# Overview: Steps of the Scientific Method and Experimental Design, With a Cheesy Example at the End

Many people struggle to understand the presentation and use of the steps of the scientific method as outlined here, because it breaks down the process into a series of discrete, specific components that may or may not be the same as what they have learned previously or in other classes. Be sure to pay close attention to these distinctions and spend time explicitly practicing the steps as outlined here—one of the most valuable skills you can learn in college is the use of good logic, and this explicit framework will help strengthen your reasoning skills.

1. <u>Puzzling observation</u>: Puzzling observations typically come from direct experiences. They can also come from other people who are familiar with the subject at hand. Basically, a puzzling observation is something that you want to figure out.
2. <u>Causal question</u>:
    a. A **how** or **why** question
    b. Asks about the **cause** of a phenomenon
    c. Fairly **general** in nature
        i. Can be used to generate **multiple** possible explanations
    d. *Do not confuse with*: **descriptive questions**
        i. = Compare/contrast, who/what/where/when, yes/no, same/different
        ii. Often, descriptive questions can be re-phrased as causal questions—just think a bit harder about what the *overall* question should be
        iii. *May already contain a hypothesis* and may therefore be *biased*
    e. **Cannot** be answered by observation alone—some sort of **manipulative experiment** (definition below) is required!
3. <u>Hypothesis/Proposed Explanation</u>
    a. A **single, specific** idea that answers the causal question
    b. There may be multiple hypotheses that explain a single causal question. Each hypothesis can be tested separately using the scientific method.
    c. Must be **testable**. There are usually many ways to test a single hypothesis
    d. *Do not confuse with***:** predictions (= specific to the given experiment; see below), *statistical hypotheses* (= actually predictions and usually stated as *null* and *alternative* hypotheses), or *if…then…* statements (= actually hypotheses and predictions smashed together with each other)
4. <u>Experiment</u>
    a. **Manipulative/causal** (again, more about this below)
    b. Must test the hypothesis/proposed explanation
    c. There may be more than one experiment that will test a single hypothesis—just choose one at a time to keep things simple!
    d. Must hold constant other factors that could affect the results
5. <u>Predicted Result</u>
    a. How you *expect* the experiment to turn out, based on your hypothesis
        i. **Careful! Statisticians re-phrase predictions as a** "*null hypothesis"* and "*alternative hypothesis."* I generally try to avoid this language, but you will encounter it in the wild. Know it for what it is.
    b. Generated from:
        i. The specific experiment that you have decided to test
        ii. Your hypothesis/proposed explanation
6. <u>Actual Result</u>
    a. How the test actually turned out—what happened?

   b. No interpretation is involved—just how the test turned out
7. <u>Conclusion</u>
   a. Your interpretation of the actual result in comparison with the hypothesis and predicted result
   b. There are only three possibilities:
      i. The experiment and results supported the hypothesis
      ii. The experiment and results **did not** support the hypothesis
      iii. The experiment and results are inconclusive (they neither supported nor failed to support the hypothesis)
   c. **NOTHING IS EVER PROVEN BEYOND A DOUBT USING THIS LOGIC SYSTEM—<u>DO NOT EVER</u> ATTEMPT TO PROVE YOUR HYPOTHESIS CORRECT, OR YOU WILL BRING PAIN AND SUFFERING UPON YOURSELF!** Also: you will hear lots of other people mis-using this term. Don't buy into it.

**Experimental Design: Manipulative (Causal) versus Mensurative (Correlational) Experiments**

To understand the difference between these experiment types, you must first understand what *variables* are and the difference between an independent and a dependent variable:

**Variable**: Any factor that has a number of different possible values

*Continuous variables*: variables that are numerical and for which there are an infinite number of possibilities (ex: stem height, leaf weight, number of eggs, duration of time)

*Categorical (Discrete) variables:* variables that are defined by categories (ex: red, orange, yellow, green, blue, purple). Note that continuous variables can be divided up into discrete variables

*Dichotomous variables*: a specific type of categorical variable: variables for which there are only two different possibilities and the two possibilities are mutually exclusive (you can only be one or the other, not both; ex: handedness, day/night, either/or, heads/tails)

**Independent variable**: A variable that is *directly manipulated (changed/varied)* by the experimenter

Independent variables will generally have *at least* two <u>levels/conditions/treatments</u>:

1. One level/condition that represents what would happen if you didn't change anything (often called the *control treatment*…but see below)
2. One or more levels/conditions for which you have altered the independent variable (the *experimental treatment*)

BUT independent variables may also have more than two levels/conditions/treatments.

