

Experimental Design for Advanced Science Projects

Print

Return

Sandra Slutz, PhD
Staff Scientist, Science Buddies

Kenneth L. Hess
Founder and President, Science Buddies

Preface

Advanced science projects and independent scientific research are invariably subject to much scrutiny. Other scientists in the field will be looking at your work and will expect that the data has been rigorously analyzed. But rigorous analysis also requires careful experimental design. If you don't spend time thinking about the types of observations you'll be making, the likely outcomes, and how you can evaluate your data to differentiate statistically between outcomes, you might fail to carry out experiments that can actually provide you with informative data—and that would be poor science *and* a waste of your time! The sections below discuss techniques, tips, and resources for creating informative experimental designs. Although this guide will mention various principles of experimental design and statistical tests, it is not meant to be an exhaustive textbook. Instead, you're encouraged to use this guide as a means of familiarizing yourself with the general principles of experimental design. Once you're familiar with the concepts, we encourage you to continue your exploration of the topics most relevant to your science project using the references listed in the Bibliography, as well as personal resources, such as your mentor and other science and math professionals, including your teachers. We also encourage you to read our accompanying articles about the [Increasing the Ability of an Experiment to Measure an Effect](#) and [Data Analysis for Advanced Science Projects](#). When used collectively the information in these three articles will put you on the path towards a well-thought-out, top-quality research project.

It is also worth noting that the same principles, described below, that allow you to design and evaluate your own experiments, can also be applied to evaluating whether *another* person's data and subsequent interpretations (published or not) are trustworthy.

Introduction

As discussed in the [Roadmap: How to Get Started on an Advanced Project](#), data analysis should be a consideration even in the planning stages of your science project. Why? Because rigorous analysis relies on *statistics*, which are mathematical calculations that describe the data and measure how confident you can be that the observations support a particular hypothesis and are not due to random variation. But if you make too few observations, collect the wrong types of data, or fail to use proper controls, the statistical analysis will always come back as "inconclusive," which puts you in the position of only being able to make vague statements like: "The data trend suggests an increase, but I'd have to conduct more experiments to be sure." Wishy-washy results will not impress anyone who reviews your work. Nor will it be considered a reliable foundation for others to build their scientific research on.

Terik Daly, an accomplished experimenter and a Science Buddies volunteer, summarized the importance of experimental design and data analysis by stating:

"Data analysis for an advanced science project involves more than bar graphs and scatter plots, it should involve statistically minded exploratory data analysis and inference. In order to perform meaningful statistical analyses, you need to design your experiment with statistical principles in mind. This includes:

- Accurately and clearly defining your variables and sample space,
- Accurately defining your factors and levels of your factors,
- Identifying the type of experiment you are running, making sure that appropriate controls are used,
- Making sure that you perform enough replications to create a representative body of data,
- Making sure you understand the likely distribution of your data, and
- Ensuring that you are aware of and familiar with the types of exploratory and inferential analyses used in your field of science.

You must design your experiments with data analysis in mind, because if you don't think about analyzing your data until *after* your experiments, you are going to run into *big* problems."

Understanding How Other Scientists in Your Field Design Their Experiments

The best way to get a handle on what your experiments are likely to entail is to look at papers in your area of science and see *what* other investigators are measuring and *how*. You should take note of things like:

- What variables they are investigating
- Which are the independent variables and which are the dependent variables. Note: for a review of independent versus dependent variables, see the Science Buddies guide on [Variables in Your Science Fair Project](#)
- What their sample size is, meaning how many observations they make
- What controls they use
- How many experimental replicates they have

Once you have an idea what the standards are in your field, you can begin designing your own experiments, taking into account variables, controls, and how to maximize your ability to see the effect of a particular variable.

Understanding the Different Types of Variables

There are two types of variables: quantitative and qualitative. Depending on your experiment, either of these types of variables may be an independent or dependent variable. It is important to recognize which type(s) of variables you are evaluating, as some calculations and statistical tests can only be performed on data containing one or the other type of variable.

Quantitative variables are ones that differ in *magnitude*. They can easily be measured and recorded as a number. Examples of quantitative variables include age, height, time, and weight. Quantitative variables are easy to summarize using numerical calculations like median and average.

Qualitative variables, sometimes referred to as *categorical variables*, are ones where the observations differ in *kind*. Qualitative variables can be placed in categories like gender (male vs. female) or marital status (unmarried, married, divorced, widowed). This makes them particularly good for summarizing as percentages in a pie or bar chart.

