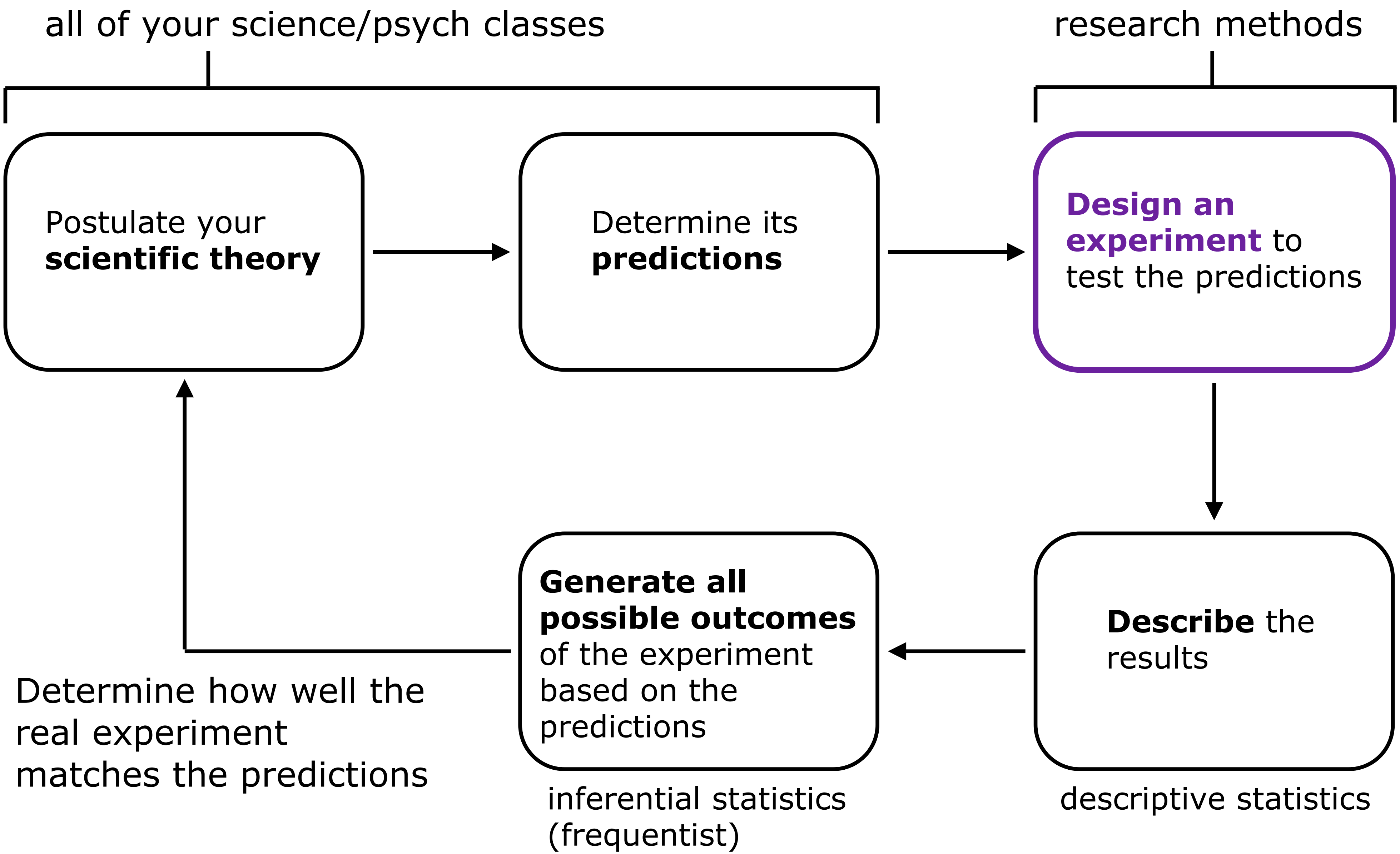
PSYCH-UH 1004Q: Statistics for Psychology

