data you need to decide which of your two hypotheses are supported. Ideally, experiments should be conducted in such a way that the experimenter has control over every **variable** that might have an influence on your results. In reality this is much harder than it sounds!

Variables

A variable is any factor that might affect the outcome of the experiment. The experimenter therefore manipulates the **independent variable** and observes the effects of this manipulation on the **dependent** - or **response** - **variable**. For example, if the goal is to determine the effects of temperature on the growth rate of a plant, then growth rate is the dependent variable because it "depends" on the temperature to which the plant is exposed. All other variable must be controlled or held constant, to the extent that you are able to. Temperature is the independent variable because that is the variable which is manipulated by the experimenter.The ideal way to perform such an experiment is to set up and run a series of tests that are all identical in all other ways - light, soil moisture, etc - except for the one specific factor that is being tested, in this case, temperature.

Control

One specimen - or group of specimens - is used as a **control** against which all other experimental samples are compared. **The importance of the control group cannot be over emphasized.** It is essential to know how the system you are investigating works under normal circumstances - i.e., before you started messing with it - before you can be sure the results obtained from the experimentation are actually due to the manipulation of the independent variable(s).

To continue our example above, if you wanted to investigate the effects of temperature - the independent variable - on the growth rate of a plant - the dependent variable - you would gather data on the growth rate of plants at their normal, expected temperature - most likely room temperature - as the control group, and then compare the data collected from this group to the growth rates of plants exposed to higher and / or lower temperatures, depending on what you're hoping to learn.The control group would provide the ”normal standard” against which all other samples would be compared.

Sample Size

Another important rule governing experimentation is that the experiments must be done on a large number of test subjects. The larger the ”sample size” or the more replicates you include, the more confident you would be in your results and the more power your study has. Every treatment in your experiment - including the control - must have replicates.

Sometimes very large sample sizes are just not possible. In our example above if you were growing your plants in the front window, you may have room for only five or six. However, if you had a greenhouse, you could grow and measure hundreds of plants. When it comes to replicates, a good rule of thumb to start from is once is luck, twice is coincidence, three times is science - and five or more times is good science!

Random Assignment

Biological variation is the inherent differences among organisms in a sample because of differences in genetic makeup, age, sex, health, etc. Biological variation is impossible to avoid entirely, but it can be minimized sometimes by limiting the characteristics of the group of organisms you are studying (e.g., by using organisms of the same age, sex, etc.). Even when potential sources of variation are minimized, it is crucial that subjects be randomly assigned to the treatment groups, so that any inherent variation among individuals will be distributed at random among all treatments, including the control group.

Study Replication

Lastly, the experiment should be **repeatable** by anyone that wishes to do so - including you! - should they wish to test your ideas for themselves. This requires careful documentation of the work that you have done at every step of the process - methods, data collection, analysis, conclusions. If your experiment is repeated and it doesn't have the same results, then maybe some details of the experiment were not sufficiently well described, or maybe the treatment effects are simply not strong enough to be repeatable given the experimental design. We said we'd emphasize note taking!

Research Data Management

Your group is expected to follow proper file and data management practices. Having your files and data organized appropriately will save you time as well as increase the reproducibility of your research.

NEVER assume that you will remember exactly what you did days, weeks or months prior… you won't, and your science - as well as your lab grade - will suffer. So being organized and meticulous is absolutely vital to the success of your project.

Considering how you are going to keep track of everything is one of the FIRST decisions that you need to make as a group. If you don't write things down right from the moment you start developing your experimental plan, how will you be able to run the experiment the exact same way twice?

Things to Consider

- Are you going to use paper or electronic files?
- How will you share methods and data with each other?
- How are you going to keep track of the changes you make over time, in case you change your plan and need to back up to an earlier version of your experimental plan?
- Once you know the data you want to collect, how are you going to record it?
- How should that file be organized to make it as easy as possible to accurately record the data your experiment generates?
- How will you make sure you know which test subject/ trial each data point belongs to at all times?
- Now that you have the results, what are the different statistical tests that you tried? What did they tell you about your data and whether your Ho is supported or rejected?
- Are the trials the same? What are your thoughts about this? Writing down your brainstorming ideas is also extremely useful.

File Naming

In **Lab 5** - online, asynchronous lab - you will be exploring best practices for file management and naming conventions. For now, here is an example of naming a data file for an experiment investigating how mealworm movement is affected by the presence of light.

`20200626_Mealworm-project_Light-movement-data.csv`

Here is an example of naming a written proposal for the same experiment:

`20200724_Mealworm-project_Proposal_V01.docx`

Visual Methods

As you design your experiment, and develop - and troubleshoot! - the method you expect to use, it can be very helpful to draw out your proposed methodology into a flow chart. This can help you better visualize what you want to do, and may help you realize where you need to think a little more critically about your proposed plan and how it will work

Note visual representations of your methods work really well on posters! So, if you start working on it right form the start, it should be awesome by the time you're ready to build your poster!

