Description of the Scientific Process

John Stewart, Andrew Doust, Julie Angle, and Robert Burnap

The textbook description of the steps of the scientific process seems straightforward: observe, hypothesize, experiment, analyze, theorize, and communicate, but every one of those steps is fraught with misconceptions and difficulties, and science doesn't always follow those steps as neatly as we would like. What experiment is the paleobotanist doing when she digs up a plant fossil and analyzes it? It's not exactly a controlled experiment, but it is science nonetheless. She can generate perfectly valid theories and communicate those to her peers. Still, the textbook guidelines for scientific work are valid and provide a great framework for your own science fair project.

In this document, we will explore the steps of scientific processes in a practical way. We will give you advice about what to do and warn you of common pitfalls along the way. We will show you examples of success and failure—and really, failure is often the best, if cruel, teacher. We will show you alternative approaches that you may take. While this guide is laid out in the order in which you will be doing your work, everything is connected, and you should read the whole thing before you get started doing your research. Well, at least read ahead a section or two, since it might help.

We are assuming that you will be doing original research and not just mixing baking soda, vinegar, and food coloring to make a "volcano." Original research is harder than just showing off some well-known scientific principle. You have to think independently and creatively. Possibly the very best thing about original science, though, is that when you do it, you become the first person to know a piece of the truth about the universe that no one else has ever known. Then, you get to tell the world about your discoveries. Trust us—that's rewarding. It's like

Before you dig into this document, you should consider who "we" are in the text. We can't help but to let our backgrounds affect our examples and outlooks.

John Stewart: I'm the primary writer, and I'm a plant guy and a genetics guy. Maybe you're a physics or geology person. That's fine and great. Just know that my examples might skew toward the biological and botanical. I don't want to exclude your interests, but the examples that I include are just the ones that came to mind and that I am more excited about. Try to think about examples from your field of interest when you see what I write. It's a great exercise, and if you can't think of one straight away, it's a good chance for you to brainstorm with your peers.

Andrew Doust: I am also a plant guy and geneticist, but with a strong interest in the genes that control development of organisms. In particular I am interested in the genes that underlie human domestication of crops such as wheat, rice, and corn. We will try and include examples from each of our areas of expertise and interest so that you get a broader range of views on what science is and what interesting projects you might tackle.

delicious lemonade after an afternoon of yard work but way better. In these pages, we will treat you as someone who wants to achieve that. The language will be informal and (I hope) warm. We will avoid getting overly technical when possible, but we don't want to talk down to you. If you have decided to do a science fair project, we truly respect you and your efforts. Now, get cracking.

Julie Angle: I am a science educator. As a high school science teacher I taught chemistry and biology, and mentored students in science research and coached students to regional, state, and international science and engineering fair competitions. Now that I teach at the college level I am interested in how to better prepare individuals to become excellent science teachers and help them become proficient in the three constructs of science literacy.

Rob Burnap: I am interested in basic molecular aspects of how cells function and most of my work involves the study of photosynthesis. We use model bacteria, known as cyanobacteria, that possess essentially the same mechanism of photosynthesis found in plants and algae, but because of the genetic simplicity of these bacteria, we are able to detailed molecular mechanisms to test our hypotheses on how photosynthesis works and how it is used to drive the growth of the organism.

Observation and Finding a Problem to Study

The first step in doing an **experiment** is finding your **system**[1]. Thinking of something specific to study that is interesting to you is often the most difficult part of the process, so you shouldn't feel like a failure at this step. (Failure comes later.) You need to ask yourself some questions[2].

A lot of scientific research is performed on model systems. These are well-studied systems that are well-understood, so scientists can control for pitfalls and more easily reference other work. In biology, they are very common, as there are a number of model organisms. While humans can be considered a model organism, it's hard to design and fill out the paperwork for ethical experiments with people, so experiments are done with *Pan troglodytes* (chimpanzee) and *Macaca mulatta* (rhesus macaque). Others include *Drosophila melanogaster* (fruit fly), *Arabidopsis thaliana* (a weed), and *Escherichia coli* (a kind of bacteria), *Dictyostelium discoideum* (a slime mold), and *Caenorhabditis elegans* (a tiny worm).

In human research, HeLa lines (human cells derived from one woman's cancer in 1951—it's a fascinating and compelling story that you should look into) are frequently used, since they have human genes.

First, what are you interested in? Do you like a particular field of science the most? This question won't get you to your experiment, but it will narrow the list of possible experiments down. If you happen to come up with an experiment that is outside of what you think of as your favorite kind of science, though, that's fine. One of the best things about being a scientist is that our work constantly changes. Let's say you like biology. That is the study of life, all life. That might be a bit too broad, so you will need to narrow it down. With something like biology, you can do that several ways. You can break it down by the kinds of life you want to study (by taxon, as a biologist would say.) That could lead you to zoology (animals), botany (plants), microbiology (bacteria and other small things), or mycology (fungi). Alternatively, you can break it down by approach: genetics, biochemistry, physiology, anatomy, etc. Preferably, you should do both. Alternatively, you can choose based on application. Consider, for example, cancer research, toxicology, or robotics. You may be limited to whatever your materials allow you to do, and you should accept that as a positive thing, since it narrows your choices.

[1] The word system will be used throughout this paper. It's kind of vague, but basically it is the thing that you are studying. It could be the lake in the park, the drought response of *Capsella bursa-pastoris* (shepherd's purse), or how four-year-olds respond to the prisoner's dilemma. It's a very general term for the thing you are studying.

[2] Note from Andrew: I would say that it is important to think of many possibilities, some of which are certain not to work when you really start to look at them closely!

That still doesn't necessarily give you a project, so here's another bit of advice: think locally. If you want to do original research for your science fair project, think about what's immediately around you. You'll have ready access to local problems. The odds that someone has studied a local kind of rock, organism, or whatever are smaller than if you think about something everyone already knows is big and important. Local quirks about wildlife and geology are often very interesting but unexplored.

Speaking of local, consider what other people have told you. Maybe a friend of your parents mentioned that something was funny about the soil near his home. That might be something to look at. Likewise, you can ask people you know what they think would be interesting to study. Most folks will look at you blankly, but you might get a few people who give you the start of an idea that you can follow.

Instead of asking non-experts, you can look at what the experts say. While a lot of real scientific research is behind a pay wall (paying $25 to $50 for access to a 10-page paper is ridiculous but common) enough of it is available to you for free that you can skim through it. You will be looking through titles that seem really technical or really vague, but you can read the **abstracts**. When you do so, you are looking for experiments that were done on systems similar to the ones you're looking at. A scientist did a study on wheat growth relative to clay in the soil. Why not repeat that with a local weed[1]? Look at the scientist's methodology and repeat it for your weed.

John's anecdote: Reading abstracts is something that scientists do all the time. I once had a boss who described his reading of papers as follows:

1. Read the title. If it interests you, then…
2. Read the abstract. It's a summary of why the work was done, what was done, what was discovered, and why that's important. (We'll talk about writing abstracts when we talk about writing in general.) If the abstract interest you, then…
3. Look at the figures and tables. If you want to understand what's going on, then…
4. Read the discussion and reference the results. If you still want background, or you need to know how things were done, then read the introduction and the materials and methods.

You could just start by reading the whole paper, but this is more efficient when you're just exploring the literature.

One of the most important things to think about at all stages of the scientific process is **variation**. At this point, think about how your subject might vary against something else that might vary. In the simplest case, you will have some measure that can vary from subject to subject, i.e., the height of a plant, and you will have another measure that you can control, i.e.,

[1] Don't do experiments on illegal things here. This is a science fair project, so your research will be publically known.

the amount of water the plant receives. We call the first measure the **dependent variable** and the second measure the **independent variable**. You may have several of each of these, and we'll talk about how to manage those when we talk about experimental design. Right now, you need to think about what different things might vary.

What if you want to examine something that you aren't controlling? Consider a lake and the soils surrounding a lake. You might find that the soils vary in consistency or type (soil science is a big and complicated field; you can find a lot of job opportunities in soil science, by the way) and that the distance from the lake has a lot to do with that. The soil type would be your dependent variable, and the distance from the lake would be your independent variable. In another experiment on wild plants and animals, both distance from the lake and soil type would be considered independent variables. We call an experiment like this a **natural experiment**, and it's something you see frequently in medicine, geology, ecology, and even in social sciences like economics. Natural experiments are often more difficult to perform and noisier[1], but they can shed light on problems that cannot be examined with traditional experiments.

So, you've found something you want to test. What next? Can you just jump into your experiment? No! You need to create a **hypothesis**. The relevant Oxford English Dictionary definition of the hypothesis is, "A supposition or conjecture put forth to account for known facts; esp. in the sciences, a provisional supposition from which to draw conclusions that shall be in accordance with known facts, and which serves as a starting-point for further investigation by which it may be proved or disproved and the true theory arrived at." But that's stuffy English professor speech. Here's what you need to know. Your hypothesis is a reasonable guess of what might be going on with your system that can be tested with an experiment. "I hypothesize that the plant species *Taraxacum officinale* (common dandelion) will have less biomass when it has less available water," is a reasonable, if boring, hypothesis. Other examples:

- I hypothesize that plant or microbial communities differ in sites with different amounts of sedimentary rock.
- I hypothesize that the seed of *Acer negundo* (boxelder) will fall more quickly than the seed of *Acer rubrum* (red maple.)
- I hypothesize that farmers will be more accepting of controlled burns than city residents will be.
- I hypothesize that beef fat content will decrease in cattle raised in drought conditions relative to normal climatic conditions.
- I hypothesize that the traffic capacity of a local avenue will decrease after a new traffic light is added.

---

[1] **Noise** is a term that scientists use to refer to variations in the data that are not meaningful. It's like static on an analog radio set. Scientists often refer to separating the **signal** from the noise. The signal is the meaningful information that you can find in variation. It's like the music on a radio station. We will talk more about how you can use statistics to separate the signal from the noise in a separate part of this document.

- I hypothesize that different quantities of water and sodium citrate will be needed to emulsify different kinds of cheeses.
- I hypothesize that molded ABS plastic will have higher flexural strength than 3D printed ABS plastic.
- I hypothesize that mutant cyanobacterial cells with defects in their light-capturing antennae will require more light to achieve the same growth rate as normal cells.

So, you've got your hypothesis. What now?

Designing Your Experiment

To test your hypothesis, you need to come up with an experiment. This is when you get to decide what work you will actually do. You need to start by taking your hypothesis and the independent and dependent variables that you think that you can test and figuring out how you are going to test them.

In a traditional **controlled experiment**, you will need what is called the control sample and the **treatment sample**, or **experimental sample**, (maybe samples.) Let's start with the word **sample**. It refers to a set of specimens, measurements, etc. that you will group together for repetition. The number of those things is called the **sample size**, and we will address that soon. Your **control sample** is the sample that is under somewhat normal conditions. The term "normal" is loaded, but the important thing is that you know what "normal" means for your experiment. Scientists argue and fuss about how valid controls are all the time, but when you feel satisfied that you have a set of normal conditions, then you should proceed. The **treatment samples** are those with conditions that somehow vary from your control in defined ways. For example, you water your control dandelions with 100 mL of water every day, you water your drought treatment dandelions with 50 mL of water every third day, and you keep your waterlogged dandelions semi-submerged. Here, you are varying water exposure relative to your control (you are also varying water volume and watering frequency, so you may have to adjust for that.)

Materials and methods sections of papers are really boring, and even scientists tend to skim over them. We do use them to figure out what we're going to do, though. I may want to use someone else's methodology on my system. Here is some advice for how to digest them:

- Materials and methods are densely written. Separately, write down what is relevant so that you can write it down and adapt it into a protocol.
- Authors often reference other articles when they describe what they are doing. You may have to go and find them to replicate the methods.
- Know what units the authors are talking about. You'll see a lot of nm's, μL's, and ω's in papers. If you don't know, then look them up (the funny symbols are Greek, typically). If you check on Wikipedia, you can find out all the different ways they are used in science.

It's important that you consider what might be realistic for your treatments. What is a reasonable range of the variable? I'd bet that those waterlogged dandelions aren't going to get very far, and it might not be a good level of exposure, since they will likely die within a few days. You have to consider this issue for your own experiment. Here is where you can dive into the literature. Scientific papers have to include their materials and methods, that is what things they used and how they used them. Check those out.

How do you find out what treatment extremes you could use? First, you should look at what others have done in similar experiments. Yes, that means you have to look at the literature again, but if you used any scientific literature to get to the point you are now, then that same literature will probably have what you need. You can also ask an expert. Any scientist worth his or her salt would be delighted to correspond with you about setting up an experiment, and really, the biggest danger here is that they make your scope of study way too big to handle. Remember that what they tell you is advice and not a dictate. Finally, you can try a pilot experiment.