Class 2: Science is our goal

Prof. Jon Sprouse
Psychology

We need to ground ourselves in science

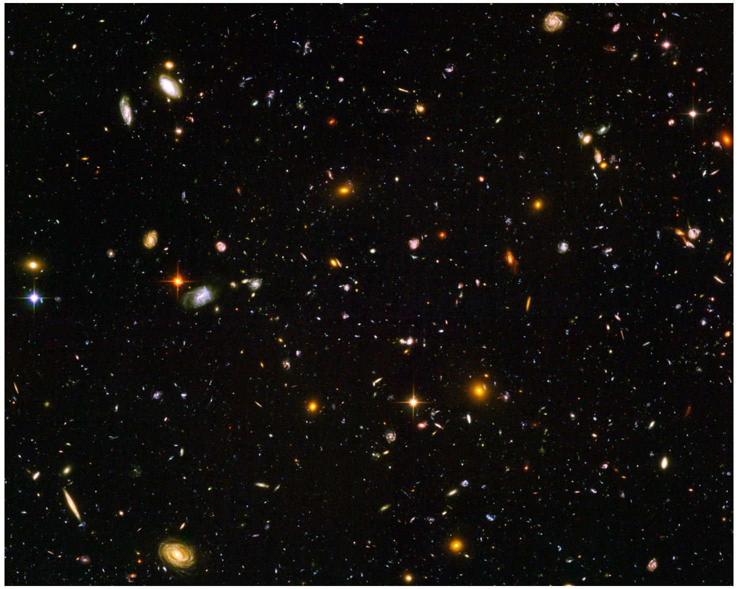


Why do we do science?

(quick review from lecture 1)

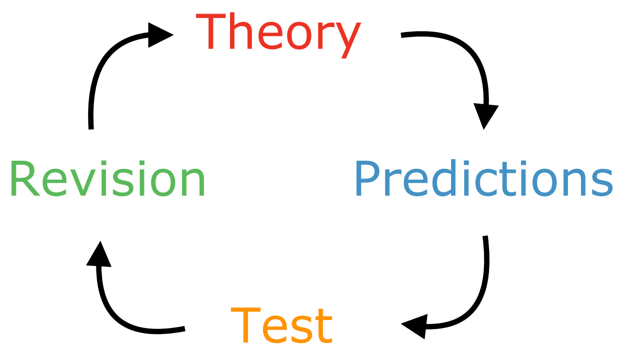
It is all about changing your mind!

We all have beliefs about how the universe works.



Science gives you a set of rules for figuring it out, and most importantly, for **changing your mind** when you encounter new evidence.

You've probably seen diagrams like this before - something showing the "scientific method". Today we are going to try to look at this in a bit of detail.



Science has never been more important

Scientific debates are becoming more and more relevant to the world we live in. As information becomes easier and easier to access, it is critical that we understand how to use evidence to prove/disprove theories.

The Impact of Vaccines in the United States

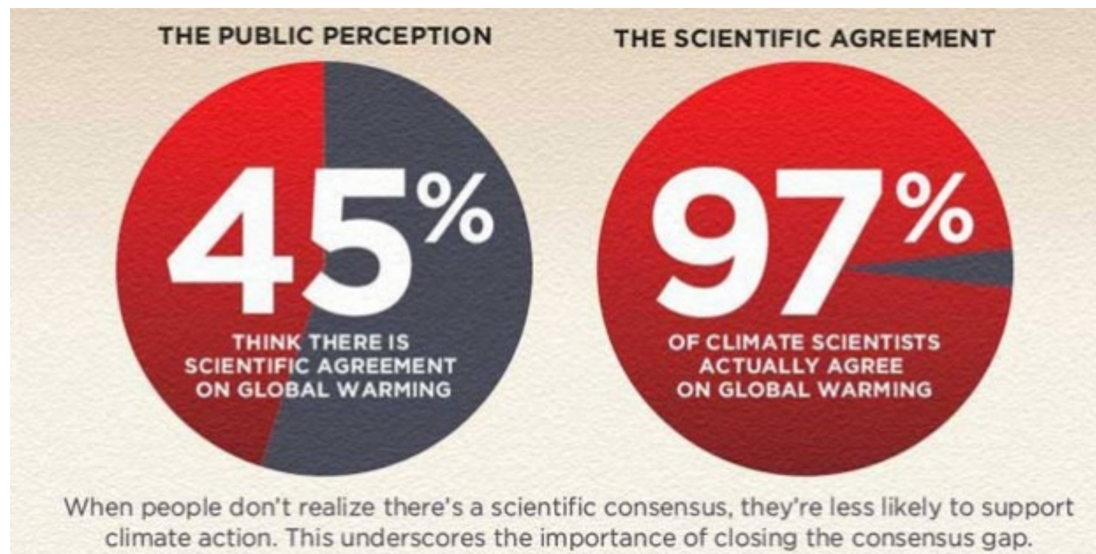
Disease	Baseline 20th Century Pre-Vaccine Annual Cases	2008 Cases*	Percent Decrease
Measles	503,282	55	99.9%
Diphtheria	175,885	0	100%
Mumps	152,209	454	95.7%
Pertussis	147,271	10,735	92.7%
Smallpox	48,164	0	100%
Rubella	47,745	11	99.9%
Haemophilus influenzae type b, invasive	20,000	30	99.9%
Polio	16,316	0	100%
Tetanus	1,314	19	98.6%