Data Analyses & Visualization

The data you collect from the experiment, is generally called the **raw data**. Once we have the raw data, we must explore it, in order to understand whether the data supports or refutes our null hypothesis.

For quantitative data, like we are going to collect in this experiment, this usually means that averaging and / or some other statistical treatment will be applied, which helps you determine which parts of your result indicate normal biological variation, and what might be hinting at the effect of your independent variable on the dependant one. Often graphs are used to display the data to others, as it helps them to quickly see the trends that you observed.

To return to our example above, the raw data may simply be the heights of the plants you measured each day, recorded in `cm`. The final data on growth rate are then calculated and expressed as the average increase in height per day, of all the plants measured, in each temperature group. You would likely also do a statistical test to see if any variation that you are observing is "normal" or "significant".

Note It is virtually impossible to do experimental science without doing some kind of quantitative statistical analysis. However, BIOL 202 is the course designed to teach you the complexities of biological statistics. Here we are giving you only enough information so that you will be able to complete your BIOL 116 research project. We recommend that you take BIOL 202 as early in your degree as possible, as it will make all of your lab work easier!

Central Tendency

In order to interpret the results of your BIOL 116 research experiment, there are really only two bits of information you need. The first is some estimate of the

”central tendency” in the response. That is, what was the ”average” response of all the organisms measured?

In our example, if you had ten plants in one temperature group, say 10 ° C, you could calculate the **arithmetic average** - or the **mean** - growth rate by adding up the growth rates of all ten plants in that group and dividing the total by ten.

There are other estimates of the central tendency that may, in some cases, be more informative. One is the **median**. This is the value that divides the number of observations into two equal groups. The other estimate of central tendency is the **mode**. It is used most often in business statistics, rather than biological experimental research, and is the observed value that occurs most often.

Variation

The other bit of information you will need in order to interpret your results is some estimate of the **variation of the response within the group.**

For example, say the mean growth rate of our plants at 25 ° C was 3.0 cm per day. It is important to know whether

a) six of the plants grew at a rate of 3.0 cm per day, and two grew at a rate of 2.0 cm per day, and two grew at a rate of 4.0 cm per day, in which case there was very little difference - variation - in their responses to that temperature, or

b) whether two of the plants grew at a rate of 11.0 cm per day, and the other eight grew at a rate of only 1.0 cm per day, in which case there is a very great difference in the responses of these plants to that temperature.

Due to the nature of biological organisms, no two will respond in exactly the same way to any given independent variable. It is important then to have some estimate of this variation in response. There are several ways to estimate this variation. One is to simply state the **range** of the responses. The range is the difference between the highest and lowest values obtained.

There are also much more mathematically involved estimates of variation in response, such as **standard deviation**. One should not be intimidated by the mathematics of the calculation of standard deviation, but simply realize that it’s just another estimate of variation in response.

Continuous, Discrete & Choice Data

Your experimental design determines what type of data you will collect, which then determines the appropriate method for analysing the data. Data that are continuous are analysed with measures of central tendency and variation - the next section explores these in more detail.

- Examples of **continuous data** are temperature, height, weight, time, distance, etc as generated by **measured-response experiments** such as the plant experiment described above.
- **Discrete data** are data that can only take on certain values, such as the number of leaves on a plant, or the number of students in each lab section. These are recorded as frequencies of observations and are analysed using a Chi-squared (X2) test.
- **Choice experiments** where the test subject is given a choice between a set number of discrete options - if mice were presented with 2-3 different bedding materials, to see which type they prefer - are usually treated the same way that other discrete data is treated, and Chi-squared tests are used.

For your experiment, the data you collect will depend on the question you're exploring, and the hypothesis that you're testing. As such, the way that you will need to analyse and present the data will likely be different than your classmates, since their hypothesis will be different from yours.

Measured-response experiments

standard deviation and 95% confidence intervals

Arithmetic mean

After calculating the mean of each treatment, you can compare the means to see which, if any, treatments result in a significant difference in response. The first step in comparing means is to determine how close all the data points in each data set are to each other. Two statistics are used to describe the variation in a data set; variance and standard deviation.

Variance

Variance (s2) is a measure of how much the data values scatter around the mean, on average. A small variance indicates the data are tightly clustered around the mean. The larger the variance, the more scattered the data. Variance is calculated by summing all the squared deviations from the mean (a deviation is the difference between an individual measurement and the mean) and dividing this sum by the number of data entries minus one.

Calculating Variance

$$s^2 >= \frac{\sum (x_i - x)^2}{n - 1}$$

Based on the UBCO Biology Guidelines for data presentation, variance values should be reported to one more decimal place than the number of decimal places that the data entries contain.

Standard Deviation

Standard deviation (s) is a related statistic that is the square root of the variance. Standard deviation can be determined on calculators or in Excel - or as you'll learn, using the R programming language) but it is still important that you understand how it is calculated and what it represents. The formula for calculating standard deviation (s) is shown below.

Calculating Standard Deviation

$$s = \sqrt{(\frac{1}{n-1})\sum_{i=1}^{n}(x_i - \overline{x})^2}$$

Let's calculate the standard deviations of the two samples of our plant growth experiment described above with the same average daily growth rates of 3cm/day.