**Dependent variable**: A variable that is *measured* by the experimenter

You are looking to see if the dependent variable changes as a result of another variable (the independent variable)

**Constant/Controlled <u>variable</u>:** Ideally, all other variables in an experiment aside from the independent and dependent variables should be held constant. This ensures that the changes you measure in the dependent variable are due to the independent variable. In actuality, it is usually impossible to control *all* other variables, but focus on important ones.

**Mensurative (correlational) experiment:** An experiment in which two dependent variables are measured to determine if there is a relationship (a correlation) between them

**Manipulative (causal) experiment:** An experiment in which an independent variable is directly manipulated (changed/varied) and the effects of the manipulation on a dependent variable are then measured. This will allow you to determine if the independent variable causes the dependent variable to change.

**…So, what's a control?…**

When conducting a manipulative experiment, you want to be really sure that your independent variable is what is causing the dependent variable to change, and not some other outside factor. Therefore, scientists say that you control for the effects of all outside factors by holding everything (that you can control) except the independent variable constant between the different levels/conditions of the independent variable. Note: this is the definition of a controlled **variable**. Sometimes people will refer to a "control treatment" when they are talking about a specific category of an independent variable that reflects the status quo somehow (what happens if you don't give that drug to that mouse, etc.).

**So what's the big deal about the difference between correlations and causation?**

As you should be able to see from the above descriptions of mensurative (correlational) and manipulative (causal) experiments, there is one critical difference between the two—correlational experiments have *no* independent variables! If you don't directly *manipulate/change* one variable (an independent variable) and *then measure* its effects on another variable (the dependent variable), you cannot demonstrate that one variable causes another variable to change!

Another way of thinking about an experiment: A series of questions:

1. **Independent variable**: What is the overall variable that I am changing/manipulating?
   a. **Levels/treatments/conditions**: How am I changing my independent variable? What different categories of that independent variable am I creating?
2. **Dependent variable**: What am I measuring?
3. **Constant/controlled variables**: What variables/factors need to stay the same across the different treatments so I can be sure that the independent variable is what's causing changes in the dependent variable and not some other factor?
4. **Replication**: How can I be sure that my results were not due to chance? (the answer to this is usually: repeat the experiment!)
   a. If an experiment is **not** repeated, the consequence is that your evidence is **circumstantial**.

**<u>And finally, a single example.</u>** This is a really simple example to think about (I hope). If you understand it, start practicing the steps by coming up with your own examples.

**Observation:** Yum. This cheese is especially delicious.
**Causal question:** *Why* is it, exactly, that this cheese is so delicious?
**Hypothes<u>e</u>s (plural of hypothesis):**

1. Maybe it's the fresh milk, from pasture-fed New York cows.
2. Maybe it's the special enzymes used in the cheesemaking process.
3. Maybe it's the fungi in the cheese's rind.
4. Maybe… (can you come up with more of your own?).

**Experiments to test Hypothesis 1:**

*A.* *Correlational experiment*: Measure taste responses to a whole bunch of different cheeses. Figure out which ones are the most delicious.

*B.* *Manipulative experiment*: Make several batches of cheese, using milk from different sources (independent variable: milk type). The cheese preparation processes should otherwise be as close to identical as possible. One batch of cheese could be made from a mixture of equal parts of all the available milks (controlled treatment). Develop a method for quantifying deliciousness, and use it to measure participants' reactions to the different cheeses.

**Predictions:**

*A.* *Correlational experiment:* Participants will hone in on and agree that the initial cheese identified in the initial observation is the most delicious compared to all the other tested cheeses.

*B.* *Manipulative experiment:* Cheese prepared with fresh milk from pasture-fed New York cows will be the most delicious. Self-test: *Can you break this prediction down into a null and alternative "statistical hypothesis"?*

**Results and Conclusions:**

*A.* For the sake of argument, let's say that the pattern we observe matches with our predictions. While this could support several of the identified hypotheses, it won't necessarily tell us which specific hypotheses are supported unless we can sort through everything that we know about every single type of cheese tested to come up with some common denominators. So the conclusions that we can draw from this correlational experiment are very limited.

*B.* Again, for the sake of argument, let's say our predicted results match our actual results. Again, our hypothesis supported? Again, I will ask, does this mean that it's true? No, it does not. Don't look so shocked! There could still be other reasons why this specific cheese is delicious. Maybe I was extremely hungry when I ate it. From here, we could go back to the drawing board to try another hypothesis. Scientists (and cheese lovers) call the iterative nature of scientific inquiry, "job security."