> John's anecdote: I once performed a study testing the rate of hybridization of shortleaf pine and loblolly pine in a shortleaf pine forest adjacent to a loblolly pine plantation. As expected, there were many hybrids in the sample site across a road from the plantation, but all of the other sites that were staggered away from the site had few hybrids. As it turned out, my results were not great, and I could only make weak claims in my paper. The lesson is that you should have a good idea of what scale your treatments act over. If my sample sites were all closer to the plantation, I may have been able to detect the distance that loblolly pine pollen was fertilizing the shortleaf pine cones.
>
> Andrew's note: I LOVE pilot experiments! You can be wacky with them, really testing out extremes. If you don't you may find yourself in the position of my first grad student who did lots of experiments being too small to get really interesting results. Of course, what other scientists have already published can be a pilot experiment for you as well, to help you understand the responsiveness of the study system.

Pilot experiments are miniature trials that you can use to test your methods. (Professionals sometimes use them to justify grant requests.) If you have the time and resources, then you can try running your experiment with a lot of treatments to work the kinks out. In addition to informing you about extremes of treatment, pilot experiments can help you work out bugs in an apparatus you built or show you that you need to take extra care in some step.

Consider our dandelion example. Let's say that you have already decided to put your plants into 2" × 2" × 3" pots. Now, you need to know how much water you should add to get reasonable plant growth. You will just want a range of watering conditions repeated two or three times to get a feel for their effects. Those conditions might be 2 mL/day, 5 mL/day, 25 mL/day, 50 mL/day, 100 mL/day, and 250 mL/day. Notice that these conditions do not scale in a linear way. That is, they are not in increments of 10 mL or something. This way, you can capture a better representation of varying growth conditions with only a few samples. You might also only grow your plants for 2 weeks in the pilot experiment instead of the full 4 weeks you would use in a planned experiment.

You will also need to visualize how your experiments will work and where they might introduce accidental **experimental bias**. Consider a flat of daisies in an experiment. It's placed on a windowsill on the east facing side of a house. In the morning, the flat gets sun, so all seems well. You then notice that the easternmost daisies are growing faster, and you realize that they are shading out the other daisies. More than likely, this experiment has been ruined by an accidental bias introduced in the design. Even if the flat of daisies was placed under a sunlamp so that all of the plants got even light, you would notice that the plants around the edges are growing better, since they are less crowded. We have to ignore our edge plants and only collect data for the others. If you had designed the experiment so that all of your treatments were in neat rows, and your control plants had all been on the edge, you just lost your control samples. Ouch!

Besides edge effects, there are a great many things that can add bias to your results. If you need to test multiple samples with a device like a probe, each exposure could change the probe's sensitivity. Again, if you measured all of your control samples first, and then measured all of your treatment samples, you might introduce a bias into your experiment through the order of sampling. Social scientists frequently use surveys and have to be very mindful of the wording of the questions they use. A given political affiliation or racial group might respond very differently from another just because some word with a loaded history was included in the question.

What do you do about this? First, randomize. Randomize the order of the plants in the flat or the order in time of the samples you measure with a probe. You can do this in a spreadsheet[1] or with a deck of playing cards. Second, include dummy samples. On your flat of daisies, just grow daisies around the edge that are not considered to be treatment or control plants. You can then ignore them in the analysis. If you are doing some sort of chemical assay on them, then they can act as practice specimens later. Be creative!

Sample size is the bane of scientific studies. We don't know how many times we've seen a scientist say that their sample size was too small. They couldn't end up making statistical claims that they wanted to make about their studies, because they couldn't find enough smallmouth bass, or the small proportion of sample locations with rose rocks is too small to make meaningful conclusions. Sample size issues can be devilishly frustrating to deal with, but you can overcome them.

First, how large should your sample size be? It depends on how much variation you expect to see in your dependent variable and on how many samples you think you will have to discard as the experiment progresses. If you don't know those things, you can again ask an expert, read the literature, or attempt a pilot experiment. When asked how large a sample size should be, most scientists will tell you that it should be as large as possible, and that's true. There are trade-offs, though. Given limited resources, you can have more treatments with smaller sample sizes or fewer treatments with larger ones. If you can manage 100 specimens, you could make a control and a treatment with samples sizes of 50, or you could make a control and 3 treatments with samples sizes of 25. Which should you do?

---

[1] See our Excel Tutorial.

If you have no prior information and just want to proceed with the experiment, going with 40 individuals is not a terrible rule of thumb[1]. You can often get good statistical significance with 30 individuals, but you must assume that you will lose some during the experimental process. If you can't get that many, do not despair. If the differences among treatments and controls are large enough, you may still get great results with only 4 samples, though that is very risky.

Finally, you've designed your experiment. Now, ask yourself once more, how will this experiment test my hypothesis? If you aren't sure, either redesign your experiment, or reword your hypothesis. Both approaches can be valid. Sometimes, you lose your way somewhere in the experimental design process, but sometimes, your original hypothesis was just too hard to test with the experimental resources you have. Do what you feel is best. It's your project.

---

[1] The website http://edis.ifas.ufl.edu/pd006 has a decent guide to finding out how big your sample should be.

Running Your Experiment

So, you have designed your experiment and gathered your materials. You are ready to begin. When you run your experiment, the most important thing to do is to *be consistent*. If you used Brand X potting soil for the control plants, then use Brand X potting soil for the treatments (unless the experiment was testing the quality of Brand X versus Brand Y potting soil, of course.) Apply even conditions across all of your samples as best as you can to avoid introducing **experimental bias**. It seems really simple, but as the experiment proceeds, it can be all too easy to get sloppy and unevenly apply some condition. Likewise, make your measurements consistent. If you use a Brand X pH probe for some of your samples, use it for the rest. Define how you will dissect each insect to measure abdomen length and do it the same way each time. Doing things unevenly will make your results highly variable and unreliable.

> Sometimes, mistakes happen that threaten to end the experiment prematurely. When that happens, step back and take a deep breath. (Go and punch a pillow and yell, too. It does help.) When you have gathered your wits and achieved a Zen-like state, consider what you can do to salvage the experiment. If you added twice as much hydrochloric acid to half of your samples, you might just redesign your experiment on the fly. You've been taking notes, right? You can use those to describe some of your samples as being new treatments. Just write it down and carry on. Who knows? Maybe you'll find something interesting. If it's early enough in the experiment, or you can afford more research, you can use much of the material from a failure, and you can certainly use the lessons so hard-won.

One way to stay consistent, especially in multi-day experiments, is to take notes. Get yourself a nice notebook dedicated to your research project. In it, write down the times of what you did, exactly what you did, how much of a chemical you used, measurements you took, funny things you observe about your samples, and so on. We don't think we've ever heard of anyone putting too many things in a scientific notebook. Do not underestimate the utility of the notebook. You can later use it to spot accidental experimental bias that you introduced or recreate part of your experiment for the sake of consistency. When you are creating your posters or papers for your project, you will probably find something you wrote down in your notebook important. Finally, the notebook is an artifact and a keepsake of your project.

During an experiment, it is often critical to label your samples. If you need to redo something or just prevent yourself from double-treating and individual, labels are really important. When you take notes, you can report exactly which individual or data point was affected by a mistake. Remember to upkeep those labels, too. Organic solvents and sunlight can erase your markings, and tape can peel off. If you are working with outdoor labels, avoid permanent markers (which aren't so permanent in the sun) and just go with pencils. Better yet, use stamped metal labels.

And those mistakes will happen. In science, the occasional spill, drop, or contamination will occur. You just have to carry on. What do you do with the would-be data points? Most often, you remove them. Keeping them around could add bias to your results. On occasion, if enough sample is destroyed, you may have to abandon the experiment, but that's tragic. When you designed your experiment, you should have made the sample size large enough to account for data loss, so most removals should not impact your study too much. Do not be too cavalier about this aspect of running your experiment, but take some comfort in having accounted for the possibility.

When you have completed your experiment, keep as many of the materials as you can. You might find unusual results later that are unaccounted for, and you can double check that there was not something physically wrong. Your peers may also be interested in exactly what you did. For science fairs especially, artifacts of your research can be very useful for your presentation.

Analyzing Data

In a way, data analysis begins with experimental design. Assuming all went well during the experiment, your analysis will start with that design. Before you run your statistical analysis[1], it is helpful to lay your data out and examine what it seems to show. However, your data will generally be a bunch of numbers, and it's hard to make anything out of that sort of thing. Depending on what you have, here are some things that you can do.

- Make a **scatter plot**. A scatter plot is a simple graph with each point positioned by two variables on a flat surface (Figure 1). Here's what you can get out of this:
    - You can spot outliers. **Outliers** are individuals that do not follow some common trend in your data. Very often (but certainly not always) outliers show that some individual underwent an experimental bias. Maybe that was the snake that you fed twice instead of once. Consult your notes and your common sense. We will show you what to do about outliers a little later.
    - You can get a real sense of the general trends. You might see where things bunch up or how increasing one thing decreases another. You can find crucial insight here.
- Make a **bar** (or **column**) **chart**. It's a simple way to visualize the magnitude of your data (.)
    - You can add error bars to indicating the confidence interval of your data as well.

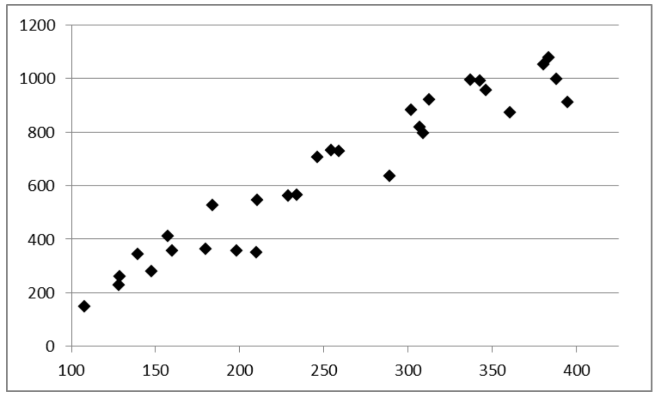

Figure 1: A scatter plot.

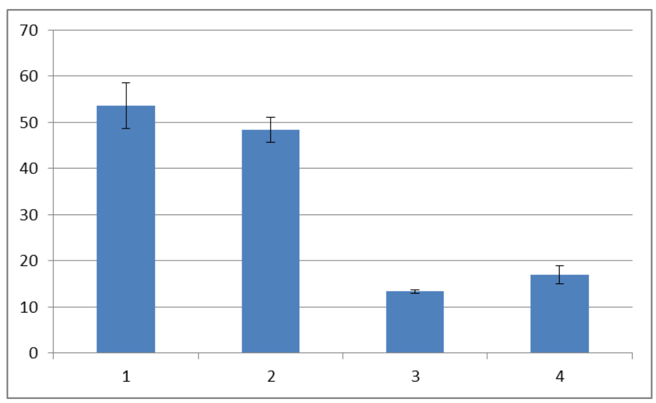

Figure 2: A column chart with error bars.

[1] Here, I am assuming that you are using statistical analysis with quantitative data. There is some science that does not require statistics. Consider a paleontologist measuring a dinosaur bone or a primatologist observing chimpanzee behavior. Medical case studies are regularly published, and medical science is all the better for it. You might have an experiment devoid of statistics, but these experiments are rarer, and scientists are more likely to accept your results, if you can apply rigorous statistical tests.

- Make a **pie chart**. Pie charts are great ways to see proportions of data (Figure 3).
- Make a **line graph**. These are especially useful for making changes over time visible (Figure 4).
- Organize and reorganize your data. Try to find patterns that you can describe in your results.

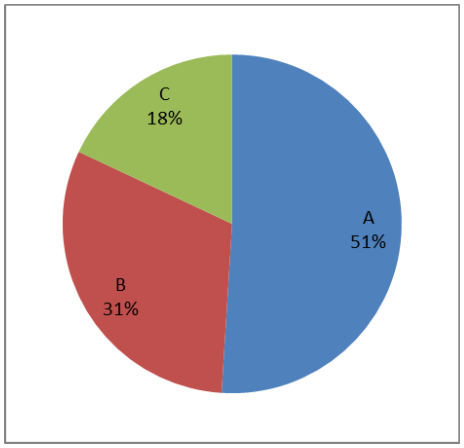

Figure 3: A simple pie chart.

Sometimes, it's only when you are analyzing your data that you realize that some of it might be bad. You took good notes, and you were consistent, but still, something doesn't look right about individual 23A. One of the most common sources of bad data points is zero that isn't a zero. Zero values often represent missing information. If a spider's height is 0 mm, then it's not a spider. (Maybe you squished it. No one would blame you.) Don't represent missing values with 0's; the convention is to use "·" or "–". In a spreadsheet, it's just an empty cell.

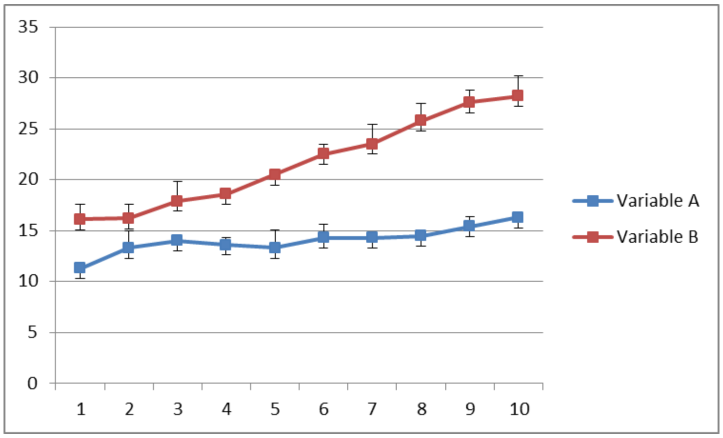

Figure 4: A line graph with two variables and error bars.