Complete the table below to find the standard deviations for each sample population.

Plants

Sample population 1

Sample population 2

Growth rate in 1 day (cm)

$x - \overline{x}$

$(x - \overline{x})^2$

Growth rate in 1 day (cm)

$x - \overline{x}$

$(x - \overline{x})^2$

1

3

0

0

11

8

64

2

3

0

0

11

8

64

3

3

0

0

1

-2

4

4

3

0

0

1

-2

4

5

3

0

0

1

-2

4

6

3

0

0

1

-2

4

7

2

-1

1

1

-2

4

8

2

-1

1

1

-2

4

9

4

1

1

1

-2

4

10

4

1

1

1

-2

4

Mean growth rate

3 cm/day

3 cm/day

Sum = 4

Sum = 160

For Sample population 1, the sum of $(x-x^2) = 4$; whereas for sample population 2, the sum of $(x - x^2) = 160$. For both populations, $n - 1 = 9$.

Therefore, for sample population 1;

$$s = \sqrt{4/9} = 0.66$$

and for sample population 2;

$$s = \sqrt{160/9} = 4.22$$

What does this tell you? Much like the measure of variance, a low standard deviation indicates that the data points tend to be very close to the mean, whereas a high standard deviation indicates that the data points are spread out over a large range of values. Both variance and standard deviation can tell us important information about the data we have, and what is ”normal” for our test subjects versus what might be the effect of the independent variable.

95% Confidence Interval

The mean and standard deviation are statistics that describe the data points in a sample. As such, they are estimates of the true mean and standard deviation of the entire population we're interested in. We can then use these estimates to calculate an interval or range of values within which the mean of the entire population - the true mean - will most likely fall. The size of this range indicates the confidence with which we can say the true mean will fall within these limits.

By calculating a 95% confidence interval, we are saying that we are 95% confident that the true mean falls within this range; this is the level of confidence biologists most commonly report.

Calculating 95% Confidence Intervals

The formula for calculating the 95% confidence interval (CI) of the mean is:

$$CI = \overline{x} \pm 1.96 \frac{s}{\sqrt{n}}$$

Where:

- $\overline{x}$ is the mean,
- s is the standard deviation; and
- n is the sample size

Note the 1.96 in the equation represents what's known as the Z-value, or Z-score. It's a little bit beyond the scope of this course to go into it, but it's related to standard deviation and the mean of your population, and for a 95% confidence interval, it's pretty much always the same, so we're going to ask you to trust us for now.

95% confidence intervals allow us to determine, with a 5% chance of error, whether the differences between two or more samples are due to random chance or sampling error, or due to some real difference in the characteristics of the samples and therefore the entire population of measurements. Most biologists use this 5% probability level to determine if there are significant differences in the means from different treatment groups in their experiments. This roughly translates to an acceptable risk of 5% or 5 times in 100 of drawing the wrong conclusion about the population from the sample data.

Interpreting 95% Confidence Intervals

If the 95% confidence intervals of any sample means do not overlap, we can be 95% confident that the samples are not from the same population of measurements. Therefore, we can reject the null hypothesis that there is no significant difference between these samples and, at the same time, provide support for our alternate hypothesis.

If the 95% confidence intervals of all sample means **are seen to overlap**, then we are less than 95% confident that the samples are from different populations of measurements. In this case, we cannot reject the null hypothesis. When the 95% confidence intervals overlap, discrepancies between sample means can be attributed to random chance, error and/or normal biological variation rather than the effect of the independent variable.

Discrete Data & Chi-Squared Tests

Graphing your data

When analyzing, and also when presenting your data to others, it is very useful to display the data or results in the form of a graph.

Line Graphs

Simple **line graphs** allow the relationships - trends or patterns - within the data to be visualized by the reader at a glance. Line graphs are often used to display **continuous data**, or measurements made on the same subjects over time.

In our example, the mean heights of the plants growing at 25 ° C could be recorded each day and plotted as a line graph (Figure 2). From this type of graph, it is easy to tell whether the plants grew at the same rate throughout the course of the experiment, or whether they grew more quickly when they were young seedlings.

Line graphs are generally best for data attained from measured-response experiments.

Bar Graphs

A **bar graph**, seen in Figure 3, is useful if you wish to display **discrete data**, such as the number of leaves present on our plants with respect to temperature. Bar graphs are generally best for data attained from choice experiments.

Regardless of the type of graph used, the independent variable - temperature in this example - is always placed on the horizontal, or X, axis and the dependent variable - growth rate here - on the vertical, or Y, axis.

All figures must have a number - and be numbered sequentially - and contain a detailed title that are placed below the figure. See Procedures and Guidelines: Figures and Tables for further guidance.

Interpreting Results

This final step in the scientific method requires that you provide a straightforward description of the conclusions drawn from the data you collected and analysed. For example:

- Do these sample means represent different populations of measurement (i.e., did you reject Ho)?