*Provisional. Widespread use of vaccines in the United States has eliminated or almost eliminated infectious diseases that were once terrifying household names. Credit: Morbidity and Mortality Weekly Report, Centers for Disease Control and Prevention, 4/2/99, 12/25/09, 3/12/10

In short, there are a number of debates in society that depend upon an understanding of what it means to use evidence.

Science has never been more important

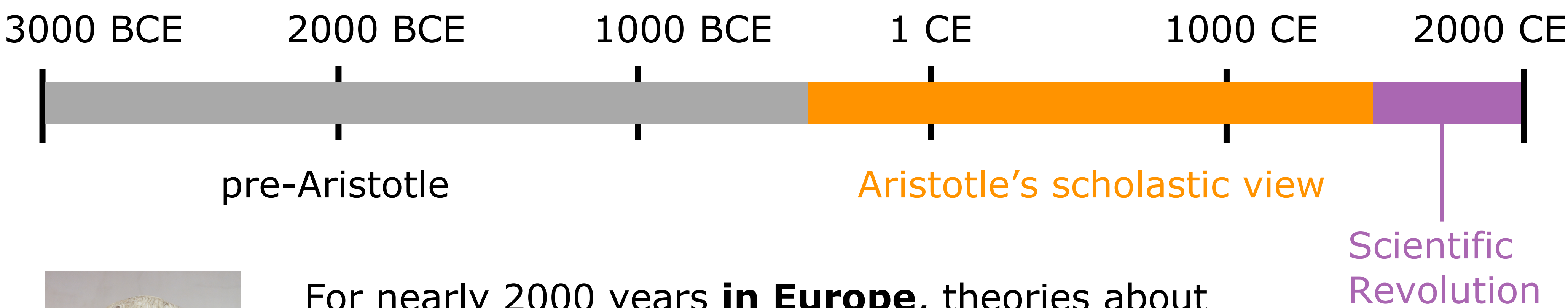
Scientific debates are becoming more and more relevant to the world we live in. As information becomes easier and easier to access, it is critical that we understand how to use evidence to prove/disprove theories.

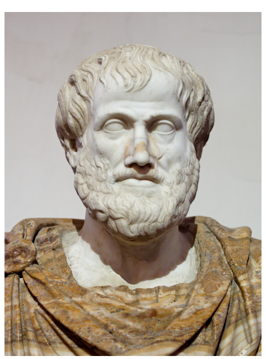


In short, there are a number of debates in society that depend upon an understanding of what it means to use evidence.

Why do we care?

The bottom line is that science is incredibly successful, and appears to be substantially more successful than other knowledge-gathering methods.





For nearly 2000 years **in Europe**, theories about the universe were dominated by Aristotle's view that there were 4 elements (earth, air, water, fire), and the earth was the center of the universe.

Then from 1550-1700 prominent thinkers started to question those theories, and even question how to build a theory. We call this the age of enlightenment or the scientific revolution. From that point forward, the expansion of human knowledge has been dramatic!

What does it mean to be a scientific theory?

Scientific theories are testable

There is a lot of debate in the philosophy of science about how to define the difference between scientific theories and non-scientific theories.

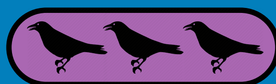
But one frequent answer is that scientific theories make testable predictions (and non-scientific theories do not).

testable	**untestable**
All ravens are black	?
The COVID-19 vaccine reduces serious illness and death	What makes a claim untestable?
[Can you think of other examples?]	

Testability comes from causal relationships

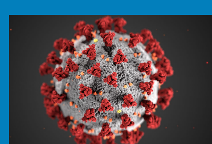
The goal of a scientific theory is to **explain something**. And the way they do this is by proposing a **causal relationship** between properties.

All ravens are black.



If this were truly a scientific theory (and not just a simple example), it would probably say that the genes that determine the species also determine the black color. There is a causal relationship between species and color.

COVID-19 vaccines reduce illness and death.