Be careful to know your instruments' limitations. You can measure something with an instrument but then exceed the instrument's measurement range. Think of a thermometer. There is some maximum and minimum temperature that it can measure. If you have exceeded the maximum temperature of 125° F, then you might only note that it was 125° F. That would be wrong. All sorts of other instruments can have their own limits, so be sure that your measurements fall within them.

Finally, you may have outlier individuals. You wrote down the wrong thing when you measured it. One of your reagents wasn't thoroughly mixed, and it got a lot more salt than the others did. The list can go on and on. If you think that you have an outlier, then you should consider scrapping it (or remeasuring it, if that is possible.) If you can explain why a data point is not representative of the experimental conditions you set out to measure for it, kick the data point out.

Sometimes, data appears skewed, only because it needs to be transformed. You can transform data by performing some mathematical trick on all of the data points. The goal of your transformation is to fit the transformed data into a normal distribution.

- Logarithms can make trends over orders of magnitude look like linear trends. The most common approaches are to use log base 10 or the natural log of a variable.
- If a histogram of your data looks skewed to the right (i.e., there are a pile of values squished up on the left), try taking the square root of your values.

Statistics

Table 1: Parameters and estimates.

Variable or Estimate	Name
$\bar{x} = \hat{\mu}$	Sample Mean
μ	Population Mean
n	Sample Size
a_i	Value of individual i
$s^2 = \hat{\sigma}^2$	Sample Variance
s	Standard Deviation
σ^2	Population Variance
σ_x	Standard Error
p	Probability of Success
$\hat{p}$	Sample Proportion
$\sigma_{\hat{p}}^2$	Estimate of Variance of Binomial Data
$\sigma_{\hat{p}}$	Estimate of the Standard Error of Binomial Data
MoE	Margin of Error
t_{n-1}	t-Test Statistic
z	z-Test Statistic
r or R^2	Pearson Correlation Coefficient

Many students will cringe at hearing the word "statistics," especially when they are being told to engage in it. Honestly, though, knowing how to use statistics is an amazingly useful skill. Even if you never do a science experiment again, understanding what statistics can do will help you be a better citizen. You will better understand opinion polling or the news of the latest study about a vaccine. Depending on your own talents and background, statistics can be hard, but you can overcome it. Or, you just find it incredibly easy, and you don't need the pep talk. Either way, let's talk about what statistics is.

The Oxford English Dictionary defines statistics as, "The systematic collection and arrangement of numerical facts or data of any kind; (also) the branch of science or mathematics concerned with the analysis and interpretation of numerical data and appropriate ways of gathering such data." That's a pretty good definition[1] for our purposes. Statistics, as we are using it here, has to do with how we arrange and interpret the raw data collected from scientific research.

Let us say that you want to know how tall oak trees are. You could tell me that the oak tree outside of your window is 12.3598 meters high. I'll believe you (well, maybe not to the tenth of a millimeter, but it could be true.) You cannot use that information to then say that all

[1] These days, people often use the word statistics (often shortened to stats) to describe an athlete's performance or the abilities of a fictional character in a video game, these are colloquial usages of the term, and those are often misleading when they are applied to science. When a scientist uses the word theory, for example, he or she is referring to an idea that has been repeatedly tested and modified. It's something that a scientist can call true to the best of our knowledge. It's testable and has withstood those tests. In common parlance, a theory is more like a hunch, someone's guess at what might be true. Beware of people who apply common usage definitions of words when they are in a scientific context and vice versa.

oak trees are 12.3598 meters high. That same one wasn't that tall last year. Then, you tell us that you cored the tree and counted the rings. It is 34 years old, you say. Therefore, all 34-year-old oak trees are 12.3598 meters high. Again, absurd. Trees, as people, vary in height according to many factors, and we should account for those. Besides, can there possibly be that many significant digits[1].

> This statistics guide is not intended to be about the theory of statistics or even remotely substitute for a statistics class. Instead, we hope to arm you with some tools and tests that you can use to interpret your results. Hopefully, as you read this, you will find an appropriate tool to use.

We should instead consider all of the 34-year-old oaks in the area, right? Well, that gets closer, but when we are talking about these things, we need to consider things that we can compare. The word "oak" refers to all of the species in the genus *Quercus*, of which there are hundreds, some of which are shrubs and not trees. We'll probably want to narrow our query to one species. Also, a tree growing in a swamp will do better or worse than one growing on a hill, depending on its preferences for water. We'll need to account for that. Shade, too. The list goes on. We will need to figure out how tall oaks of those types are. Still, even if we accounted for all of those variables, we would still find that the oaks came in different sizes. What we need is an average!

Means and Variation

There are different sorts of average. The main ones we will refer to here are the mean, the median, and the mode. The **mean** (technically, the arithmetic mean, as there are other sorts of means) is the sum of all of the values of interest, divided by the number of things you measured. Here it is mathematically:

$$\bar{x} = \hat{\mu} = \frac{1}{n} \times \sum_{i=1}^{n} a_i$$

Where $\bar{x}$ is the mean of your sample, $\hat{\mu}$ is the approximate mean, n is the sample size and a_i is the value measured for individual i. Got that? Yeah, we're going to have to talk about that little mess. $\bar{x}$ ("x-bar") is the mean of your sample, the best estimate of the true mean, μ ("mu") which we call $\hat{\mu}$ ("mu-hat," as we'd say; it's the Greek letter mu with a hat on it.) It's

---

[1] Significant digits, or significant figures, are the digits in a number that you can trust not to be too precise but inaccurate. As a student, I had trouble understanding them at first, but the truth of the matter is that when you multiply or divide numbers, you will often end up with many more digits than you can expect to accurately reflect what you are talking about. It's not so much that they are wrong, but they can be misleading about how accurate your measurements could possibly be or can simply distract the reader from the value of the number. That is, long numbers (with lots of digits) aren't necessarily the same as big numbers. While there are rules for determining significant digits, you should consider how accurate your instruments are and how readable your text is. Also, you should not have more significant digits than your instrument can precisely measure.

approximate, because we don't have every single example in nature of whatever it is we're measuring. We're sure you get the 1/*n* bit, as that's just one divided by *n*. The weird E looking thing is the capital Greek letter sigma which we use to notate a sum. The a_i bit is what is being totaled. We measure value (let's stick with height for our example) *a* for individual 1, and we call that a_1. The $i = 1$ term tells us to start at 1. For individual 2, the height is a_2, and we continue until we get to the last one measured, height value a_n. You already knew how to do a mean, though, so why did we make it more complicated? We'll be using similar notation going forward, and you may see similar things in books and scientific papers. It's good to be caught up with the basics.

Let's do an example. We've gone out and measured 10 34-year-old *Quercus alba* (white oak) trees that were growing in identical conditions. Their heights were:

Table 2: Oak tree heights.

Tree ID	Height (meters)	Tree ID	Height
1	12.1	6	12.1
2	12.4	7	11.9
3	10.9	8	11.2
4	11.8	9	11.8
5	12.0	10	12.1

To create the average, I take the equation:

$$\hat{\mu} = \frac{1}{n} \times \sum_{i=1}^{n} a_i$$

and we substitute $n = 10$ and all of the heights to get:

$$\hat{\mu} = \frac{1}{10} \times \sum_{i=1}^{10} a_i$$

$$= \frac{1}{10} \times (12.1 + 12.4 + 10.9 + 11.8 + 12.0 + 12.1 + 11.9 + 11.2 + 11.8 + 12.1)$$

$$= 11.83 \cong 11.8$$

The mean is most meaningful when your data has a normal distribution (Figure 5). That is, most of the values are close to the mean, and less frequent values are much larger or smaller than the mean. We usually assume that data has a normal distribution, but that might not be the case. If you are using Excel, the Average function will calculate the mean of a range of cells.

The next most frequently encountered kind of average is the **median**. If you arrange all of your data from the smallest value to the largest value, and then you pick the middle value, you have a median. If you have an even number of values, then calculate the mean of the two middle ones. It's pretty easy to determine, and spreadsheets will readily generate your median for you (the Median function in Excel is very straight-forward.)

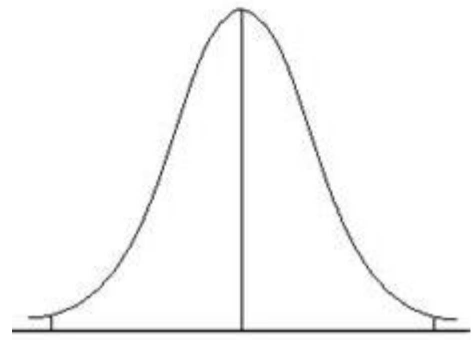

Figure 5: Data in a normal distribution. The x-axis represents the values for the data, and the y-axis represents the frequency of the data.

Why would you want a median? The classical example is income. In 2004, the United States Census Bureau calculated the mean American household income as $60,528. However, the median income was $44,389. Why

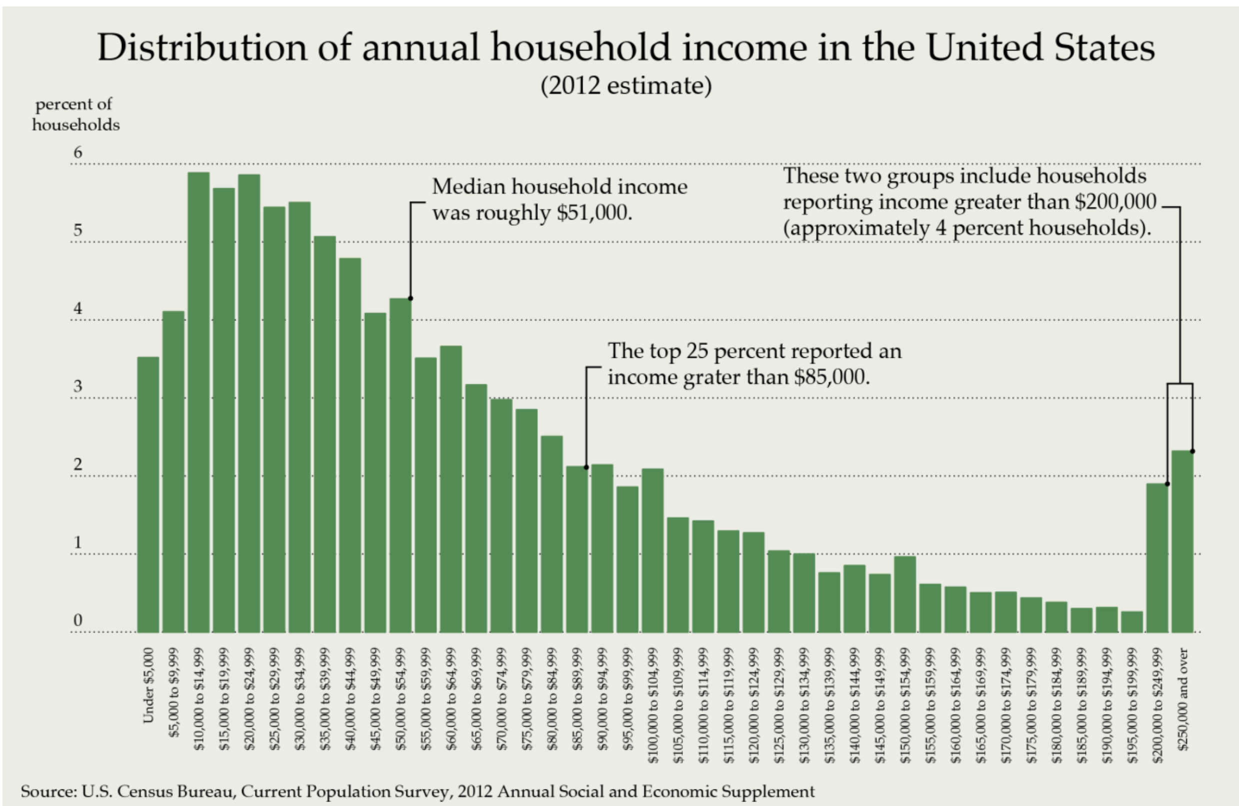

Figure 6: Annual income distribution in 2008, as estimated in 2012. This is not a normal distribution (though it is very typical for an income distribution.) Note that the final 2 columns are relatively large, because they contain wider ranges of incomes.

are these numbers so different? Income does not follow a normal distribution (Figure 6[1]). Both the mean and the median are valid ways to describe a normal or average individual, but their meaning depends on context.