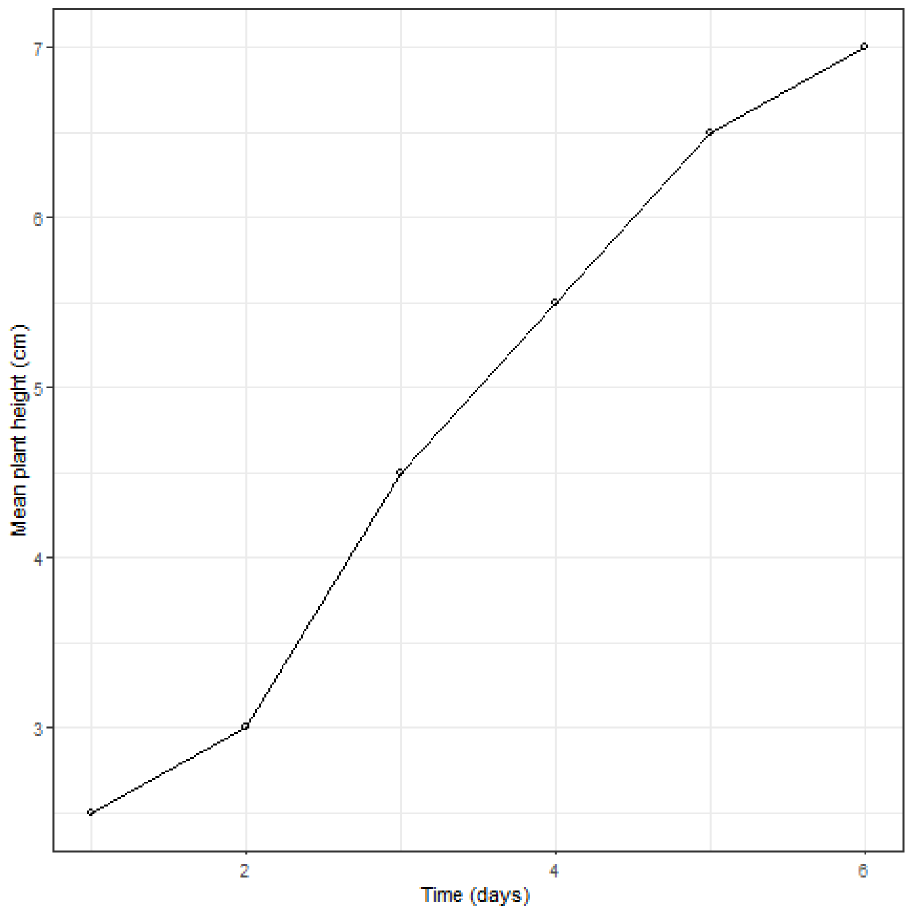

Figure 5: Figure 5. The effect of time on mean plant height (mm) for plants grown at 25 ° C.Image by Clerissa Copeland, licensed under CC BY-NC-SA 4.0

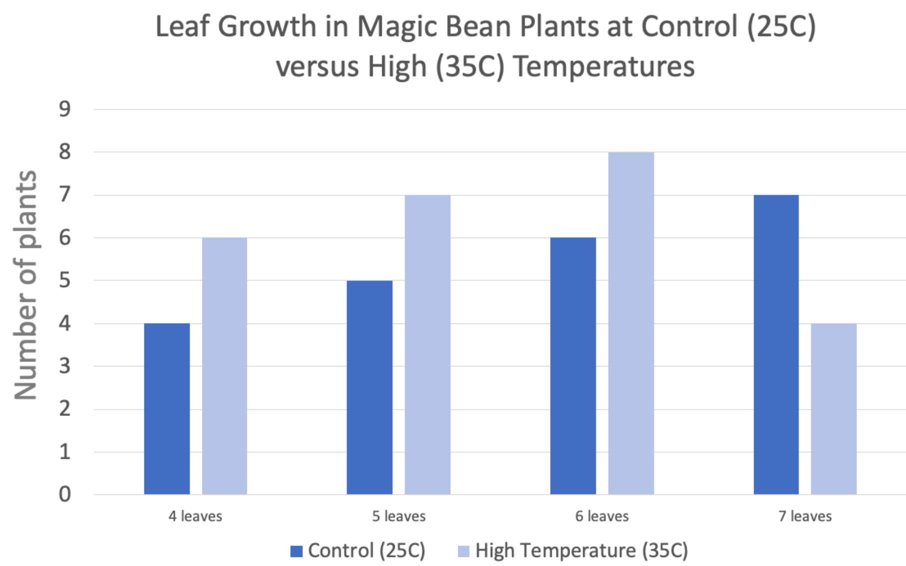

Figure 6: Figure 6. The effect of varying temperatures on the number of leaves on magic bean plants.Image by REY, licensed under CC BY-NC-SA 4.0

- Or, do the sample means represent the same population of measurements (i.e. is Ho supported by this evidence)?
- Do the 95% confidence intervals overlap or did you find a significant Chi-squared value?
- Are the results from the different trials that you ran consistent with each other? How so?