Here we know that the claim is that the vaccine induces an immune response that allows the immune system to attack the sars-cov-2 virus before it has a chance to replicate too much. There is a causal relationship between the vaccine and the immune response to COVID-19.

property 1 —causes→ property 2

Testability comes from being constrained

“A theory that explains everything, explains nothing”
-maybe John Playfair ~1802

What this famous quote is telling us is that **theories must be constrained**. If your theory predicts that ravens are black, that is great. But if you then encounter red ravens, and you say “my theory can explain that too”, then your theory is no longer constrained.

If a theory can accommodate any possible fact that you uncover, then it is not teaching you anything about the world.

Non-scientific theories are untestable because they are unconstrained. People add options to them to try to explain away evidence that they do not like. (Ghosts are real, but only believers can see them, etc.)

But science is committed to making every question testable. It may not work. There may be questions that can’t be answered by science. But we try.

Formalizing the process

(This is just a first pass - we will keep expanding on this throughout the semester)

Let’s figure out the logic we want to use

Let’s say you want to test for a causal relationship between COVID-19 vaccination and the severity of illness:

COVID-19 vaccination —causes→ (lower) severity of illness

How would you test this?

Maybe you’d try something like this:

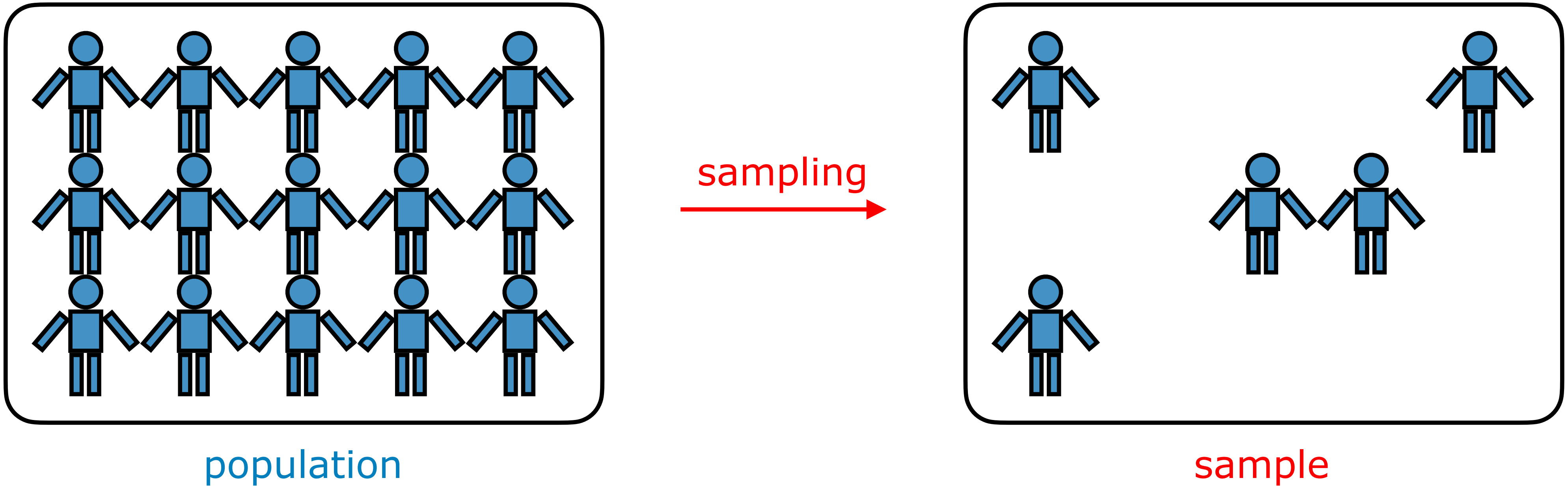
1. Find two groups of people.
2. Give the vaccine to one group and a placebo to the other.
3. Measure the severity of illness in the two groups.

This is a basic experiment. It sounds logical, right? What we need to do is formalize this a bit with some precise terms and concepts so we can see why this sounds logical to us, and what the options are when designing a study.

Populations and Samples

1. Find two groups of people.
2. Give the vaccine to one group and a placebo to the other.
3. Measure the severity of illness in the two groups.

In an ideal world, we would study the entire population that we are interested in. But we usually don't have the resources (time, money) for that. Instead, we select a subset of the population called a sample:



Populations and Samples

Population: The complete set of items/values. This is most commonly thought of as people (e.g., all of the people in the US is the population of the US), but it can also be other units that you want to study. A population can be defined using whatever criteria you want (e.g., the population of people born in NJ; or the population of people who have been vaccinated).