Sometimes, qualitative data can be ranked. For example, a fruit survey might rank the taste of the fruit as:

Ranked qualitative variables are often called *ordinal variables*. Although the observations are qualitative, the ranking allows some numerical calculations, like averages, to be made. This can be particularly important at times when you want to compare how different people categorize data before and after an event. For instance, evaluating people's mean change in opinion for "How do you think this fruit tastes?" *before* they actually taste the fruit and *after* they get to try a sample. Ordinal variables are particularly common in social and behavioral studies.

In some circumstances, it is possible to choose whether you want to collect quantitative or qualitative data. For example, you can either ask people their exact age (quantitative) or have them select whether they are a child, teenager, adult, or senior citizen (qualitative). By pre-planning your data-analysis methods, you can choose the type of data and thus, the experimental design that is most appropriate for your research goals. When planning your experiments, try consulting Table 1, below, which gives an outline of several different types of variables, examples of data that fits them, and some of the common statistical summaries used with each type of variable.

Quantitative Variables			
Type of Variable	Definition	Examples of Data	Common Statistical Tests and Summaries
Discrete	The data are described numerically on a finite scale. There is a logical limit to the precision.	<ul style="list-style-type: none"> • Number of children in a family • Bacterial colonies on a plate • Coin toss • Shoe sizes 	<ul style="list-style-type: none"> • Mean • Median • Mode • Chi-squared • Standard deviation • Standard error of the mean • Regression • Correlation
Continuous	The data are described numerically on a continuous scale that can be broken up into infinite measurements. Theoretically, there is no limit to the precision.	<ul style="list-style-type: none"> • Temperature • Age • Weight • Time • Length 	<ul style="list-style-type: none"> • Mean • Median • Standard deviation • Standard error of the mean • Regression • Correlation
Qualitative Variables			
Type of Variable	Definition	Examples of Data	Common Statistical Tests and Summaries
Nominal (also called <i>categorical</i>)	The data are described by words or categories. They are not numerical and cannot be automatically ranked from high to low.	<ul style="list-style-type: none"> • Colors • Gender • Occupation • Location 	<ul style="list-style-type: none"> • Mode • Chi-square • Anova • Paired t-test
Ordinal (also called <i>ranked</i>)	The data are described by words or categories. Although they are not numerical in the sense that the values can be added or subtracted, the categories can be ranked from high to low.	<ul style="list-style-type: none"> • Amount of pain on a scale of 1 (low) to 10 (high) • Moh's scale of hardness for minerals • IQ • Degree of like or dislike 	<ul style="list-style-type: none"> • Median • Mode • Kruskal-Wallis • Ordinal logistic regression

Table 1. This table includes examples of when and how to use the four most common types of variables.

How the Number of Interacting Factors in Your Experimental System Impacts Your Experimental Design

The goal of an experiment is to measure the effect of a particular variable, or set of variables, on a system. The most important first step is to sit down and think about all the possible variables, also called *factors*, which might contribute to your results. For example, if you want to test to find out which brand of tires, A or B, results in the best gas mileage when driving on the highway, you first need to identify all the possible factors that might influence the gas mileage. These factors might include: the car, the weather conditions, the type of road surface, and the speed of the car. Once you've identified all the variables, you can design a fair test where you hold all variables constant, while only changing one factor. In this case you'd change which tires, brand A or B, was on the car for each trial, but keep all the other variables the same; use the same car, on the same day, in the same type of weather, on the same road, at the same speed. This would allow you to assess just the effect of the tire brand. For more information, visit the Science Buddies page [Doing a Fair Test: Variables for Beginners](#).

However, there are times in more-complex experimental systems where you need to assess either the impact of several interacting variables on the final outcome, or it is prohibitively expensive or physically impossible to change only one factor at a time. Let's go back to the question of "Which tire brand, A or B, results in the best gas mileage when driving on the highway?" as an example. The fair test outlined above would tell you which brand of tire resulted in the best mileage for a particular car, the one you used to do the test, under very specific conditions. But what if you wanted a more-general answer to the question of which tire brand results in the best mileage—one that was applicable to different cars and conditions? You might suspect that the "best" tire depends on the type of vehicle in question (a minivan, a sedan, or a pickup truck), or how worn down the tires are (new or after 5,000 miles of wear and tear). Now you have three factors (tire brand, car type, and wear and tear on the tires), which may interact, resulting in different outcomes. For example, tire