If your data allow you to reject Ho and provide support for Ha then you need to explain the specific effect the factor had on your organism and the processes that occurred to result in this response.

- Does this agree with all aspects of your prediction? If not, why not?

When possible, you should not only repeat your experiment as many times as is reasonable, but also compare your results to those of other investigators working on the same problem. This helps you to determine how reproducible you would expect your result to be, which also helps provide evidence for how likely your results reflect the "truth".

It is important to remember that scientific investigations do not always yield the anticipated results. If there are discrepancies between your results and those of others, or what you expected to find based on your reading of the scientific literature, this is the place to try and explain those discrepancies. As a general rule, this means looking at the published results of other scientists, and critically comparing the work they did to yours, to see if similarities make sense, and discrepancies can be explained.

All information obtained from other sources, or any ideas that are not your own, must be properly cited in the body of your poster presentation and included in a ”Literature Cited” or ”References” list at the end. This is an essential part of science, and academic endeavours. No scientist ever works in a vacuum, and comparing to others is expected, so it’s perfectly normal and expected that you will learn from the work of others, and cite them.

Citing

Citing the work of others

Science is, at its heart, a collaborative endeavour. No scientist ever works alone, especially in the modern times we live in. As such, all scientists share ideas and information in one way or another. Information is our currency: we trade it, share it, and make it grow. This also means that giving credit to others for their ideas and information is a vitally important part of science, not to mention academic life and the Open Science movement. Again, you will learn more about Open Science and how it works next week, in the asynchronous material of Lab 3.

For your written work, every time you mention an item of information or any idea that is not your own, the source must be credited in the text. This refers not only to published material but can also include personal communications from colleagues and professors. In scientific writing, we avoid using direct quotations and footnotes wherever possible. We do not copy *verbatim* from our sources - i.e. copy & paste. Instead, put the source away when you write, so that you naturally rephrase the material into your own words. Finally, acknowledge the source using the appropriate style and format for the work you’re trying to cite.

While many different formats exist for citing your sources, for the purposes of this project the APA reference and citation style is to be followed. The Procedures and Guidelines: APA Citations has a quick reference guide and other resources that you can follow up with.

Academic Integrity & Copyright

Citing falls under two broader categories, academic integrity and copyright. Academic integrity is about honest, responsible conduct in academia. Copyright is the legal framework that governs how the things we make - the things we write and draw - can be copied and distributed by others. The Procedures and

Guidelines: Academic Integrity and Copyright sections provide further guidance on each of these.

Closing Remarks...

Time to start doing science!

You will conduct your research project over the course of the term, alternating between in person sessions, when you will work with your group to plan and / or implement your research plan, and asynchronous online sessions, when you will work through Canvas modules to help you gain the skills and information you need to be successful in this project.

The weekly plan is written in your course syllabus, along with which weeks are in person - called synchronous in the syllabus - and which ones are online - called asynchronous in the syllabus. Most weeks you will have something to submit that is related to this project in one way or another (except for the weeks you're actually running your experiments and collecting the data!. You will get more information about this in you first in-person lab this week.

One final thought… while science is serious and noble and logical and all of the rest… it's is also exciting and fun! We can be rigorous and careful in the work we do, and also be really excited by the cool questions we're exploring. So even while you are doing your best to learn the principles of science, it is our sincere hope that you also find the joy in research and discovery!

Happy Sciencing!

Assignment

Please use the following template for this assignment:

20210813_Lab2b_Experimental-Design-PreLab-Assignment.docx (18 KB)

NOTE This assignment is due at the start of your next in-person lab (i.e. Lab 4).

Lab 3: Open Science

Last updated 2021-08-16

This week's lab content can be found here. You are asked to cover Part 1, Principles of Open Science.

The accompanying quiz can be found in Canvas.

Lab 4: Experimental Research Pilot

Last updated 2021-08-16

This week you will be trying out your experimental design.

This week is all about learning! Be prepared that things may not go as expected and that is okay. This week is about taking the time to figure out what works, what doesn't work and what we need to change.

This is all part of research but will help better prepare you for your next few weeks of data collection.

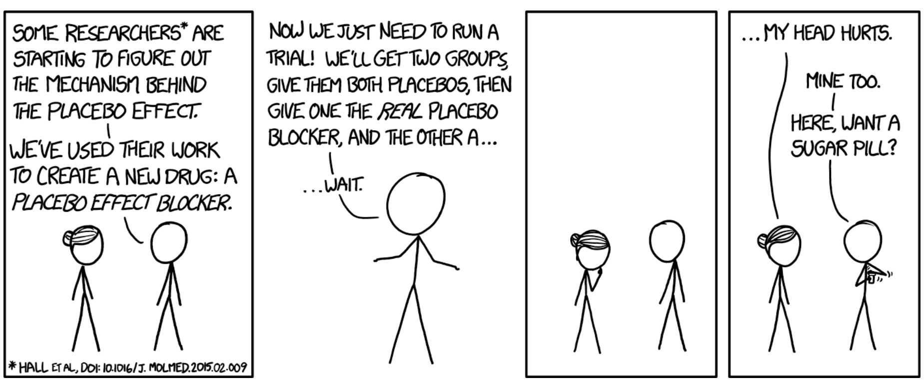

Figure 7: Figure 1. Placebo comic.Image by xkcd. Licensed under CC BY-NC 2.5

Questions

Remember, there are 20 of you and only one teaching assistant in your lab. You may be feeling anxious and frustrated at times as you may not be able to get the help you need immediately but please be patient.