Sample: A subset of a population. The process of selecting the subset from a population is called sampling. Sampling is usually necessary because most populations of interest are too large to measure in their entirety. Samples can be chosen randomly, or they can be chosen non-randomly. How a sample is chosen matters for the types of inferences you can make. (Random is best… everything else limits your inferences.)

Random Sample: In a random sample, every member of the population has an equal likelihood of being selected. This is the ideal for experiments for a number of reasons we will learn about later - for now I just want us to know this is the ideal.

Why do we care?

We care about the distinction between populations and samples because:

We will only ever get to work with samples. We will rarely ever get to test an entire population.

We typically want to draw conclusions about populations. So we will need to learn how to **draw inferences** about the population from the sample.

The logic of experimental manipulation

1. Find two groups of people.
2. Give the vaccine to one group and a placebo to the other.
3. Measure the severity of illness in the two groups.

What we want to do is show that vaccination causes a change in illness.

COVID-19 vaccination —causes→ (lower) severity of illness

To do this, we need to manipulate vaccination. The simplest way to do this is to have two **conditions** in the experiment: one with vaccination and without vaccination, and compare them:

condition 1	vaccination	—causes→	lower severity
condition 2	no vaccination	—causes→	higher severity

Variables - traditional names

In general, for any experiment, what we want to do is show that the independent variable causes a change in a dependent variable:

independent variable —causes→ dependent variable

variable: Property of an object or event that can take multiple values.

independent variable: An independent variable is one that can be manipulated. Its values are independent of any other variable.

dependent variable: A dependent variable is one whose value is determined by another variable.

Any property you can imagine can be a variable. Whether someone is vaccinated or not is variable (with two values). How many days someone was sick is a variable (with lots of possible values). How tall someone is is a variable (with lots of possible values).

Variables - modern names

It is becoming more common for people to use modern names for variables. Here is a pair that some people like better:

predictor variable → (causes) → outcome variable

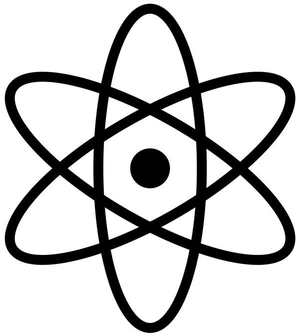
Another term for an independent/predictor variable is **factor**, and another term for the values of a factor is **level.** Our statistical software, **R**, uses these terms, so you will have to learn these terms too. I am sorry for so many different terms!

factor → (causes) → dependent/outcome variable

factor: Another term for an independent variable.

level: Another term for “value” of a variable.

Why do we care?



We care about the distinction between independent variables/predictor variables/factors and dependent variables/outcome variables because:

This is the logic of our studies: we are looking for a causal (directional) relationship between the two variables. The names encode that directionality for us.

Variables - continuous vs discrete

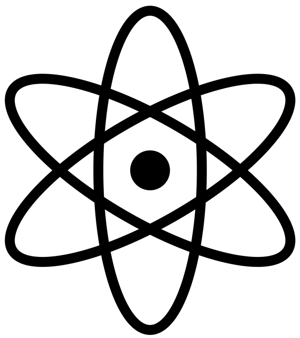
continuous: A continuous variable can take any value within some range (in principle, continuous variables can take an infinite number of values).

discrete: A discrete variable can only take a finite number of values.

	independent/predictor variable	causes →	dependent/outcome variable
discrete:	vaccinated/unvaccinated		number of days ill
continuous:	10mg dose/20mg dose		time ill (h, m, s, ms, ns, etc)

In practice, no measurement will be truly continuous (and perhaps no measurement could be according to quantum mechanics). So the distinction is mostly about what your theory assumes - does your theory treat the variable as discrete or continuous?

Why do we care?



We care about the distinction between continuous and discrete variables because:

When we look at descriptive statistics, we will see that the plots that we create will be subtly different for the two types of variables.

When we look at inferential statistics, we will see that the way we think about statistical tests may differ depending on the nature of the variables (e.g., regression vs ANOVA), but ultimately the underlying math and logic is the same.

Don’t worry, the differences in how we plot our data and how we think about statistical tests will follow relatively naturally from the difference between continuous and discrete variables. You won’t have to work very hard to memorize anything about that. In my daily life as a scientist, I rarely explicitly think about continuous vs discrete, and I never explicitly discuss it.