The final way that we will consider the average is the **mode**. The mode is simply the most common number that comes up in your sample. If you consider the white oak height data that we made up, you will see that the number 12.1 comes up 3 times. That is our mode. If there is a tie for most common number, we have two modes. If there isn't a number that comes up most frequently, then there is no mode. We can round our numbers to get a new mode, but you will definitely want to report doing so in a paper. However, few papers use modes, so you will need to justify why you think it is interesting.

Just because you have an average, that doesn't mean that you have described your data completely. Even ordinary individuals will be different from the mean. They vary. There are several related measures of this phenomenon: variance, standard error, and standard deviation. Let's talk about **sample variance**[2] first. You estimate variance of your sample this way when we have a normal distribution:

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

Where s^2 is the sample variance, n is the sample size, x_i is the value of individual i, and $\bar{x}$ is the sample mean. Once again, we have a summation, but this time, we will be doing something a little more involved. Each part of the sum will involve subtracting the sample mean from the value of individual i and squaring the result. For our oak experiment (Table 1), it will look like this:

$$s^2 = \frac{1}{10-1} \times [(12.1-11.8)^2 + (12.4-11.8)^2 + (10.9-11.8)^2 + (11.8-11.8)^2 + (12-11.8)^2 + (12.1-11.8)^2 + (11.9-11.8)^2 + (11.2-11.8)^2 + (11.8-11.8)^2 + (12.1-11.8)^2] \cong 0.205$$

Sample variance is not necessarily useful when we analyze data, except that it is used in other calculations. It has units, but those units are square units relative to the mean, so if you are measuring the length of crystals in an evaporation experiment, your lengths could be in millimeters (mm), but variance would be in square millimeters (mm^2), even though it is not a reflection of area. A small variance indicates that the individual values in the sample are very

---

[1] Figure 6 is a **histogram**. It's a kind of bar graph in which individuals are pooled into groups that fall within a range. Instead of counting all of the people who made $45,322.37 per year into one tiny group and repeating that for every income level down to the cent, we can group similar values into a range.

[2] We distinguish sample variance from population variance. Remember that most studies use a representative sample of the population instead of the whole population.

similar to the mean, whereas a large value indicates that they are more widely distributed (Figure 7).

If variance isn't immediately useful, how do we use it? The first thing that we can do is calculate the **standard deviation**. The equation for that is:

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

Whoa! That's even more complicated. Yes, but you could also say that standard deviation is just:

$$s = \sqrt{s^2}$$

That's easy. Since standard deviation (s) is in the same units as your sample measurements and mean, you can use it to help you analyze your data.

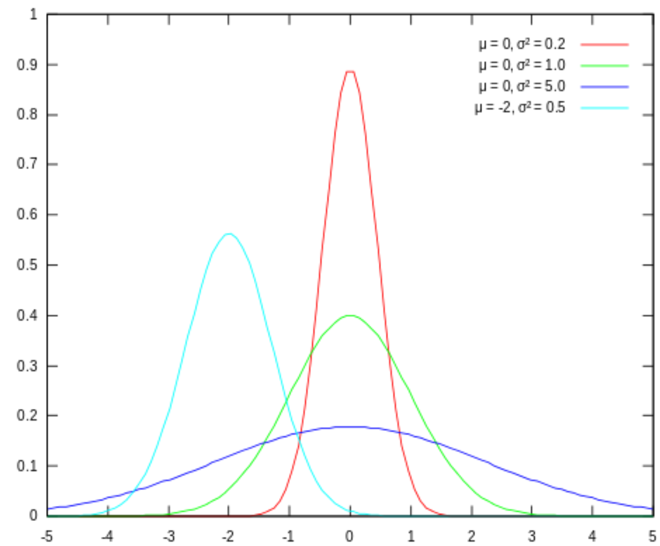

Figure 7: Smaller variances (like the tall red curve in the center) indicate that the individuals have more similar values. Larger variances (like the flat blue curve that spans the graph) indicate that individuals have less similar values. From Wikimedia Commons user MarkSweep.

Not all data are recorded as continuous values as we have been using, let's consider **binomial data**. Binomial data values are typically represented as being 1's and 0's, and the probability of success of a sample or population is a value from 0 to 1. The most common binomial information you see is in the form of percentages, as you see in public opinion polls, as in 47% of people approve of the President's plan to increase kumquat consumption. Any time you say that X% of something is Y, you are using binomial data. On your record sheets, you have recorded 0 or 1, dead or alive, in favor or opposed, brown or not brown, etc.

The variables for binomial data are slightly different from those used for continuous data. With binomial data, the theoretical mean[1] is:

$$\mu = np$$

Where p is the "**probability of success**," a term referring to the fraction of the time a condition is true. For coin flips, the probability of success for heads is 0.5, since half the time, heads comes up. If we flip a coin 100 ($n = 100$) times, then we expect the mean μ to be 50. However, we can estimate p with the **sample proportion**:

---

[1] Theoretical mean is another word for the mean of the population.

$$\hat{p} = \frac{X}{n}$$

Where X is the number of successes. It is the closest equivalent of an approximation of the population mean μ. If we actually flipped the coin 100 times, we might get 52 heads by chance. In that case, $\hat{p} = {}^{52}/_{100} = 0.52$. In this case, we are estimating the probability of success by measuring the sample proportion ($\hat{p}$).

The variance for the binomial distribution is:

$$\sigma^2 = np(1-p)$$

The standard deviation is thus:

$$\sigma = \sqrt{\sigma^2} = \sqrt{np(1-p)}$$

The variance of your sample proportion ($\hat{p}$) can be computed with the equation:

$$\sigma_{\hat{p}}^2 = \frac{\hat{p}(1-\hat{p})}{n}$$

Keep in mind that these figures are based on the sample proportion and not the number of successes. That is, it is based on the fraction and not the count. Notice that the larger your sample size (n), the smaller the variance of the sample proportion and the standard error of the sample proportion.

Confidence Intervals

When you calculate means and variances, you describe important properties of your data, but that does not mean that you can make any useful claims about it. You have often heard the phrase, “give or take,” as in, “It’s about 50 pounds, give or take.” That phrase can describe an approximate mean with wiggle room for uncertainty. When we use statistics, we can assign meaning to “give or take” with **confidence intervals**. These are minima and maxima[1] of estimated means. That word confidence is very important. We can only be some percent confident that our real mean or probability of success falls between two numbers. Usually, we use the cutoff of 95% confidence, because in order to be 100% confidence, we would have to settle with, “It’s a number.”

Let’s start with continuous data. You will need to calculate the **standard error** (σ_x, the Greek letter sigma with an x subscript):

---

[1] Minima is the plural of minimum, and maxima is the plural of maximum.

$$\sigma_x = \sqrt{\frac{s^2}{n}} = s \times \sqrt{n}$$

If your data has a normal distribution, then 68% of all of the individuals will have values between $\mu \pm \sigma$, or approximately $\bar{x} \pm \sigma_x$. Likewise, 95% of the individuals will have values between $\mu \pm 2\sigma$, or approximately $\bar{x} \pm 2\sigma_x$ (Figure 8). In other words, you can often say that 95% of the individuals in a population are two standard errors away from the mean. Thus, the confidence interval is $\bar{x} \pm 2\sigma_x$. Now, consider two averages, $\bar{x}_1$ and $\bar{x}_2$. Can we say that they are different from each other? Using confidence intervals, we can check their lower and upper bounds. Let's say that $\bar{x}_1 < \bar{x}_2$. If $\bar{x}_1 + 2\sigma_{x_1} > \bar{x}_2 - 2\sigma_{x_2}$, then their confidence intervals overlap, and we cannot conclude that they are different with 95% confidence.

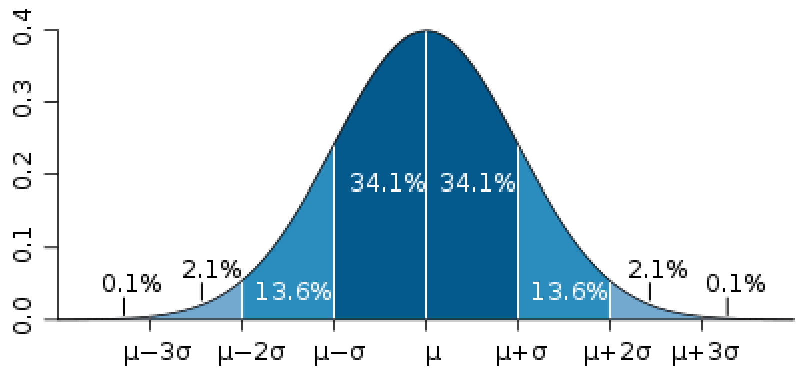

Figure 8: Standard deviation around a mean in a normal distribution. From Wikimedia Commons user Ainali.

Error bars in graphs are very important applications of confidence intervals. You can use them to visually display how confident you are of your data, and your audience will be able to tell just how different your various means are from each other. You can use these in bar charts or line charts (Figure 2 and Figure 4, for example.)

When we use binomial data, the confidence interval is based on the **margin of error**:

$$MoE = 1.96 \times \sigma_{\hat{p}} = 1.96\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

We are using the number 1.96 for what is called a *z*-value, a concept that is beyond the scope of this guide[1]. Once you have the margin of error, you can say that you are 95% confident that the population proportion p is between $\hat{p} - MoE$ and $\hat{p} + MoE$ (or, $\hat{p} \pm MoE$). If you have a polling sample of 100 of your peers, and 45 of them said that they preferred country music over rhythm and blues then you could state with 95% confidence that the actual proportion of like-minded youth who like the twang:

[1] You can substitute 2.33 for the *z*-value to get 98% confidence or 2.58 to get 99% confidence.

$$0.45 - 1.96\sqrt{\frac{0.45(1-0.55)}{100}} = 0.35 \text{ and } 0.45 + 1.96\sqrt{\frac{0.45(1-0.55)}{100}} = 0.55$$

That is a big range, but if you increase your sample size to 1000, then it shrinks to 0.42 and 0.48. You can report that as "0.45 ± 0.03".[1]

Hypothesis Tests

In statistics, the term hypothesis has a somewhat different definition than it does in the sciences, but it is rather similar. As in science, a statistical hypothesis is a testable prediction, but it is also a mathematical expression, and we can perform a **hypothesis test** to help analyze our data. A statistical hypothesis might be something like, "The average length of *Amythas* spp. (common Asian earthworm) found in Oklahoma County is a different size than the length of common Asian earthworms collected in Payne County;" "The average specific gravity of selenite crystals from the Great Salt Plains State Park is the accepted specific gravity for selenite minerals: 2.3;" or, "The average falling time for these bowling balls is the same." In further analysis, we will talk about those being **null hypotheses**. The shorthand for the null hypothesis is H_0. In opposition to the null hypothesis, we also have the **alternative hypothesis** (H_1 or H_A). While the null hypothesis is always $\mu_1 = \mu_2$ (or "mean length of earthworms in Oklahoma County = mean length of earthworms in Payne County," "specific gravity = 2.3," or, "bowling ball 1 falling time = bowling ball 2 falling time," from the examples above), the alternative hypothesis can be:

$\mu_1 \neq \mu_2$ or $\mu_1 > \mu_2$ or $\mu_1 < \mu_2$

To test these hypotheses for continuous data, we can use *t*-tests. We start by calculating the *t*-test statistic:

$$t_{n-1} = \frac{\bar{x} - \mu_0}{s/\sqrt{n}}$$

Where t_{n-1} is a value from what is called the *t*-distribution[2] with $n - 1$ degrees of freedom; $\bar{x}$ is the mean of your sample; μ_0 is what you have hypothesized μ to be (the length of the earthworm or the specific gravity of selenite from the examples above); s is the standard deviation of your sample; and n is the sample size of your sample. The term "degrees of freedom" refers to the number of independent ways that the numbers may vary given the constraints of the system. For variance and standard deviation, we use $n - 1$ **degrees of freedom**.

---

[1] This formula for the margin of error is weaker when $\hat{p}$ is near 0 or 1. One solution when you have small or large proportions is to use the Wilson score interval.

[2] The vagaries of the *t*-distribution are beyond the scope of this guide. In essence, it is a normal distribution based on uncertainties arising from estimations. We use it when we are making claims based on data.

Next, you will need to use this to calculate or look up a *p*-value. You can either use a *t*-table, as is tradition[1], or you can use software like Excel to calculate it. The ***p*-value** is a number between 0 and 1 that you can use to accept or reject the null hypothesis (that $\mu = \mu_0$.) If your *p*-value is more than 0.05, then you can claim that it has passed the 95% confidence test, and you fail to reject the null hypothesis. When we find that two means are different from each other using a *p*-value, we say that their difference has **statistical significance**.

The way in which the *p*-value is checked depends on the your alternative hypothesis ($\mu_1 \neq \mu_2$ or $\mu_1 > \mu_2$ or $\mu_1 < \mu_2$). When the alternative hypothesis is that $\mu_1 \neq \mu_2$, the *p*-value is derived from the area that makes up 95% of the *t*-distribution (Figure 9). The left and right "tails" of the *t*-distribution each comprise of 2.5% of the remaining area, and if the *t* that you calculated falls in one of those two areas, then you can reject the null hypothesis and declare your mean to be different from the hypothesized mean. For the alternative hypothesis that $\mu_1 > \mu_2$, you will instead need to use one tail, and that tail makes up 5% of the area under the right curve. The Excel guide includes instructions for differentiating among the different alternative hypotheses.