Your teaching assistant will come around to provide whatever support they can, however there may be a bit of a wait. While you are waiting on your teaching assistant see if you and your partner aren't able to figure things out together. You may find you do not need your teaching assistant when you put both heads together.

Materials & Safety

All the materials needed will be along the far bench with some items being held in the fume hood. For those items in the fume hood please ensure you follow the guidelines indicated by your teaching assistant. There are safety issues that need to be considered and so we must ensure we are following all laboratory safety rules. If you are in doubt of how to dispose of something or where it should be placed please ask rather than guess.

Tidy Up

There are other students coming in immediately after you so please be considerate.Your station should look the same as it did when you arrived.Take the time to clean up your area and put everything away.

Have Fun

Yes, this may be a challenging week for you but we expect that it will also be a very enjoyable experience… you get to test your own theories! Have fun with it!

Lab 5: Naming Conventions

Last updated 2021-08-16

This week's lab content can be found here. You are asked to read

- Chapter 1: File and Data Management and
- Chapter 2: File Naming.

The accompanying quiz can be found in Canvas.

Lab 6: Data Collection

Last updated 2021-08-16

Now that you have conducted your pilot experiment and saw how it went, you are ready to collect some data.

Make sure you have worked through any major changes to your experimental methodology with your teaching assistant and you have carefully documented everything that happened during your pilot week (Lab 4). This information will come in handy when it comes time to explaining your results, but for this week you can just focus on collecting data for your research!

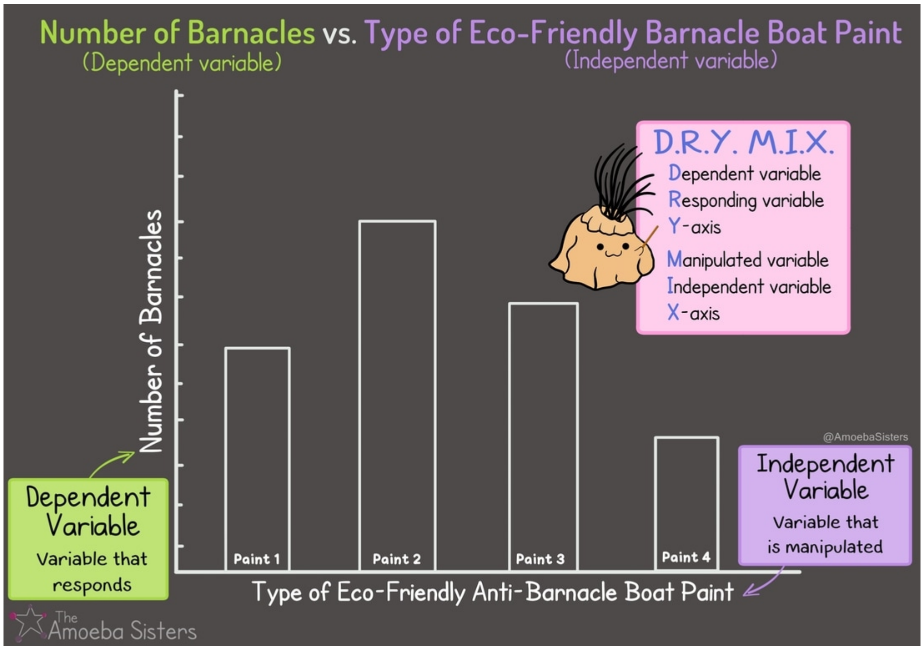

Figure 8: Figure 1. Types of variables comic.Image by Amoeba Sisters. Used in accordance with the creators' terms of use

Lab 7: Intro to R & Shiny Apps

Last updated 2021-08-16

This lab will have you putting into practice through visualizations and data tables much of what was covered in Lab 2B.

0.1 Data Tables & Figures

Before we tackle data visualization and analysis, we should talk about data presentation conventions.

The UBCO Biology faculty has a Procedures and Guidelines document outlining conventions for data presentation based on current ”best practices” in Biology. Keep in mind that you may encounter small differences when working with data or reading the results of research from other disciplines. That being said, the aim of these conventions is to achieve consistency among faculty, instructors, and students in how data are summarized and presented within lab reports and research papers.

Before continuing through this lab, please take some time individually or with your groups to read sections 5.1, 5.2, and 5.4 of the UBCO Biology Procedures and Guidelines document. For this lab's assignment and all future assignments in this course, you wil be expected to follow these guidelines!