brand A may lead to the highest gas mileage on a pickup truck, regardless of wear and tear, but may be outperformed by brand B on a sedan after the tires have 5,000 miles of wear-and-tear. The 12 combinations in this example (2 brands x 3 car types x 2 wear and tear options) might all seem testable, but if you added even just one more factor, like air temperature (below 50°F, 50–75°F, or above 75°F) the total number of combinations grows exponentially to 36 (2 x 3 x 2 x 3), which might be too many to test individually! If, like in the tire example, your experimental system relies on the interaction of multiple factors, you'll need to design your experiment(s) so that you can use statistical tests to systematically evaluate the combined and individual effects of each factor. There are a variety of experimental design methods that lend themselves to this type of systematic evaluation, including: *orthogonal arrays* and *multivariate or factorial analysis*. In essence, if the experiment is properly designed, these techniques enable you to test multiple factors at one time. Consult the references in the Bibliography, below, for more details about how to set up and analyze these types of experiments.

# of factors to test	# of choices per factor				
	2	3	4	5	
1	2	3	4	5	Total # of "fair tests" needed
2	4	9	16	25	
3	8	27	64	125	
4	16	81	256	625	
5	32	243	1024	3125	

Table 2. As the number of factors and the number of choices per factor increase, the number of "fair tests" needed becomes prohibitively large. When you have many factors, and/or choices per factor, it is necessary to use a different experimental design, like orthogonal arrays or factorial analysis.

Creating Well-Controlled Experiments

Regardless of whether you are conducting experiments to evaluate one or multiple factors, you will need to design a well-controlled experiment. Controls allow you to:

- Evaluate, on a technical level, whether an experiment is working.
- Help you interpret the results by giving you *standards* to compare against.
- Guard against unforeseen factors that might bias your results.

For better or worse, the word *control* is used by researchers in several different, but related, ways. Table 3 summarizes the different usages, followed by more-complete descriptions below.

Use of the Word <i>Control</i>	Brief Description	For More Information
Positive Control	One or more experimental samples, which are known from previous data to give a positive result in the experiment. A positive control is used to confirm that the experiment is capable of giving a positive result.	Read below
Negative Control	One or more experimental samples, which are known from previous data to give a negative result in the experiment. A negative control is used to confirm that the experiment is capable of giving a negative result.	Read below
Control Group	An experimental trial where the independent variable is set at a pre-selected level, often the variable's natural state, for the purpose of comparing to all other experimental trials.	Read below
Controlled Variable	Quantities that a scientist wants to remain the same between trials so that the effects of <i>only</i> the independent variable are being measured. Sometimes called <i>constant variables</i> .	Visit the Science Buddies page Variables in Your Science Fair Project

Table 3. In science, the word *control* is used in many ways. This table summarizes the most common usages.

Positive controls are used to determine if your experimental design and your testing method are capable of detecting the effects you are trying to evaluate. Positive controls consist of one or more experimental samples, which should behave in a known manner in your experiment. If you conduct your experiment and see that your positive control behaves in an unexpected manner, you have cause to doubt the validity of your experiment. For example, if your research question was "Will this new circuit design work to turn on a lightbulb?", you would want to confirm that the lightbulb was able to be

turned on. What if it were burned out? Then the new circuit would always give you a false result (no lighted bulb) even if it was capable of turning the bulb on. To avoid a *false negative* result like this, you need to build in a positive control; in the aforementioned case, you'd want to have a circuit that you know works, like a trusted lamp, that can be used to check the function of the lightbulb.

Just as positive controls are used to minimize the impact of false negatives in an experiment, *negative controls* are used to minimize the impact of *false positives*. Negative controls confirm that the experimental procedure is not observing an unrelated effect. In the case of the new circuit design example above, a negative control would be to make sure that if the new circuit can turn on a lightbulb, that disrupting the circuit then turns off the lightbulb. This rules out the possibility that there is another power source, perhaps another circuit that you forgot was still connected, powering the lightbulb.

In cases where the experimental question is more complex than a simple "yes" or "no," it is often also useful to have standards to compare test samples against. *Standards* are products or practices in a particular field that are collectively thought of as working well. Going back to the new circuit example, if the question wasn't just whether the new circuit resulted in powering the lightbulb, but also how efficiently it did so, then it would be useful to know the amount of power used by the new circuit. Not only would you want an absolute measurement of power consumption, in terms of watts used per hour, but also a comparison to another circuit design generally accepted in the field as working well and efficiently. This other circuit would be the standard against which you expressed the efficiency of your new circuit design.