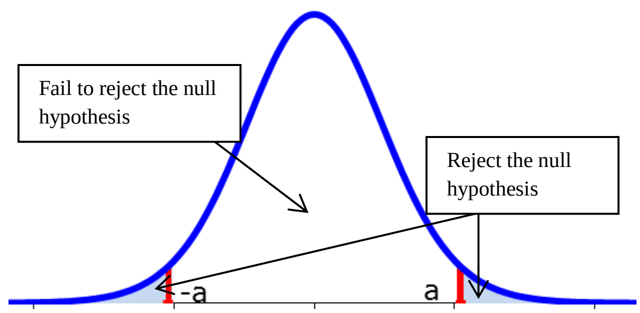

Figure 9: *t*-distribution with decisions for the alternative hypothesis that $\mu_1 \neq \mu_2$.

We use a *z*-test for determining the *p*-value for binomial data[2]. In that case, the null hypothesis is $p = p_0$, that is that the proportion of the population for some trait is p_0. We will then calculate:

$$z = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}}$$

Where $\hat{p}$ is your sample proportion, p_0 is the hypothesized proportion, and n is the sample size. We can then use Excel[3] to determine the *p*-value.

Very often, we do not have some automatic expected value of μ or p. Instead, we are comparing two or more samples to each other, i.e., a control and a treatment. You will want to

---

[1] Textbooks still insist on using tables for all sorts of things, probably because you can't use the technology of books to do these things. Scientists these days don't use tables. We have computers to do the same thing even better. If you take a statistics course in college, expect to use *t*-tables.

[2] Remember that binomial data is stuff like yes/no or 1/0.

[3] See the Excel guide on the website.

know if they are different from each other. If you have two samples, you can use a paired *t*-test. You will have to choose a formula, depending on your results as shown in Table 3.

Table 3: How to calculate *t*-statistics given different conditions. Notice that we use a *z*-statistic instead of a *t*-statistic when we calculate with proportions.

Conditions	*t/z* calculation	*s*	degrees of freedom
Equal sample sizes and equal variances $n_1 = n_2$ and $s_{x_1}^2 = s_{x_2}^2$	$t = \frac{\bar{x}_1 - \bar{x}_2}{s_{x_1x_2}}$	$s_{x_1x_2} = \sqrt{\frac{1}{2}\left(s_{x_1}^2 + s_{x_2}^2\right)}$	$2n - 2$
Unequal sample sizes and equal variances $n_1 \neq n_2$ and $s_{x_1}^2 = s_{x_2}^2$	$t = \frac{\bar{x}_1 - \bar{x}_2}{s_{x_1x_2}\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$	$s_{x_1x_2} = \sqrt{\frac{(n_1 - 1)s_{x_1}^2 + (n_2 - 1)s_{x_2}^2}{n_1 + n_2 - 2}}$	$n_1 + n_2 - 2$
Unequal variances $s_{x_1}^2 \neq s_{x_2}^2$	$t = \frac{\bar{x}_1 - \bar{x}_2}{s_{\bar{x}_1 - \bar{x}_2}}$	$s_{\bar{x}_1 - \bar{x}_2} = \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$	$\frac{(s_1^2/n_1 + s_2^2/n_2)^2}{(s_1^2/n_1)^2/(n_1 - 1) + (s_2^2/n_2)^2/(n_2 - 1)}$
Proportions	$z = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\hat{p}(1 - \hat{p})\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$	Note that $\hat{p}$ is the proportion of the combined sample.	

Generally, the Excel Guide to Statistics will show you how to do these very simply in Excel.

Pearson Coefficients

Often times, when we look at data, we look for correlations. Data correlate when they have some sort of relationship with each other. Maybe they both ascend with each other, or when one variable is larger, the other is smaller. **Pearson coefficients** are the primary and simplest means of analyzing correlations. We start by calculating the correlation coefficient:

$$r = \frac{1}{n-1}\left(\frac{\sum_x \sum_y (x - \bar{x})(y - \bar{y})}{s_x s_y}\right)$$

For each sample, you will need the value for two variables, x and y. You will need to calculate the means and standard deviations for all individuals for each variable to get $\bar{x}$ and $\bar{y}$, as well as s_x and s_y. For each pair, you will need to calculate $\frac{(x-\bar{x})(y-\bar{y})}{s_x s_y}$, sum those, and then divide by $n-1$. This whole process can be done in Excel with a simple function. The resulting correlation coefficient is very useful. Numbers near 0 indicate weak or poor correlation. If r is close to 1, then you are seeing a positive correlation, so that as one variable increases, the other decreases. When r is nears -1, there is a negative correlation, indicating that as one variable gets larger, the other one gets smaller (Figure 11).

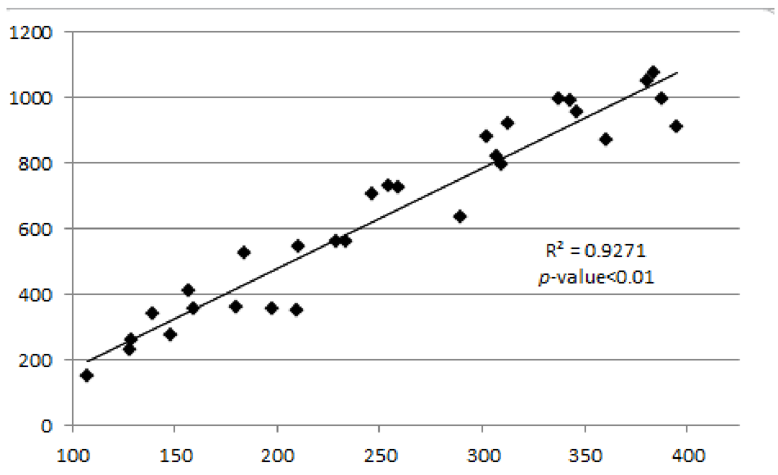

Figure 10: Scatter plot with trend line, R^2, and *p*-value.

Keep in mind that this is only a metric for linear relationships, those that behave as lines. There are more sophisticated tests for other sorts of relationships, but those are beyond the scope of this document. If you think that there is another sort of relationship, you can attempt to correlate transformed data. For example, if you think that $x \sim \log y$, then you should replace y with $\log y$.

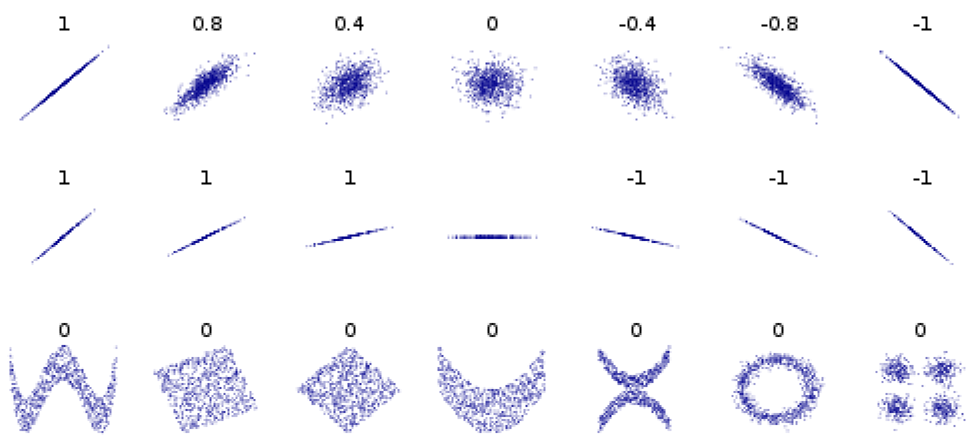

Figure 11: Pearson coefficient of various scatter plots. The Pearson coefficient is primarily used to determine how linear sets of paired data are to each other. Image courtesy of Wikimedia user DenisBoigelot.

Pearson coefficients are frequently reported as R^2 instead of r. Funny enough, $R^2 = r^2$. Keep in mind that because it is a square number, R^2 is a value between 0 and 1 and not -1 and 1, like r is. You may also generate *p*-values from r values. To do so:

$$t = r\sqrt{\frac{n-2}{1-r^2}}$$

The *t*-test involves $n - 2$ degrees of freedom.

Other Statistics Tests

Analysis of Variance (ANOVA) is a set of tools that can be used to compare multiple population means as well as multiple variables for the same populations. Think of it as a beefed up, caffeinated *t*-test. If you have lots of test populations and lots of variables, then you may want to look into using ANOVA to analyze your data. If you have multiple variables, then it can also be used to determine if there are interaction effects, situations in which the effects of the two variables are worth more (or less) than the sums of their individual contributions.

Chi-squared tests (X^2) tests can be used to find out if a set of categorical data fits expectations. If you have can categorize individuals into certain kinds of things, and count how many fit into each category, you can test whether those counts are similar to what you would expect. For example, you could collect flowers and categorize them by color. You might expect that all colors would have equal counts, so you could compare your counts and see if they are statistically different numbers from the expectations. Chi-squared tests are frequently used in genetics studies.

If it is important to calculate the formula for a regression line (as seen above with the Pearson coefficients), then you can use the least squares method. It is relatively straightforward, and it can be easily calculated in Excel. If you have multiple variables that you are comparing in regression analysis, you can calculate regression lines in multiple dimensions using multiple linear regression.

There are hosts of other important tests that have been designed in the last 150 years. Information about them are scattered around the internet and in textbooks. If you don't know what to do with your data or are unsure how to proceed, then you should absolutely ask a teacher, a college professor, or even a discussion forum on the internet.

Interpreting Results

You have done your statistical analysis, and you can show that two things are likely related (or not.) You just did the mechanical part, the boring part, the thing that does not yet have the inspiration of human spirit. Here you are now, results in hand. What in the world do you do now? You have to interpret what you have found. You have to take the results and tell yourself and others what it means, what the implications are, and how this changes our understanding of the world.

Statistical Significance

If you performed a test that told you that two or more means were statistically different, then you need to explain why you think that they are. Were you expecting them to be different? Were you expecting them to be the same? How different are they[1]? If your tests did not show statistically significant differences, then what do you make of them?

These questions can dog professional scientists, and honestly, there is no hard and fast rule that you can use. If you have a *p*-value of 0.049 (just under the 0.05 threshold), then you have shown statistically significant differences according to conventions, but you have barely done so. When you interpret that difference, you need to acknowledge that it is a weak one. Perhaps, more replications of your experiment will clear this problem up, but you may not have the time or resources to do more experimental work.

The bottom line is that you can make a claim that your mean or proportion is different with statistical significance. You will have to interpret what that means or implies. Be sure to relate what you found with what others have found in similar systems or with other measures in your system. If you found that the selenite crystals in a local park have greater specific gravity than is normal, you will need to speculate and perhaps do follow-up chemical tests to determine why it is so different.

Correlations

If you did correlation tests, then you will need to interpret what those correlations mean and suggest what other tests need to be considered. In general, correlations show that two measurable things vary with each other or against each other (or not at all.) You have all heard the phrase, “Correlation isn’t the same as causation,” or something to that effect. This statement is true, but correlation can imply causation.

[1] Since you are working on a high school student’s budget, if your results are significantly different, then they are probably meaningfully different. In studies with many replications and huge sample sizes, statistical differences can be found that are so minor in real world terms that they don’t really imply any worthwhile conclusion. This is most common in gigantic medical studies.

When two things (A and B) correlate with each other, we can come away with four different conclusions: A causes B; B causes A; something other than A or B simultaneously causes both of them; or A and B correlate due to coincidence. Scientists are not interested in the last conclusion, except that they need to rule it out. Let's start with the first two conclusions, A causes B and B causes A. You will have to use reason and common sense, along with whatever evidence you can find, to tease out which is true[1].

Consider Oklahoma's recent spate of earthquakes. Starting in 2009, the frequency of earthquakes with a magnitude of 3.0 or higher has gone up dramatically (Figure 12). As many have observed, this date is the approximate date that waste injection well use in fossil fuel extraction also increased[2]. The two things correlate. Let's compare A causes B to B causes A. Would it be reasonable to say that injection wells could cause earthquakes? Waste injection wells push liquid deep underground at high pressure, which geologists surmise could alter the state of underground faults, thereby initiating earthquakes. Could we suppose that earthquakes cause injection well use? That would be absurd, of course. No one feels the ground shake and then decide to go and drill because of that.