0.2 Intro to R

Visualizing data is a key part of science communication. Until now, it is likely that you've used or interacted with visualizations built using tools like Excel. In Biology, it is increasingly common to not use Excel, but instead to use an application, or programming language, called R.

R offers several advantages when it comes to building visualizations. R is what we call a scripted progamming language. When we use scripts to manage our data and create visualizations, we're engaging in reproducible workflows, specifically computationally reproducible workflows.

0.3 Computational reproducibility

What is computational reproducibility?

When you load data into an application like Excel, you click through a series of options to clean up or organize and then visualize your data. If you wanted to repeat this process, you would have to manually go through the same series of steps over and over again, which is a lot of mouse clicking. It's also very difficult to communicate your workflow from raw data to visualization as you would have to write down every step and do so in a way that was clear enough for someone else to reproduce exactly.

Working in a scripted programming language like R, instead of clicking buttons, we write code that tells the program how to clean, organize, and then visualize our data. When we want to repeat the process, we just run our script - or mini program - and it exactly, computationally, reproduces our workflow. It's less work for us and it's reproducible. And, since anyone we share our script with can read our script to see the steps taken to go from raw data to visualization, it's transparent. We just don't get this level of reproducibility and transparency with Excel.

0.4 Shiny Apps

In BIOL 116 we're not going to learn how to write R, we'll save that for later. What we are going to do is start working with learning to visualize data in an application built in R called Shiny Apps. Shiny Apps allow a user to interactively visualize a data set. It will also show you the R code that is running in the background to build the visualization. This will be your first window into R.

In this lab, we'll work with a couple of data sets that are preloaded into the Shiny App, so you can explore mapping different variables to a visual space.

The first is a data set about penguins, called `palmerpenguins`. From the authors: "The `palmerpenguins` data contains size measurements for three penguin species observed on three islands in the Palmer Archipelago, Antarctica." Read more about the `palmerpenguins` data set here if you'd like.

The second is about cars, called `mtcars`. From the authors: "The data was extracted from the 1974 *Motor Trend* US magazine, and comprises fuel consumption and 10 aspects of automobile design and performance for 32 automobiles (1973–74 models).” Read more about the `mtcars` data set here if you’d like.

0.5 Getting Started

The Shiny App that we’ll be using can be found at this link.

The following is a descriptive overview of what you’ll find in the two panes of the application.

Left Pane

On the left navigation you have four menu options: Welcome; Upload Data; Plot; Descriptive Stats.

Welcome

This is an overview of the app itself.

Upload Data

For now, ignore the option to upload your own data. Instead, use either the `palmerpenguins` data or the `mtcars` data, selected from the drop down menu in the Plot menu. In **Lab 9**, you’ll upload your own data set from your experimental research project to build visualizations for your report.

Plot

Here we select the variables we’d like to visualize.

You’ll want to first select a dataset, `penguins` or `mtcars`, and then select an `x` and a `y` variable from the drop down menus. Lastly, tell the Shiny App if your variable is `quantitative` or `categorical`. In Lab 2 we learned about each of these variable types, but here is a brief refresher:

Quantitative, or numeric, data may be either discrete or continuous.

- Discrete quantitative data are whole integers - population numbers are a good example of this, as we can’t have half a person!
- Continuous quantitative data on the other hand are data that lies on a continuum - an example is temperature, where there are infinite potential temperatures between 25 and 26 degrees Celsius, but our data collection tool or convention determines to what decimal point we’ll record a given temperature.

Categorical data, per its name, deals with categories of things and may be either nominal or ordinal.

- Ordinal categorical data has an intrinsic order where one thing is more or less than another - storm severity is often classified by stages - Stage 1, Stage 2 … Stage 5 - where Stage 1 is less severe than stage 5. We don't know how much more severe one stage is than the next - that is we can't quantify the difference - but we know there is an intrinsic order.
- Nominal categorical data has no natural order - the ordering of the colours blue, pink, or white makes no difference to the data or analysis.

Once you've selected your data sources and identified the data types you're working with, you'll be presented with a series of plotting options. You may also be presented with the option to group one of your variables. You will also be presented with the option to save a copy of your plot.

Descriptive Stats

Here you can run basic descriptive stats on the data that you've plotted.

Similar to the instructions under the Plot tab, here, you must first select a dataset, `penguins` or `mtcars`, and then select two variables from the drop down menus, `Variable 1` and `Variable 2`. Again, you'll need to tell the Shiny App whether each of these variables are `quantitative` or `categorical`.

Once you've made your selections, the Shiny App will automatically calculated the preferred descriptive statistics for the type of variables you have selected. For example, if you choose

- **two quantitative variables**, the Shiny App will provide you with the sample size `n`, mean, standard deviation `sd`, median, and inter-quartile range `iqr`.
- **one categorical and one quantitative variable**, you will be provided with the sample size `n`, mean, standard deviation `sd`, median, and inter-quartile range `iqr` for each group of your categorical variable.
- **two categorical variables**, then a frequency table will be displayed showing you how many subjects fall into each group.

