



## From Correlation to Causation

---

---

---

---

---

---

---

---

## Quest for finding causes

When something happens, we ask "Why?" We want to know what caused the event

Why are we interested in causes?

Knowing the causes frequently provides understanding

Knowing causes empowers us to intervene

These two tend to go together

Why do these barrels produce better beer?

Learning the reason is more hops provides understanding

And a procedure for making better beer

---

---

---

---

---

---

---

---



## What is a cause?



The roots of talk of causation is found in our doing something to produce an effect

We want to move a rock, so we push it

We want to stay warm so we put on a jacket

Independent of our own action, a cause is something which **brings about** or **increases the likelihood** of an effect

The cause of the explosion was the spark from the generator

---

---

---

---

---

---

---

---

## Correlations Can Point to Causation

Correlation is not the same as causation, but correlations do call for explanations

The challenge is to determine what explains the correlation

does one of the variable cause the other?

is there some other cause of both?

Examples:

Consumption of red wine and reduced heart attacks

Good study habits and good grades

---

---

---

---

---

---

---

---

## Major Difference: Correlation Symmetrical; Causation Asymmetrical

Being run into in a traffic accident might be a cause for the big dent in your car



Having a big dent in your car is correlated with having a car accident, but it is not the cause of having a car accident

Causation is **directional**, correlation is **symmetrical**

So when correlation points to causation, we still need to establish the direction

---

---

---

---

---

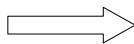
---

---

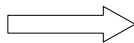
---

## Challenge of Establishing Directionality

Does watching violence on TV result in aggressive behavior in children?



Or do the factors that generate aggressive behavior cause children to watch more violence on TV



---

---

---

---

---

---

---

---

## Causal Loops

Sometimes X causes Y and then Y causes more X  
The causation here is still directional, but works in both directions

Back pain may be the cause of a person limping  
but walking with a limp may cause further back pain



---

---

---

---

---

---

---

---

## Independent/Dependent Variables

- In correlations, the two variables are equal, but in causal claims they are not
  - **Independent variable**
    - The variable that is thought to be the cause
    - The variable that is altered/manipulated in an experiment
    - The treatment in a clinical trial
  - **Dependent variable**
    - The variable that is thought to be the effect
    - The variable that one is trying to predict/explain
    - The outcome in a clinical trial
- The dependent variable *depends on* the independent variable

---

---

---

---

---

---

---

---

## Clicker Question

If average driving speed is the independent variable  
Its value depends upon the dependent variable  
It is the variable that is manipulated in the experiment  
It is the variable that is affected by the manipulation  
It is to be explained by finding the cause

---

---

---

---

---

---

---

---

## Measured versus Manipulated

The strongest tests of causation claims involve manipulation of variables → Experiments

In some contexts, a researcher does not or cannot manipulate the independent variable

Immoral to assign people to categories such as having unprotected sex

Cannot assign people to categories such as being female

All one can do is measure the independent variable refer to it as a *measured independent variable*

When it is possible to manipulate the independent variable (conduct an experiment)

refer to it as a *manipulated independent variable*

---

---

---

---

---

---

---

---

## Clicker Question

Which of the following makes no sense?

Manipulated independent variable

Measured independent variable

Manipulated dependent variable

Measured dependent variable

---

---

---

---

---

---

---

---

## Remember Operational Definitions (Measures)

Often causal relations are specified in general terms:

Violence on TV causes violent behavior in school

The variables used to operationally define such variables are sometimes referred to as *measures*. The specific values on these variables are *data*

“The number of gun firings on a given TV show is a good *measure* of violence on the show. We have related *data* on gun firings to *data* on two *measures* of aggressive behavior by those watching the show.”

The measure: Violence operationally defined as # of gun firings

Data on # of gun firings

---

---

---

---

---

---

---

---

## Extraneous Variables

Given the number of possible variables to consider, in any given *inquiry* some variables will be correlated with the dependent variable of interest

If these are not the variables we are focusing on, we term them *extraneous*

But

What we term *extraneous* may in fact be the causally relevant variable

So, in testing a causal hypothesis, care must be taken to rule out any causal link between these extraneous variables and the dependent variable

---

---

---

---

---

---

---

---

## Clicker Question

Does the following argument represent the logic of experimental confirmation?

If X is a cause of Y, then there will be a statistically significant difference in Y when X is present

There is a statistically significant difference in Y when X is present

∴ X is the cause of Y

No, the first premise is usually false

No, one cannot determine statistical significance in an experiment

No, the argument affirms the consequent

No, the argument form is modus ponens whereas modus tollens should be used

---

---

---

---

---

---

---

---

## The Logic of Causal Research

To confirm or falsify a causal claim based on a correlation, we use *modus tollens*. The first premise in each case, though, is different

Confirming a causal claim:

If X is not a cause of Y [and there is no alternative plausible hypothesis], then there will not be a statistically significant difference in Y when X is present

There is a statistically significant difference in Y when X is present [and there is no alternative plausible hypothesis]