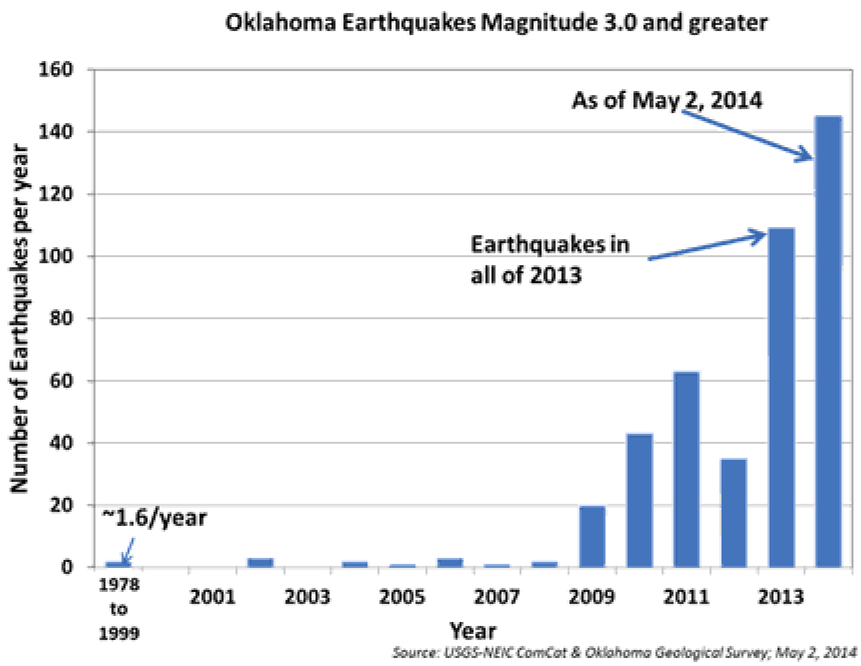

Figure 12: Recent Oklahoma earthquake frequency.

Could something else cause both the recent increase in earthquakes and increased waste well injection? That seems almost as absurd as earthquakes causing drilling, so we should rule that out. The final possibility is that these two things are mere coincidence. That is not as

---

[1] You know, both could be simultaneously true. This would be an instance of feedback. Consider one plant shading another plant. They both need sunlight to grow, and the bigger plant is getting more of it. Because of that, it can grow faster, continuing to shade the smaller shaded plant. Thus, the shaded plant's small size is preventing it from getting sunlight (small size is causing the plant to be shaded by larger plants), and the shade is stunting growth (shade is causing the small size.) Plants have means of escaping these situations, of course, but this might be true for some plants some of the time.

[2] Some modern fossil fuel extraction has a lot of waste water that needs to be disposed of. Drillers simply inject it deep underground to prevent it from harming the environment.

absurd, and it should be seriously considered. The joint study performed by the United States Geological Survey and the Oklahoma Geological Survey concluded that it was unlikely to be coincidence after more rigorous statistical tests and the use of information from other locations, but you should always consider the possibility that correlations are coincidence.

Frequently, two observed trends will correlate, because some other factor is driving them. Sour foods can lead to tooth decay, if you're not careful. How does sour hurt your teeth? Well, you taste acidic foods as being sour, and acids can wear away at tooth enamel. Both the sour flavor and the tooth decay are caused by acidity. Let's try another. Sometimes, silly hats and fireworks co-occur[1]. While silly hats could lead to fireworks, it might just be that it's New Year's Eve. People like to wear silly hats on the holiday, and people like to display fireworks on certain ones, so both silly hat wearing and fireworks are caused by people's reaction to New Years Day. Okay, one more. Alligators and cypress trees correlate. Does one cause the other? No, they both prefer swamps. When you look at your data, make sure that you find instances of these things.

Now and then, correlation is total happenstance. Two things have positive or negative correlations, and it has absolutely nothing to do with them relating to each other. Sometimes, it's really easy to tell when that is true (Figure 13)[2]. As with finding common causes for correlating variables, uncovering when they are simply happenstance is also an exercise in common sense and reason. If you cannot explain why two things are related, then you should assume that they are not. In fact, designing experiments to test whether two things are actually related and possibly why is a great enterprise in-and-of-itself.

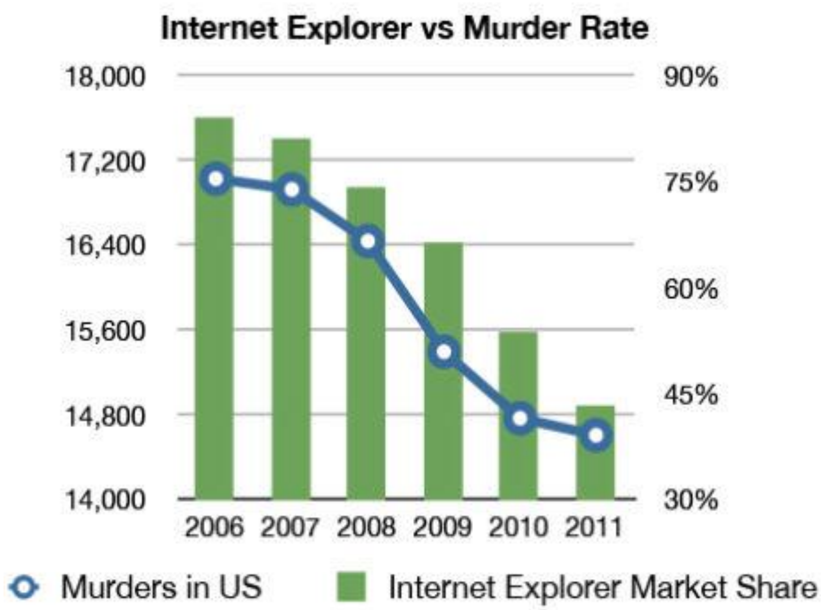

Figure 13: Correlation sometimes has nothing to do with causation. Say what you will about Microsoft Internet Explorer, but it would be unreasonable to assume that it caused murders or that few murders caused people to stop using the software. Source: Gizmodo.com

[1] Scientists often use the words "occur" and "co-occur" in ways that are kind of funny to non-scientists. You might say that wolves live in the forest, but a scientist would write in a paper that they occur there. If two things happen together or live together, we will say that they co-occur. This often implies correlation.

[2] Here are lots of spurious correlations that you can waste your time justifying.

Context

Science is not conducted in a vacuum. Your results relate to someone else's work, and you worked on something that interacts with something else. You need to put your results in the context of those things. Here, your research of the systems and methods you used in your experiment will pay further dividends. What did other authors say about their systems? What did they say about the results they got? Borrow[1] those ideas and make them your own.

There is not any one-size-fits-all advice for finding the context of your work. It is generally a matter of knowing why you researched, hypothesized, experimented, and analyzed. Here are some tidbits that we can provide you, though:

- If another experimenter used the same methods that you did on another similar system, and you share results, what did he or she conclude? Do you think that you can make similar claims?
- What have other researchers discovered as unique about the system that you are working on? Could they be relevant? Why or why not[2]?
- How will this work impact what we understand about other systems related to your system?
- Is there an impact on people[3]?
- Does anything make you wonder what is really happening? You could have found something unusual that you can encourage others to look into.

When science is done, you test a hypothesis to challenge a theory. A theory's resistance to those challenges is what makes it strong. Sometimes, when you have done your work, and you found no differences between treatments like you thought you should, you are tempted to despair. The lack of differences is called "negative results," and even the wording implies failure. You have to put that into context, too, however. Do those negative results support or challenge the theory? Maybe they imply that the theory is not as strong as others think it to be. Depending on how well supported that theory is, you may have shown that it does not apply to your system, to a class of systems, or to anything at all. After all, if only one experiment in all of

---

[1] We say borrow and make your own, and that's important for two reasons. If you want to be good at what you are doing, you have to internalize the ideas you are using and coming up with. This is a matter of practice and talking to other people. When you have aspects of your work internalized, you can really shine. Also, it will help you to avoid **plagiarism** when you write. Take note of where you got ideas, but also having them as your own ideas will help you to write original text, and that is the essence of avoiding plagiarism.

[2] Generally, if something isn't relevant, you don't need to report that it is. There are cases in which people have assumed something to be relevant, and research showed that it was not. Maybe you found something like that. If so, you'll want to report it.

[3] Not all science is immediately relevant to human welfare, and that's okay. If you can make a claim that yours is, though, it tends to get folks' attention. As a dispassionate scientist, you shouldn't care about that, but as a dispassionate scientist who needs to get a grant proposal accepted in order to get a paycheck, you probably should care. Just saying.

history has supported the theory, you may have just cast enough doubt on it to disprove it. On the other hand, if it has been supported by millions of prior experiments and has been shown to explain phenomena time and time again, it's probably your experiment.

Negative results are difficult to make conclusions from, but they are an underappreciated part of science. It's hard to get them published in the more prestigious journals, and many scientists don't bother submitting papers about them to even the least of journals[1]. That's a shame, really. If nothing else, it prevents other scientists from trying the same experiment when they could be doing something else. As a high school student, you should be proud of your results, even if they are considered negative. You might have discovered something new anyway.

---

[1] Peer-reviewed scientific journals represent the most important way that scientists communicate their results. They are basically magazines (more online entities these days) that contain articles describing what scientists did and what they found. Each article has been examined by other experts in the field and vetted for its rigor and importance. Very prestigious journals such as Nature, Science, and Proceedings of the National Academy of Sciences have a very high bar for what they consider important. Every field has its own set of journals ranging from very prestigious to those that accept any paper, as long as the science was rigorous enough.

Communicating Your Results

The final step in science is communicating your findings. While the public thinks of scientists as people who work in labs and who write confounding equations on chalk boards, scientists are also writers and speakers. You can do all of the lab work, field work, observations, calculations, research, and analysis you want, but if you never get your work out for, you may as well never have done your research. Science fairs are excellent ways to practice this aspect of science, and we hope that you enjoy communicating your results. We will discuss the primary ways that scientists communicate: with writing, with posters, and with presentations.

Writing

Good writing is a skill that you can use in many walks of life, and even if you never really use it professionally, you can use it personally and even romantically! How can science writing be romantic? Mostly, it can't, but writing is writing, and being good at one kind can help you be good at other kinds. In your English classes, you should be learning all of the rules of grammar, punctuation, and spelling. Those are reasonably universal, and even when you write poetry or the sloppy slang of your fictional cowboy, it really helps to know what rules you are actually breaking (since you can break them better that way.) Just developing the swagger to put words onto your page is really helpful. Learn how to write, as it is a great skill.

Before digging into the nitty-gritty of science writing in particular, it is worthwhile to relay some advice to you about writing in general. We're not going rehash, subjects, predicates, gerunds, and commas, but you should remember all of those rules. It saves time in the editing process, and it makes you look smart. Having graded some college papers, we can tell you that students make very different kinds of mistakes when they write. While one student will frequently use incomplete sentences, another will have run-on sentence after run-on sentence. You have to know your strengths and your weaknesses. Remedy those grammatical weaknesses, and you will save yourself a lot of grief.

As is true with many things, practice is necessary to develop as a good writer. While you may only produce a few papers for your classes every semester, you should consider come other ways to practice.

- Write for fun and profit. Make up short stories, poems, or opinion articles. Even if they're not good, it's still practice. Maintain a blog or a gallery of your work. As you write, you will understand writing more, and if you do it for fun and not in fear of teacher's red pen, you can better associate writing with pleasure.
- Use good writing in text messages and on social media. Instead of abbreviating everything and not caring how your sentence comes off, make a solid effort to use proper grammar, spelling, and punctuation with your peers. Your generation writes a prodigious amount of text, so you have no excuses for writing it poorly. If your friends make fun of

you for doing so, just tease them for their bad spelling and move on. Writing well should be a habit.

- Read. Read books, magazine articles, and blog posts. Read fiction, poetry, science, and history. When you read, have opinions not just about the content of what you are reading but the quality of the authorship. Copy styles you like and understand why you don't like what seems bad to you.
- Read a few writing guides. John's very favorite is "Politics and the English Language" by George Orwell, but there are lots of others out there. The National Center of Biotechnological Information (NCBI) also has a guide to writing scientific papers.

When you write, you create a voice. Just as it is with a person's audible voice, there are lots of ways to describe a written voice[1]. You can also write in different voices depending on your audience and your intent. When we write scientific papers, we follow a very different style and present a very different voice than what we are writing with here. In this work, we are trying to be familiar and warm. We are avoiding too many unnecessary technical terms without sounding condescending. We can use the words "we" and "you" with abandon. Scientific papers are very different. In those, we have to be precise and very detailed. The writing is colder, but that's okay, because the purpose and the audience are different. It is more important for a scientific reader to very clearly understand *exactly* what we mean by something, so there is more jargon. The priorities are different, and you have to appreciate the different motives of the reader.

[1] Frequently, we talk about voice in writing in terms of active and passive. In the active voice, subjects do things. In the passive voice, things are done to the subjects. "Plants are taller with more fertilizer," is an active sentence. "Addition of fertilizer has resulted in taller plants," is a passive sentence. Avoid the passive voice. Embrace the active voice. Too many scientists use the passive voice in papers, and that just makes me sleepy. Don't make me sleepy.

Let's get down to brass tacks in discussing what you need to do when you write a scientific article. Overall, you need to be precise, explanatory, and concise. It needs to be clear what you did, why you did it, and what you found out. The whole point of the exercise is to explain all of the little details, but you should also be fairly brief. Like I said it takes practice. We recommend reading some good scientific articles to learn about this.