Why did we select two groups?

1. Find two groups of people.
2. Give the vaccine to one group and a placebo to the other.
3. Measure the severity of illness in the two groups.

vaccine **placebo**

It makes no sense to give both a vaccine and a placebo to the same group of people. They would all simply be vaccinated.

vaccine **placebo**

In this case, if we want to see the different effects, we need to test them in different people

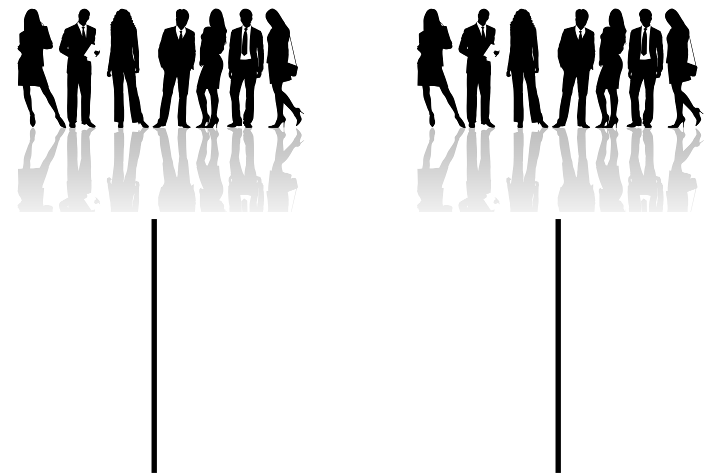
But this isn't true for all experiments. Can you think of experiments that could use a single set of participants?

In general, we have two options

Repeated Measures

condition 1 **condition 2**

Independent Measures

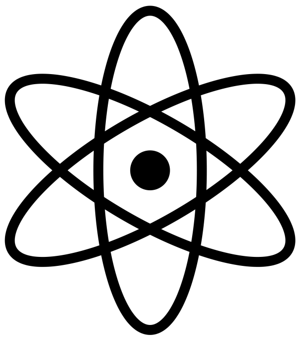


condition 1 **condition 2**

Repeated Measures: If each participants sees every condition, we call it repeated measures. It is also called a **within-subjects** design.

Independent Measures: If each participants sees only one condition, we call it independent measures. It is also called a **between-subjects** design.

Why do we care?



We care about the distinction between repeated-measures and independent-measures because:

When we learn inferential statistical tests, we will see that there are differences between the calculations for repeated-measures and independent-measures designs.

(We will discuss why in detail when we get to that part of the course!)

There are four possible types of measurements

1. Find two groups of people.
2. Give the vaccine to one group and a placebo to the other.
3. Measure the severity of illness in the two groups.

nominal: The values on the scale are named; but there is no order, no distances, and no meaningful zero point.

ordinal: The values on the scale are named and ordered, but there are no distances, and no meaningful zero point.

interval: The values on the scale are named, and ordered, and the distances between points are equal; but there is no meaningful zero point.

ratio: The values on the scale are named, ordered, regular distances, and there is a meaningful zero point

Examples using temperature

nominal: The values on the scale are named; but there is no order, no distances, and no meaningful zero point.

One way we could measure temperature is by categorizing it as pleasant for humans or unpleasant for humans. Notice that this has names, but there is no order (in fact, unpleasant temperatures can be less than or more than the pleasant ones).

pleasant	unpleasant
20 C	35 C
25 C	-10 C
18 C	50 C

There can be any number of categories. This example just has two. Another example could be the countries that people are from. There are 195 of them! But there is no inherent order for country names, and no distance between country names.

Examples using temperature

ordinal: The values on the scale are named and ordered, but there are no distances, and no meaningful zero point.

We could order temperatures by how hot (or cold) they are. Imagine we have 5 temperatures that we have measured. We can place them in order:

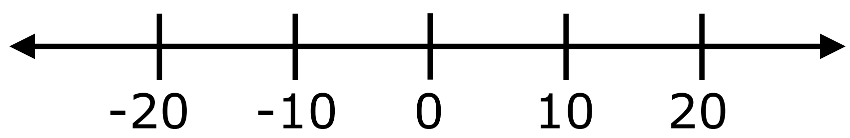
order	Temp C
1st	71 C
2nd	69 C
3rd	50 C
4th	40 C
5th	10 C

The critical thing to note here is that the order does not capture anything about the distances between the measurements. The difference between 1st and 2nd is only 2C; the differences between 2nd and 3rd is 19C; etc.