∴ X is a cause of Y

Whether the first premise is true depends critically on how we set up the test of the causal hypothesis—whether we make it very unlikely that anything else could produce a difference in Y

---

---

---

---

---

---

---

---

## The Logic of Causal Research - 2

### Falsifying a causal claim

If X were the cause of Y [and the auxiliary assumptions are true and the experimental set up is adequate], then there would be a statistically significant difference in Y when X is present

There is no statistically significant difference in Y when X is present [and auxiliary assumptions are true and the experimental set up is adequate]

∴ X is not the cause of Y

The truth of the first premise depends critically on how we set up the test of the causal claim

## Causal distinctions

1. Necessary, Sufficient, and Contributory Causes
2. Proximate and ultimate causes

## Sufficient Causes

A sufficient cause specifies how to bring about an effect

it provides us a recipe

Example: Exercise is a sufficient cause of losing weight

Although not everyone likes the recipe

I have never taken any exercise, except for sleeping and resting, and I never intend to take any. Exercise is loathsome.

Mark Twain, Essays: Seventieth Birthday

However, this only works if one doesn't increase what one eats  
Exercise suffices for losing weight if a person doesn't change anything else



## Necessary causes

Necessary causes are required to bring about an effect

Blocking them will prevent the effect

Example: engaging in intercourse can result in pregnancy

So not engaging in intercourse can prevent the pregnancy from occurring

However, sometimes there are other ways of bringing about the effect

Sex is not strictly necessary for producing babies—in vitro fertilization can replace it



## Clicker Question

Putting a jar over a candle is

A necessary cause for the candle burning

A necessary cause for putting out the candle

A sufficient cause for putting out the candle

Both a necessary and sufficient cause of putting out the candle



20

## Recap

	Quick description	Value of identifying	Inadequacy as a full characterization of 'cause'	Practical difficulties
<b>Sufficient Causes</b>	A sufficient cause can bring about an effect, (e.g. ordinarily, pulling a trigger of a loaded gun suffices to fire a gun).	If we know sufficient causes, we can be empowered to bring about desired effects, (e.g. if we know how to fire the gun, we can.)	Events which we may want to call causes aren't strictly speaking sufficient, since we can imagine contexts in which the causes fail to suffice, (e.g. pulling a trigger fires a gun, but not if the bullets are duds).	Sufficient conditions are difficult to identify, and, perhaps, impossible to perfectly specify, since such a specification would require an exhaustive description of the relevant background conditions.
<b>Necessary Causes</b>	A necessary cause is a precondition that must obtain for the effect to manifest, (e.g. guns need a hammer to fire.)	If we know necessary causes, we can prevent effects, (e.g. remove the hammer and the gun will not fire).	Preconditions don't do anything (just because a gun has all its parts, doesn't mean it will fire.)	Necessary conditions are also difficult to identify since the scientific enterprise can reveal, unexpectedly, that what was once supposed necessary is not

## Partial or contributory causes

A factor that increases the likelihood of the event occurring but may not be **either necessary or sufficient** for the effect

Icy roads are a cause of car accidents

but are neither necessary nor sufficient causes of them

Genetic factors are a cause of heart attacks

but are neither necessary nor sufficient causes of them

22

---

---

---

---

---

---

---

---

## Uncle John Exception

My uncle John smoked two packs of cigarettes a day for 75 years, and he never got lung cancer. See, smoking doesn't cause lung cancer.

This would be an effective counterexample if the claim were that smoking is a sufficient cause of lung cancer

But if the claim is that smoking is a contributory cause, one or even many counterexamples are not telling

Rather, what one must do is show that over a population there is no increase in lung cancer among those who smoke

23

---

---

---

---

---

---

---

---

## Why care about contributory causes?

Even though changes in contributory causes cannot either:

Totally prevent the effect

Bring about the effect by themselves

They can significantly increase or decrease the likelihood of the effect

Eliminating fast foods from your diet can reduce the risk of heart attacks



24

---

---

---

---

---

---

---

---



## Proximate/ultimate causes



What was the cause of the Cedar Fire?

A hunter's flare?



25

---

---

---

---

---

---

---

---

## Proximate/ultimate causes

Failure to remove the dry thinning?



Dry conditions

26

---

---

---

---

---

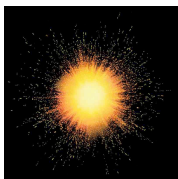
---

---

---

## Proximate/ultimate causes

The growth of seedlings into trees?



The big bang?

27

---

---

---

---

---

---

---

---

## Proximate/ultimate causes

There is generally a history of events, each of which plays a role in bringing about the event of interest

Sometimes we are interested in events in close proximity to the effect we wish to explain—proximate cause

Sometimes we are interested in events further back in the chain of causation—ultimate causes

Proximate and ultimate causes are not competitors

They are relevant for different explanatory projects

28

---

---

---

---

---

---

---

---

## Clicker Question

Compared to planting seeds, watering a plant

Is a more proximate cause of plant growth

Is a more ultimate cause of plant growth

Is a more important cause of plant growth

Is a sufficient cause of plant growth

29

---

---

---

---

---

---

---

---