The standard scientific paper goes something like this: title, abstract, introduction, materials and methods, results, discussion, acknowledgements, and literature cited. Most papers aren't actually written in that order, though. Instead of going over them in order of when you see them in a journal or in your final paper, we will go over them in order that you should write them.

Start with the **introduction**. Even though it's not the first thing you see when you look at a paper, it's very easy to write it first. The introduction should cover some basic information about the system you are working on, what sorts of related things that others have done with your system, what sort of problem you are addressing, and a hint at what you found. For a science fair project (or most science writing in general) the introduction should be short. Most are 4 to 8 paragraphs long. You can write an early draft of your introduction before you begin your experiment and use it as a personal guide for how you will think about your project.

Here are some scientific papers that will be relatively easy to read based on their content and style:

- Curley LP, Hoseney RC. 1984. Effects of corn sweeteners on cookie quality. Cereal Chemistry 61: 274-278.
- Eide ER, Showalter MH. 2012. Sleep and student achievement. Eastern Economics Journal 38: 512-524.
- Cvetkovic C, Raman R, Chan V, Williams BJ, Tolish M, Bajaj P, Sakar MS, Asada HH, Saif MTA, Bashir R. 2014. Three-dimensionally printed biological machines powered by skeletal muscle. Proceedings of the National Academy of Sciences 111: doi:10.1073/pnas.1401577111.
- Lichti NI, Steele MA, Zhang H, Swihart RK. 2014. Mast species composition alters seed fate in North America rodent-dispersed hardwoods. Ecology 95: 1746-1758.

Please note that some of these are likely behind a paywall. To get ahold of them, you may need to contact a local college. Your teacher may be able to help.

Next, write the **materials and methods** section. Herein, you will need to describe what kinds of samples you are working with and where you got them. You will need to describe the conditions you used, the ways that you did your experiment, and any special methods of analysis you used. Materials and methods sections are unusual in that they are often written in the passive voice. This is because you are the one who did the things to the samples, and a lot of scientists avoid using first person writing when they write anything, but this is changing. Either, "50 mL of water was added to the solution, and pH was taken again," or, "We added 50 mL of water to the solution and remeasured pH," are usually acceptable. Check with any guidelines about this for your science fair before you continue.

Many papers include a conclusion section. Sometimes, it is required by the journal. A conclusion is typically 1 to 3 paragraphs that distill the most important points raised in the discussion. If you include one, keep it brief and poignant.

For the **results** section, you will need to include numbers, observation, and so on. You will need to include information about statistical significance, if that is important, as well as brief interpretations. You will likely reference your most important figures and tables in the results section as well. Keep your sentences simple, since you are only stating fact here and not in depth analysis. Like materials and methods, results sections are often broken down into subsections describing the results of different tests. Give each one a heading.

The **discussion** section is the coup de grace of your paper. In the discussion, you get to tell your readers what this is really all about. You will take the most important results and explain why they matter, what they could really mean, and what sorts of questions they unveil for future investigation. You will need to defend your claims here, and since science is about disproving theories, you will need to suggest some potential problems and pitfalls relating to what you discovered. Then, you will need to explain why they might not be valid or relevant for your research. In the discussion section, you will need to put your neck out and really stand for your results.

There is a little section that you can write any time you like, really, but it usually follows the discussion called the acknowledgements section. This is usually a very brief paragraph that thanks any non-author who helped in gathering materials, running experiments, giving key advice, creating graphics, etc. If you received outside funding, you should explain who gave it to you, and provide an agreement number, if that is applicable.

Combined, the introduction, materials and methods, results, and discussion comprise the body of your paper. Now, it is time to squish them all together into one paragraph, the **abstract**. While this is one of the first things that a reader will encounter, it should be one of the last things you write. The essential gist of the abstract is that it explains what is important about the problem you are addressing, how you approached the problem, what the headline results were, and what their implications are. Here's a basic approach to writing an abstract[1]:

---

[1] I have proposed a basic approach here, but occasionally, different journals, or perhaps your science fair guidelines, will specify how they want to abstract to pan out. Typically, the only requirement is a maximum word count, anywhere from about 150 to 300 words. Sometimes, the journal will dictate style and form.

1. Use one or two sentences to describe your system. Just highlight the aspects of the system that you are addressing.
2. Use one or two sentences to describe your methods. You don't need to include specific numbers like how many mL of water you added to each sample, unless that's what you were varying. You need to include kinds of materials you used and the ways you varied your methods.
3. State what your basic results are in 1 to 3 sentences. Only include headline numbers. If your results are more qualitative, then do not worry about the specific numbers.
4. State what this means to the field of study in one sentence. This final sentence encapsulates what it is you did, and it is the single most important sentence of your entire paper. What are your implications? What has this changed about how we understand the world? No pressure.

Here are some titles from an issue of the Proceedings of the National Academy of Sciences (PNAS):

- Some inconvenient truths about biosignatures involving two chemical species on Earth-like exoplanets.
- Radiometric ^{81}Kr dating identifies 120,000-year-old ice at Taylor Glacier, Antarctica.
- Switchable S = 1/2 and J = 1/2 Rashba bands in ferroelectric halide perovskites.
- Juvenile hormone regulates body size and perturbs insulin signaling in *Drosophila*.
- Field experiments of success-breeds-success dynamics.

Note that some of the titles are more accessible than others. It largely depends on whom the authors wish to reach.

Every sentence of your abstract should be an active and bold sentence. It should be clear and concise. Save your uncertainties for the discussion section and distill everything down to its fundamentals.

You have one last little bit to write: the title. Most authors write this last, though if a pithy or really effective title comes to mind at any point during the process, write it down, and it could work. There are a lot of approaches to writing titles, too. Some of them simply state the finding in a sentence. Others describe what the authors were investigating. You will find the occasional title filled with puns. While there is a lot of flexibility, you should carefully examine what it is that you are trying to do. Puns and humor get attention, but do they get the target audience interested in your topic? If you give away what you found, will anyone be interested in reading the rest of your paper? You just have to figure out what is most appropriate and what you think best reflects your work.

Academic writing, and scientific writing by extension, requires **literature citations**[1]. Here's the deal: in a scientific paper, when you make a claim, you have to back it up. There are three ways to do that: use pure reasoning, use your results, or cite literature. Pure reasoning is helpful, but it's not actually that common in scientific papers, because science is largely evidence-based. You already know about reporting your results. Citing literature is just you saying, "Hey, this guy already showed that this is true, so I don't have to totally rehash whatever it is he said."

Literature citations serve several purposes. First, they show you did your homework. Second, they show that what you have to say is backed up by the scientific community at large, placing your own work into the larger scheme of science. Third, they pointreaders to other articles that they might be interested in. Consider when you were doing the literature research putting together your experiment or even writing the paper. You might have read a paper and seen their own literature citations, and those pointed you toward something else that is useful. You need to pass on the favor.

Every journal has a style for literature citations. In the text, there are two common ways to call out a citation. The most common is to use the author and date. You can make a claim and then finish it with "(Smith 2012)," "(Chu and Mustafa 2009)," or "(Talbot et al[2]. 1993)." If it makes a sentence work better, you can write something like, "Smith (2012) reported that…" That is an especially useful way to keep the active voice when you write. The other format that is occasionally called for is the use of numbers. Numbers between parentheses or in superscript[3] are used after phrases to indicate which reference in an enumerated list you should use. Keeping an active voice when using these or even talking about them can be difficult, but you can use something like, "A previous study[134] reported that…" Note: if you have more than 100 references, you need to cut back.

The format you use for the literature cited section varies widely across journals, so you will need to look up what your science fair guide tells you to do. Each entry will include the authors' last names and initials, the year of publication, the title, and the work it was published in. Note that journals, books, theses, conference proceedings, and so on will require different formats within the literature cited section. It can take some work getting them right, so be prepared for that extra time.

Most scientific works include maps, charts, photographs, drawings, and other images. Collectively, these are called **figures**, and they are really, really important. Even a serious reader

---

[1] John's note: when I was a younger student, I hated citing literature, and that was probably because I didn't understand what I was doing. Now, I'm kind of addicted to them.

[2] Et al. is an abbreviation for the Latin phrase "et alii" which means "and others." You use it when there are more than two authors. A lot of papers have three or more authors. Some papers discussing really large projects like draft genomes or gigantic particle physics accelerators can have dozens of authors.

[3] Like the number referencing this footnote.

of your paper will gravitate toward the figures first. Figures can be referenced in any part of the body of the text. If you are formatting your own work, you will need to keep the figure near the text that first references it, but your publisher may request you include all of your figures at the end of the document.

These days, maps are relatively easy to generate. You can grab screenshots from satellite images using software like Google Earth[1]. If you are mapping a study location, you can use drawing software to generate proportional lines and features. You can even use something like PowerPoint to create maps with its drawing tools! If you use a map, you need to justify why you are including it. Are you showing where your study sites are? Ask yourself if it's important for the reader to see exactly where they are or how they are laid out. It might not be especially important for your conclusions. Sometimes, you can use maps to show results, and if you can, bully for you. Maps are a truly great way to relate results to the real world.

There are a lot of types of charts, and it can be a bewildering thing to figure out which charts you should use. Here is a list of the most common charts your will see with the reasons that you might want to use them:

- **Bar/Column[2] Chart**. These charts include masses representing data from different categories so that the reader can visually compare magnitude. When it's important to display how different treatments produced different (or similar) magnitudes in a response variable, you should use a bar or column chart. If you have statistical significance information, you can also include error bars that can indicate how different your responses are from each other.
    - Stacked Bar/Column Chart. This variation on theme includes multiple contributing factors to response variables. You might show increased vegetation for the treatment in general, but with a stacked bar/column chart, you can show which kinds of vegetation increased along with the overall trend in one figure.
- **Line Chart**. These charts typically show how a response variable changes over time or space. The x-axis needs to involve a variable that changes continuously for your sample. Too often, people use line charts when they should use bar/column charts or scatter plots.
- **Scatter Plot**. Use scatter plots when you intend to show how two variables relate to each other. Most often, scatter plots are used for correlations, in which case, you may want to include a trend line as well as an r value or R^2 value.
- **Pie Chart**. Pie charts are used to show what proportions different categories make up of a whole. They are excellent for surveys of wildlife, soils, opinions, and such.

---

[1] Make sure to cite your use of mapping software and images.

[2] Technically, bar charts have horizontal bars, and column charts have vertical bars, but people often refer to column charts as bar charts.

Many papers include photographs or drawings of specimens or apparatuses. If you want to include some, consider why you are doing so. Do you need to convey how you did an experiment, and words just won't do the job? Are you describing the feature of interest some test subject and need to display a typical example? If you answered yes, then you should include a photograph or drawing. Keep in mind that most papers do not include them, and those that do use them sparingly[1].

A drawn model is one special kind of figure that is useful in some more conceptual papers. Sometimes, scientists include cartoon[2] drawings of their systems or hypothetical connections among different components of bigger concepts. These might appear as flowcharts or Venn diagrams. If you can describe your theory using a drawn model, then you might be able to better convey your ideas. If you can't, do not fret, for such things are not found in most papers, even most great papers.

Tables are like figures, but papers tend to treat them as a separate class of things. Tables can be used to organize materials and methods or to display results. In general, tables are good at displaying relatively small quantities of numerical or textual information that are really boring to try to explain in the text. You can take any table ever created and describe its contents in words only, but that text would be really boring. You should then use tables to organize basic information for the reader's sake.

For all of your figures and tables, you will need captions. These little bits of text need to describe how to read their reference and not necessarily interpret what the item says. You will do that in the body of the work. More elaborate figures need legends for your reader to quickly translate colors and symbols.

> How do you make a poster? Most scientists use PowerPoint to make the poster, since it's easy to move graphics and boxes of text around a slide. If you use PowerPoint, make sure that you set the slide to be the appropriate height and width.

Posters and Displays

While the public perceives science fair projects with their trifold posters and experimental representations to be the realm of the elementary or high school student, professional scientists more or less do the same thing professionally. We often see bewildered looks on the faces of non-scientists when we describe the quintessential scientific activity of the poster session. We scientists often take our work and sum it up on a 4'×3' printed poster that we take to a scientific meeting and stand there while other scientists wander by with plates of hors d'oeuvres to consider and discuss our findings. It is really not unlike a science fair. Indeed, the

[1] One major reason for this is that journals like to charge for colored figures. They like to charge a lot, and black and white photographs might not be good for what you are showing.

[2] Not silly ones like Garfield, The Far Side, or xkcd… usually.

poster or display may be your most important final product for your science fair project, so you should really consider what you are doing at this point.

You might not have this problem, but many scientists struggle with their posters. Few of us have been formally trained in graphic arts, so we first approach our poster as a paper, and that is quite bad. If you walk the halls of a science department a major university, you might encounter some posters that were considered afterthoughts to the all-important papers[1] on which they were based. These posters are difficult to approach, since they are gigantic columns of text spread out on a big piece of glossy paper.