Examples using temperature

interval: The values on the scale are named, and ordered, and the distances between points are equal; but there is no meaningful zero point.

We could measure temperature using a scale like Celsius.



In an interval scale, the distances between numbers is regular. The distance between 0C and 10C is the same as the distance between 10C and 20C.

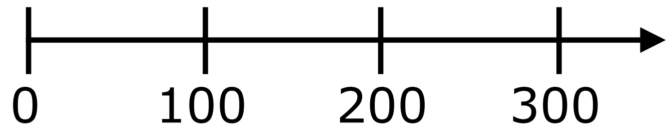
But, critically, the 0 in C does not mean zero temperature. It is just the freezing point of water. It is not a true zero point.

This means 20C is not twice 10C.

Examples using temperature

ratio: The values on the scale are named, ordered, regular distances, and there is a meaningful zero point

We could measure temperature in Kelvin.



0 in K is absolute zero. It does mean the absence of temperature. It really means zero.

So 200K is twice as hot as 100K. And 100K is twice as hot as 50K.

Examples using temperature

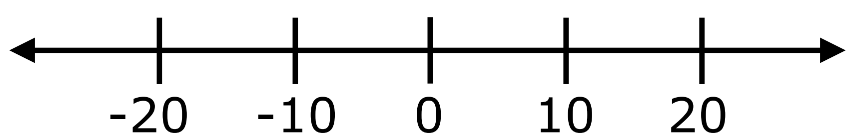
nominal: We could categorize temperature as pleasant or unpleasant

pleasant   unpleasant

ordinal: We could order temperatures by how hot (or cold) they are.

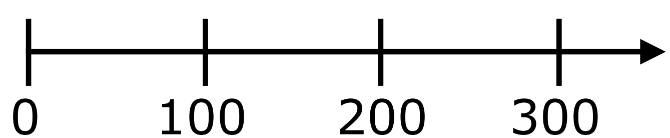
1st, 2nd, 3rd, 4th, 5th, ...

interval: We could measure temperature in Celsius.



0 in C does not mean no temperature. It is just the freezing point of water. So 20C is not twice 10C.

ratio: We could measure temperature in Kelvin.



0 in K is absolute zero. So 200K is twice as much as 100K.

Examples for our covid experiment

nominal: We could categorize participants as well or unwell.

well unwell

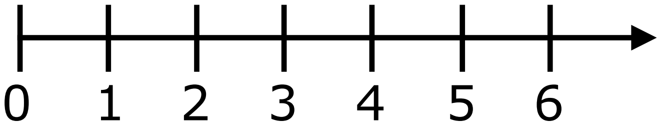
ordinal: We could order participants by how sick they are.

1st, 2nd, 3rd, 4th, 5th, ...

interval: We could rate participants' health on a scale from 1 to 5.

very ill 1 2 3 4 5 very healthy

ratio: We could measure how many days participants showed symptoms.

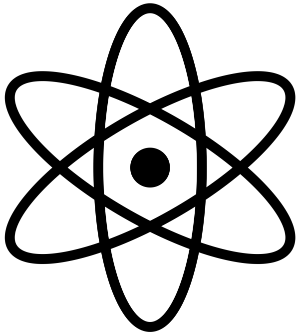


Relationships among the scales

There are four properties. The types of scales are defined by which properties they have. And the properties can only be added in a specific order because they entail each other: a zero point entails distances, distances entail orders, and orders entail the existence of names.

	names	orders	distances	zero point
nominal	✓	X	X	X
ordinal	✓	✓	X	X
interval	✓	✓	✓	X
ratio	✓	✓	✓	✓

Why do we care?



We care about measurement scales because:

First and foremost, we want to choose the scale that best fits our theory.

Second, we will see later that there are different statistical tests available for different scales.

Sometimes people worry that some tests (called parametric) can only be used for ratio scales. That is technically true, but the practical effect is really small, so most people don’t worry about it any longer.

Experimental versus Observational

Up until now we have just been assuming that you want to do **experimental research**. But we haven't defined it.

Experimental Research: The researcher directly manipulates the independent/ predictor variable(s).

The goal of experimental research is to establish a causal relationship between the independent/predictor variable and the dependent/outcome variable,.

Experimental Research:

manipulates ↓ independent variable —causes→ dependent variable

The only way to be certain that there is a causal relationship is to run an experiment (i.e., directly manipulate the independent variable).

Experimental versus Observational

But what if we cannot manipulate the independent variable? This is common - we can’t manipulate some properties of humans.

