Causal Fallacies

by Sophia

WHAT'S COVERED

This tutorial will explore a few reasons why variables that are correlated may lack a causative connection Our discussion breaks down as follows:

1. Lurking Variables
2. Reversed Association
3. Coincidence

BEFORE YOU START

In the previous lesson, you learned that even a strong correlation does not guarantee a causative link. As you review this lesson, consider the following scenario, and consider whether lurking variables, reverse association, or coincidence may be involved.

Researchers notice a strong correlation between the amount of time individuals spend on social media platforms and their reported levels of loneliness and anxiety.

Does social media cause loneliness and anxiety?

1. Lurking Variables

One reason that we cannot say there is causation with a strong correlation could be a lurking variable. This other variable could be confusing the relationship between the explanatory variable and the response variable.

EXAMPLE In many families where parents left the light on in their infant's room as they slept, the infant developed nearsightedness. This is an actual studied scenario, where researchers noticed that there was a positive relationship between sleeping with the light on and having nearsightedness. Therefore, researchers concluded that sleeping with the light on might cause nearsightedness.

Is this conclusion correct?

Upon follow-up studies, this conclusion was shown to be incorrect. The nearsightedness of the children was

genetic and was therefore caused by their parents' nearsightedness, not by sleeping in a room with the light on. In fact, the parents' nearsightedness caused them to leave the light on in the child's room so that the *parents* could see.

Therefore, the nearsightedness of the child and the light being left on were both due to the lurking variable of their parents' nearsightedness. It wasn't the light that caused the child's nearsightedness.

EXAMPLE In a study on the effect of exercise on bone density, researchers initially found no correlation. This was unexpected because there is a known mechanism whereby exercise causes increases in bone density. Upon further analysis, researchers discovered that weight was a lurking variable. Individuals that exercised more had lower weights, which results in lower bone density. Because the research team had not considered the weight variable initially, they nearly made an erroneous conclusion.

Just as in the case of nearsightedness and sleeping with the light on, there's a lurking variable behind the scenes confounding the relationship between exercise and bone density. Once researchers controlled for weight as a variable, they did indeed find that exercise had a positive effect on bone density.

TERM TO KNOW

Lurking Variable

An unidentified variable that may confound the relationship between the variables being studied, leading to incorrect conclusions if not accounted for.

2. Reversed Association

Another reason that we cannot say there is causation with a strong correlation could be that the association is reversed. If we don't know the direction of the cause-and-effect of two variables, we cannot say that it is a causal relationship, only that they are strongly correlated.

EXAMPLE As the number of firefighters at a fire increases, so does the damage the fire causes. Suppose you come up with this conclusion: "Sending firefighters is counterproductive because they only increase the size of the fire."

This is obviously a ludicrous conclusion to draw. In fact, the true association is just the other way around. The association is reversed. It is cause-and-effect relationship; however, it is a severe fire that causes the firefighters to arrive, not the other way around.

3. Coincidence

Finally, it's possible that a correlation is simply the result of chance.

EXAMPLE Paul the Octopus, hatched on January 26, 2008, became an international sensation as an animal oracle during the 2010 World Cup. Residing in a tank at the Sea Life Centre in Oberhausen, Germany, Paul demonstrated an uncanny ability to predict football match outcomes. His method was delightfully simple: presented with two food-containing boxes adorned with the flags of competing teams, Paul would choose the one he ate from first as his prediction. Remarkably, he correctly foresaw the results of Germany's seven World Cup matches, including their third-place play-off win over Uruguay and Spain's victory in the final. In total, Paul achieved an impressive record of 12 correct predictions out of 14, boasting an approximate success rate of 85.7%.

Paul the Octopus correctly picks Spain over Holland in the 2010 World Cup Final

It's not reasonable to think Paul had any real insight into these games, but it sure if fun to follow! A more serious example can be found in science.

EXAMPLE When researchers explore very large datasets containing many variables, they will inevitably find correlations between some of the variables. This is common in DNA studies where scientists look for correlations between a particular condition and many thousands of DNA markers. In this sort of study, we would expect that some proportion of the markers would correlate by chance. The initial list of DNA markers alone does not imply any causation. Scientists must follow-up on these candidates with controlled, randomized experiments to identify causative relationships.

SUMMARY

Sometimes two variables will be related because one causes the other, whereas other times they will be well-correlated, but the association isn't what we call "causal"; this is the difference between correlation and causation. In many cases, there's a lurking variable--something behind the scenes that's causing an increase or decrease in both variables, or maybe a decrease in one and an increase in the other. Finally, sometimes there appears to be a relationship between two variables, but it is only a coincidence. Thus, the most valid way to prove causation is with a controlled, randomized experiment. However, strong evidence for causation can be made with an observational study.

Good luck!

Source: THIS TUTORIAL WAS AUTHORED BY JONATHAN OSTERS FOR SOPHIA LEARNING. PLEASE SEE OUR TERMS OF USE.

TERMS TO KNOW

Lurking Variable

An unidentified variable that may confound the relationship between the variables being studied, leading to incorrect conclusions if not accounted for.