Right Pane

Under each menu option (i.e. Plot), the right pane displays three sections: Instructions; Output; Source Code

Instructions

This green box shows detailed instructions for how to use that specific page. For example, under the `Plot` menu option, there are detailed instructions for how to create, visualize, and save your plot.

Output

This blue box shows the output that the Shiny App creates based on your selections. For example, under the `Plot` menu option, this box will show the plot you created.

Source Code

This orange box shows the script used by R to generate the output. If there are any packages that need to be loaded in order to run the script, they are shown within the `library()` function at the beginning of the script. You don't need to worry about doing anything with this part for now. Rather, the script is included to enhance the transparency and computational reproducibility of your analysis. For example, since you know the script used to create your plots and calculate descriptive statistics, if you or someone else wanted to reproduce your methods, they could load your data, copy this script into R, and re-run everything exactly as shown in the Shiny App.

0.6 Screencast Demo

The following is a brief screencast introduction to using the Shiny App with the included datasets.

0.7 Activity

This is your opportunity to play around with several different data types, visualizations and descriptive statistical summaries. The assignment will ask that you submit a couple of examples of both data visualizations and tables, all properly formatted according to Section 5 of the UBCO Biology Procedures and Guidelines document.

0.8 Assignment

Please use the following template for this assignment:

20210810_Lab7_Shiny-App-Visualization_Assignment_V1.docx (22 KB)

0.9 Grading Rubric

Question 1

1.5 marks for each correct figure based on types of variable inputs. Since there are 2 figures this part is out of 3 marks in total.

- Two categorical variables = mosaic or barplot
- Two numeric variables = scatterplot
- One numeric and one categorical variable = stripchart (if < 20 observations in each group), boxplot (if > 20 observations in each group)

1 mark for each figure caption that follows guidelines. Since there are 2 figures this part is out of 2 marks in total. At a minimum a caption:

- is located below the figure (0.25 marks)
- includes a figure number (0.25 marks)
- describes both variables (0.25 marks)
- includes sample sizes (0.25 marks)

Question 2

5 marks for a correctly formatted table.

- Correct descriptive statistics were used for each variable type (1 mark)
- The heading is placed above the table (1 mark)
- The table should be interpretable on its own thanks to an informative heading (1 mark)
- Sample sizes and units are always included (1 mark)
- Use horizontal lines only; these are often placed above and below headings, and at bottom of table (1 mark)

Lab 8: Data Collection Continued

Last updated 2021-08-16

By now you should have already successfully run your entire experiment once, and collected some data.This is great!Take your time in your lab to attain as many trials as you can.The more trials you conduct the more assured you can be in the trends you are seeing.This is your last opportunity to collect your data so make the most of it.

Remember that you should be working to minimize the differences between today's trials and the trial(s) you ran in the last data collection period. Your methods should be as similar as possible - assuming, of course, your experiment worked last time.

Note making changes adds a new variable to your experiment, which will have an unknown impact on your results. If you must make changes for some reason, make sure that you carefully document everything once again.

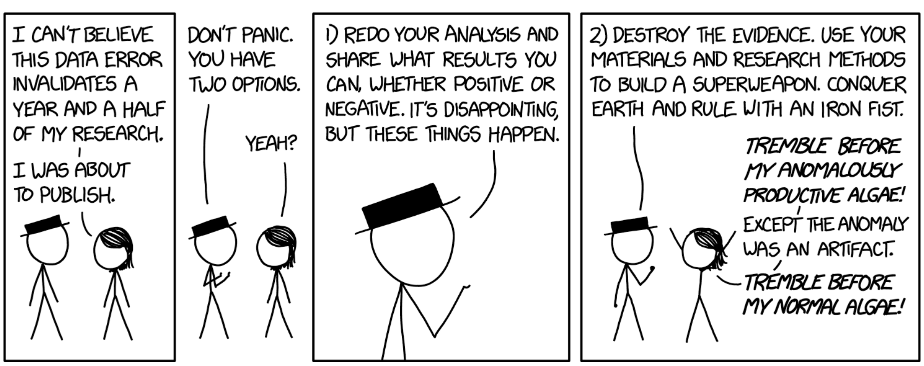

Figure 9: Figure 1. Data error.Image by xkcd. Licensed under CC BY-NC 2.5

Collaboration is a key part of science. During this time, you may want to start informally discussing what you are seeing with your partner, peers and / or your teaching assistant to see if you can explain what you are seeing.

- If you are seeing what you hypothesized, ask yourself, is it for the reason you thought? Why or why not?
- If you are seeing something different than what you had originally hypothesized why is that?

Scientists often use each other to discuss their findings and / or issues in order to help better understand what is going on. Your peers and your teaching assistants are resources for you so take advantage. Collaboration provides the opportunity to learn much much more about science and the scientific method so use it wisely.

Lab 9: Data Analysis & Shiny Apps

Last updated 2021-08-16

CONTENT PENDING

Lab 10: Poster Presentation

Last updated 2021-08-16

CONTENT PENDING