Observational Research: The researcher selects groups with different levels of the independent/predictor variable(s), but does not manipulate them directly.

Observational research can establish a correlation between the independent/predictor and dependent/outcome variables

Observational Research:

selected ↓

Variable X ⟷ correlates with ⟷ Variable Y

Correlation does not imply causation! Correlation is a logically weaker conclusion. The way to convince yourself of this is to ask yourself what possible states of the world could lead to correlation.

What do we mean by “logically weaker”?

Observational research is logically weaker in that there are multiple ways that a correlation could arise. Only one of those ways is the causal relationship that we want to demonstrate.

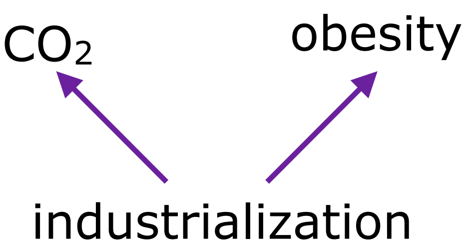
Examples

1. Variable X causes variable Y.

2. Variable Y causes variable X.

High national debt tends to correlate with slower economic growth. Economists argue about which causes which!

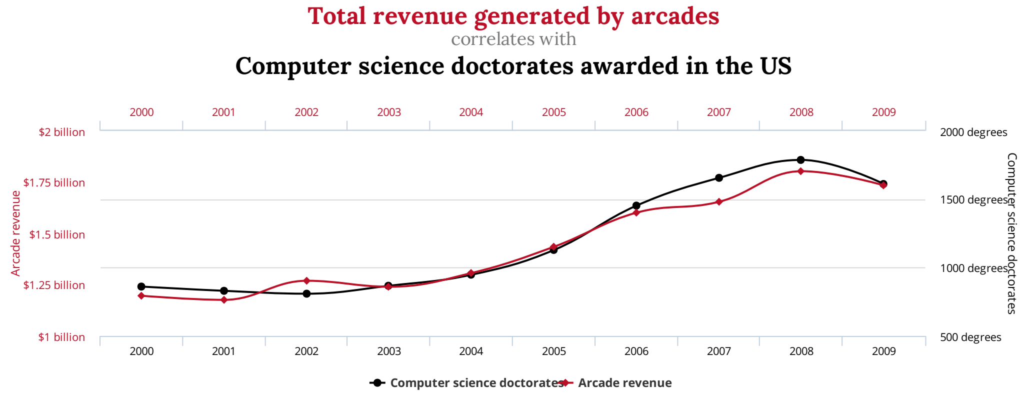
3. A third variable, Z, causes both variable X and Y.



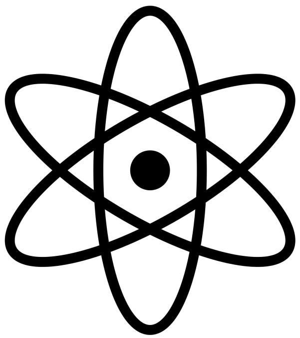
CO_2 and obesity correlate, but neither causes the other!

4. Variable X and variable Y correlate accidentally.

https://www.tylervigen.com/spurious-correlations



Why do we care?



We care about experimental vs observational because:

First and foremost, we want to establish causal theories in science. So, experiments are our strongest tool.

But, sometimes, we can’t manipulate the independent variable. This is common when it comes to properties of people or societies (social science, economics, etc).

Ultimately, both experimental and observational studies use the **same statistical methods**. The only difference is in the strength of the conclusions that we can draw.

Questions we can ask about any study!

What is the **population** of interest?

How many **samples**? (**repeated** measures vs **independent** measures)

What is the **independent variable/predictor variable/factor**?

What are the **levels** of the independent/predictor/factor?

What are the **conditions** of the experiment?

What is the **dependent variable/outcome variable**?

What type of measurement is it (**nominal, ordinal, interval, ratio**)?

Is the research **experimental** or **observational**?

A science workout

Science workout

Science is learned through apprenticing with other scientists.

The first skill to learn is to train your brain to generate ideas. You want to reward it for thinking.

The trick is to learn to generate ideas without immediately shutting them down as not worthwhile. We will learn to refine ideas later. For now, we just want to generate them.

Workout:

Let's come up with theories… or even just claims people might make. Then let's figure out how we would test them. Let's identify all of the parts of the study: the variables, the conditions, the population, repeated/independent measures, the measurement scale, and whether it is experimental or observational.