There are some research questions where neither positive and negative controls, nor standards, may be applicable. This is often the case with sociology experiments, or other research where you are surveying behavior or preferences. In these cases, it is often important to employ a *control group* to compare to your *test* or *experimental group*. This comparison helps insure that the changes you see when you change your independent variable are, in fact, caused by the independent variable. For example, if you hypothesized that "aspirin works to alleviate a headache," you'd want to have two groups of people in your study. The experimental group would take aspirin when they had a headache and a short while later answer an experimenter's questions about whether the medicine helped their headache. The control group would take a sugar pill instead of an aspirin and answer the same questions from the experimenter. Then if both the experimental group and the control group reported that taking their respective pills alleviated their headaches, you'd realize that the simple act of taking a pill had an effect, even if it was merely a psychological one, and that your data had a bias in it that you'd have to account for when analyzing it.

For control groups to be effective means of guarding against unforeseen bias, the control and experimental groups must be as identical as possible. One way to do this is to randomly assign test subjects to each group. For a more involved discussion of how this works, visit the Science Buddies page [Increasing the Ability of an Experiment to Measure an Effect](#).

Bias in an experiment can also be controlled by conducting blind experiments. In *blind experiments*, critical information is kept secret from the participants in order to guard against conscious and/or subconscious bias. There are two types of blind experiments: single-blind and double-blind. In *single-blind experiments*, critical information is kept from the test subjects, but the experimenter knows everything. In our aspirin example above, a single-blind experiment could be conducted by not revealing to the test subjects whether they are taking sugar pills or aspirin. This would prevent a scenario where people who took aspirin reported feeling better merely because they expected aspirin, but not sugar pills, to help them. The psychological expectation bias is erased if they don't know which type of pill they took. In *double-blind experiments*, both the test subjects and the experimenters are unaware of critical information. The information is only revealed to the experimenter once the data has been collected. For example, if neither the test subjects, nor the experimenters interviewing the test subjects, knew who had been given aspirin and who had been given sugar pills, it would be a double-blind experiment. The advantage is that the experimenters, because they were ignorant about the independent variable, would be unable to bias the test subject's answers by doing subconscious things like questioning the subjects who were given aspirin more vigorously than the subjects who were given sugar pills. Because they guard against a wider range of potential biases, double-blind experiments are considered the most scientifically rigorous; however, they can be challenging to implement as they require additional help by a third party who holds the key to which group—control or experimental—each test subject belongs.

Bibliography

These resources provide additional information about how to design experiments:

- Anderson, M.J. and Anderson, H.P. (1993, July/August). Applying DOE to Microwave Popcorn. *PI Quality*. Retrieved August 25, 2009, from <http://www.statease.com/pubs/popcorn.pdf>
- Gauch, H.G., Jr. (2006). Winning the Accuracy Game. *American Scientist*, 94, 2, 133.
- Khare, R. (n.d.). *Three Romeos and a Juliet: Our Early Brush with Design of Experiments*. Retrieved August 25, 2009, from <http://www.symphonytech.com/articles/romeos.htm>
- Trochim, W.M.K. (2006). Introduction to Design. *Research Methods Knowledge Base*. Retrieved August 25, 2009, from <http://www.socialresearchmethods.net/kb/desintro.php>
- Wilson, E. B. *An Introduction to Scientific Research* New York: Dover Publications, Inc.

Additional information about statistical tests and summaries is available in these online statistics textbooks:

- McDonald, J.H. (2008). *Handbook of Biological Statistics*. Retrieved August 25, 2009, from <http://udel.edu/~mcdonald/statintro.html>

- NIST/SEMATECH. (2006). *NIST/SEMATECH e-Handbook of Statistical Methods*. Retrieved August 25, 2009, from <http://www.itl.nist.gov/div898/handbook/>



It's *free!* As a member you will be the first to receive our new and innovative project ideas, news about upcoming science competitions, science fair tips, and information on other science related initiatives.

[Science Fair Project Home](#) [Our Sponsors](#) [Partners](#) [About Us](#) [Volunteer](#) [Donate](#) [Contact Us](#) [Research Grants & Outreach](#) [Site Map](#)

[Science Fair Project Ideas](#) [Science Fair Project Guide](#) [Ask an Expert](#) [Blog](#) [Teacher Resources](#) [Parent Resources](#) [Student Resources](#) [Science Careers](#) [Join Science Buddies](#)

[Privacy Policy](#)

Copyright © 2002-2010 Science Buddies. All rights reserved.
Reproduction of material from this website without written permission is strictly prohibited.
Use of this site constitutes acceptance of our [Terms and Conditions of Fair Use](#).