That leads to the first important lesson about the poster that you will no doubt include in your display: base it around your graphics. Your key charts, photos, and tables will draw the eye before your big and bold title will, so they have to be the centerpieces of your poster. They need to be big enough to decipher from four or five feet away with clear captions. While some photographs like pictures of you and your peers laboring away at your study site are not appropriate for papers, they often make great fillers of space and conversation starters when they are appear on posters.

Still, posters need text, and you will have to do some writing. You will need a title that catches the attention of your audience (and especially your judges.) You are more free to be a bit funny with poster titles than you are with paper titles, since you will need to be approachable as you man your display. You will also need to include the names of the authors of the poster, typically right under your title. While the text sections are less formal than one finds in papers, they are usually broken down into introduction material, materials and methods, results and discussion, and conclusions.

In the introduction, you will need to include very brief background material for your project. Limit this to six sentences maximum. You will also need to show a goal or hypothesis that you tested. Again be clear. The materials and methods should be extremely short, listing the kinds of tests you performed and the kinds of materials you used. Results and discussion sections are usually combined into one section to briefly highlight what important things were found and how they related back to the original goals of the experiment. Conclusions are usually limited to four to six sentences.

---

[1] And yes, peer-reviewed papers are more important, but good posters promote good dialog, and good dialog promotes good science.

You can use bullet points instead of flat text on your poster, since you are only highlighting important information. This method can save you a lot of writing effort, since you needn't concern yourself with transition phrases and the like. Readers tend to like bullet points, too, since they draw the eye. How many times have you seen paragraphs with text interspersed with bullet points and skipped straight to the bullet points?

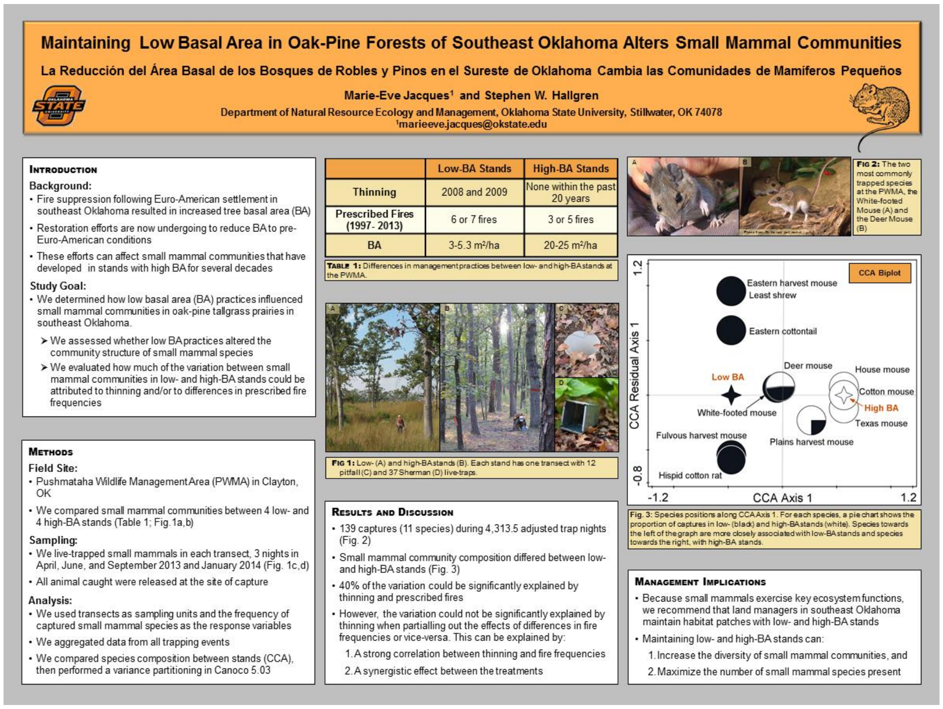

Figure 14: A poster for a professional scientific conference. Notice the layout that includes clear sections and columns. The photographs catch the eye, and the graph is attractive. The text is broken up into clear bullet points for easy reading. Courtesy of Marie-Eve Jacques and Stephen Hallgren, Oklahoma State University.

Pay attention to fonts and font size. In the world of fonts (or typefaces, as they are more accurately called) there are two classes: serif and sans serif. Serif fonts (like Times New Roman, the font in which this text is written) have little feet adorning them. In general, they are considered to be easier to read in large quantities, as you would find in a book or news article. Sans serif fonts (like Arial or Helvetica) lack these and are considered to be more attractive for things like signs and titles. You can consider using a serif font for the text blocks and bullet points and a sans serif font for the title and headers, but if you do, make sure to play with them until they seem to match. Font size should be large enough that your less youthful instructors and judges can read your sections without overstraining. I would recommend keeping most text about at an 18-point font. Make sure that your font sizes are consistent amongst sections and headers, too.

You need to be especially careful in making the widths and margins of your poster consistent. The easiest way to do this is to pick numbers and use the number fields instead of manually moving them around your poster with your mouse. Text boxes and figures may look okay on your computer screen, but you need to ensure that they are consistently proportioned and positioned, since after they are blown up, those small differences become big differences.

Presentations

Oral presentations are the third way that scientists present information. These days, we speak to groups of people using PowerPoint slideshows, possibly demonstrating aspects of our research with props. You may have stage fright, and you would not be alone. Imagine carefully assembling your results and conclusions for an audience who will stare at you as you talk about them for about twenty minutes. Now, imagine that they will ask critical questions about what you did (and didn't do.) That's giving a presentation, and we don't mean to scare you, but you have to prepare yourself emotionally as well as technically for the exercise.

Your talk will be divided in a similar way as that of a paper, but since the reader can't skip around or go back as you give your presentation in real time[1], you will have to very deliberately go through certain steps. First, present the basic system you are working on, describing what is known about it and any interesting information you know about it. You can talk about a few things that you wouldn't mention in a paper, since you have to entertain as well as inform, so now is the time to mention how many cool things your research subjects can do.

After your background information has been covered, you will need to describe what your objectives are and what you hypothesized. This step informs your audience of the scope of your study and what they should look out for as you progress through your story. You should then spend a short amount of time on your materials and methods. Students[2] are frequently tempted to spend far too much time on materials and methods, since they feel more intimately familiar with them. Just talk about how you approached the problem briefly. Only linger on parts that you think your audience might find novel or difficult.

You should focus the bulk of your presentation on your results, because this is the most important part of a talk. You will need to communicate exactly what you found and what your statistical analyses uncovered. Be descriptive and interpret your numbers into real-world terms that your audience can appreciate. Next, summarize what you found and speculate about the meaning of your work. That speculation can bring out later dialog, so don't be shy about it. You can then move on to suggest what needs to be done next to address what you speculate. Finally, thank your collaborators and helpers and open yourself up to questions.

---

[1] This isn't true of recorded presentations, of course, but even when you give one, you will likely have a live audience. Webinars are relatively recent sorts of presentations. The word "webinar" is a portmanteau of World Wide Web and seminar. Webinars are essentially virtual meetings, and they can often be recorded for future audience members.

[2] This includes graduate students, so don't take it too personally when we say "students."

Typical scientific talks involve accompanying slideshows (almost always made in PowerPoint.) Since we are visual creatures, it is important that your slideshow itself be very good—appealing, not distracting, informative, and supportive of your words. Here is some miscellaneous advice for making good slideshows:

> John's anecdote: I once made a slideshow for a minor presentation, and the file had to be converted from an OpenOffice format on a PC to a PowerPoint format on a Mac. All of my text disappeared, leaving only my graphics. The presentation went swimmingly, partly because I focused all of my attention on describing the figures.

- Minimize text. People don't read and listen well at the same time. Any text you include should simply outline what you are going to talk about. There are exceptions, of course, but few slideshows are improved through the addition of more text.
- Try to include at least one graphic per slide. If people have something to look at on the screen, then they will do so. It is especially important to include lots of charts and such as you present your results. People like to see what they are hearing.
- Use big fonts. Projectors vary in quality, and people often sit far away from the screen.
- Transitions and animations are nice, but don't overdo it.

Giving presentations well takes experience, practice, and raw talent. You are young, so you must be forgiven for inexperience, and you cannot help it if you lack raw talent (but bully for you, if you have it.) You can, however, practice. There are several reasons to do so. First, you can find the spots in your presentation in which you are unsure of what to say and fix them. Everyone has these points in their talks in which they are not sure how to transition from one topic to another, and you can practice different approaches. Second, things that often sound good in our heads sound really stupid when we talk. Third, you have to get the timing right. If you are given 20 minutes for a talk, you need to make sure you time it right. Also, you can get a lot of help through practicing with an audience of people who want to watch you succeed like your parents, teachers, and friends[1]. They can give you good feedback to improve your talk.

When you present, you need to do several things to look and sound good. Make sure to wear appropriate clothing and be well-groomed. When you speak, try to suppress the urge to fidget or pace. Tics can be hard to suppress, and everyone has their own, but you can fight them and do a better job. Speak confidently. You may not feel confident, but you have to pretend that you are. Smile. If you force yourself to smile and project positivity, eventually you will internalize that. You may need to modulate your speed, and that is what practice is for.

[1] You can use pets as audiences, too. They are pretty non-judgmental, though they can be rather insulting when they just curl up and take a nap during when you reach a really cool figure.

Glossary of Terms

Abstract – In a paper, a short summary of the full paper.

Alternative hypothesis – In statistics, the hypothesis that one mean is larger than, smaller than, or unequal to another mean with statistical significance.

Analysis of variance (ANOVA) – In statistics, a way to test multiple means against each other for statistical significance.

Bar / column chart – A chart with classes of data and their magnitudes.

Binomial data – In statistics, data that is an either/or proposition, for example, 0 or 1, or yes or no.

Chi-square (X^2) test – In statistics, a test to determine whether quantities of different classes match expected quantities.

Confidence interval – In statistics, the upper and lower bounds of the estimation of the population mean, given the sample data.

Controlled experiment – An experiment in which one sample is set up as a control, and other samples are set up to vary one or more independent variables.

Degrees of freedom – In statistics, the number of values in the final calculation of statistic that are free to vary.

Dependent variable – The variable that is measured as the result of the experimental effect.

Discussion – In a paper, the section that is used to describe the impact of results and speculate about their meaning.

Error bar – On a graph, the upper and lower bounds of a confidence interval.

Experiment – A systematic approach designed to test a hypothesis.

Experimental bias – An undesired effect leftover from accidents in the experiment.

Experimental sample / treatment sample – In an experiment, the sample(s) that has (have) variables different from the control sample.

Figure – In a paper, poster, or presentation, a chart, photograph, drawing, map, etc.

Histogram – A special kind of bar / column chart that compares the frequencies of ranges of a variable.

Hypothesis – A testable idea that could explain an observed phenomenon.

Hypothesis test – In statistics, a test to determine whether two estimated means are different from each other with statistical significance.

Independent variable – A variable that is altered by the experimenter or otherwise varied across treatments.

Introduction – In a paper, the section that describes the system and introduces the problem addressed.

Literature citation – In a paper, a callout to previously produced work used to support a point.

Margin of error – In statistics, the value added to and subtracted from the sample proportion to generate the upper and lower bounds of the confidence interval.

Materials and methods – In a paper, the section that describes sources for materials and the parameters of the experiments.

Mean (arithmetic) – In statistics, the sum of all variable values for individuals in a sample divided by the number of individuals in the sample.

Median – In statistics, the value of a variable for one individual separating the upper half of the distribution from the lower half of the distribution.

Mode – In statistics, the most common value for a variable in a sample.

Natural experiment – An experiment comparing existing natural samples using natural variation to determine independent and dependent variables.

Noise – Random variation in a value.

Null hypothesis – In statistics, the hypothesis that two sample means are not different from each other with statistical significance.

Outlier – A value for an individual's variable that is unusually much larger or smaller than the mean or otherwise appears to be non-normal.

***p*-value** – In statistics, the probability of getting a test statistic result at least as extreme as what was observed, assuming that the null hypothesis is true.

Pearson coefficient – A measure of the linear correlation between two variables for a sample.

Pilot experiment – A miniature experiment used to test the viability of and potential pitfalls for a real experiment.

Plagiarism – Unattributed use of another's text or other materials.

Probability of success – In a binomial variable for a population, the likelihood that an individual will have a particular value.

Results – In a paper, the section that describes what the experimenter(s) found.

Sample – A group of individuals that were all exposed to the same treatment.

Sample size – The number of individuals in a sample.

Scatter plot – A chart in which individuals are organized for two variables, one of which varies on the x-axis, and the other of which varies on the y-axis.

Sample proportion – In statistics, the proportion of individuals that have a certain value for a binomial variable.

Sample Variance – In statistics, a measure of the variation in a variable across the sample. In square units of the original variable.

Signal – The non-random variation in values that conveys useful information.

Standard deviation – In statistics, a measure of the variation in a variable across the population. In the same units of the original variable.

Standard error - In statistics, a measure of the variation in a variable across the sample. In the same units of the original variable.

Statistical significance – In statistics, having a low probability of randomly reproducing at least as extreme a result as found within the data.

System – The organism, mineral, interaction, etc., that you are working on in your research.

Treatment – A set of conditions under which an experiment is being performed.

Variation – The difference or divergence within or among samples.