# Chapter 10: Causal Reasoning

## I. Introduction

Human beings see causation everywhere. It’s how our brains are hard-wired; it’s not just a habit of ours, it’s a survival technique. When we touch a hot stove and it hurts, we immediately form the causal connection in our brains, “hot stoves cause pain.” Thereafter, we’ll try to avoid touching them when we can. We see causation so often, however, that we *over-see* it – we make causal connections when there really isn’t any connection between the two things.

In this chapter, we’ll accomplish three things. First, we’ll go over the false cause fallacies – typical mistakes people make in causal reasoning. Next, we’ll go over different kinds of causal connections – specifically necessary conditions and sufficient conditions. Then, we’ll go over Mill’s Methods – John Stuart Mill’s methods that we can use to provide evidence for our causal assumptions.

## II. False Cause Fallacies

False cause fallacies are mistakes humans make in causal reasoning. There are many, but I want to introduce three: post hoc ergo propter hoc, non causa pro causa, and oversimplified cause. These are the fallacies we’re trying to avoid by using Mill’s Methods to provide evidence for (or against) our causal assumptions. The first two completely misunderstand a causal link; the third one is partially correct about a causal link, but in a way that is often misleading.

### 1. Post Hoc Ergo Propter Hoc

This is Latin for “after this, therefore because of this.” It is often shortened to just “post hoc.” The post hoc fallacy is where we see causation where there is none, because one thing happened right after another one. Usually, there are two significant events that caught our attention, and our brain forms a causal link between them. (If two events are insignificant, we usually don’t leap to a causal connection. This morning I brushed my teeth and then a little later, I put my shoes on. I had to stop and think what I did after I brushed my teeth; neither of these events are notable, so I don’t even think about them. I certainly don’t think that brushing my teeth caused me to put shoes on.)

Here are a few examples of the post hoc fallacy.

In 2013, Beyoncé played the half time show during the Super Bowl. It was a spectacular show, with great music, dance routines, and special effects. About twenty minutes into the second half of the game, the power went out in the Superdome. It took about 40 minutes before the power was fully restored. Fortunately, there was enough power that the commenters could talk about how the lights were still off…

Of course the immediate assumption was that Beyoncé’s show caused the power outage. (It did not. Beyoncé’s crew brought their own generators, they didn’t even use the Superdome’s power).

See the “after this, therefore because of this” structure? Beyoncé played, *and then* the lights went out. So people (falsely) assumed that her show *caused* the lights to go out. Here’s another one:

The governor of California gave a speech, and then an earthquake hit Los Angeles. He needs to stop giving speeches; he’s putting people in danger!

He spoke, *and then* the earthquake hit, so the assumption is he caused the earthquake, maybe by bringing bad luck to the residents of Los Angeles. Superstitions are often reinforced by the post hoc fallacy. If a black cat crosses my path, *and then* I get in a car wreck, I will assume that the black cat crossing my path *caused* my bad luck.

### 2. Non Causa Pro Causa

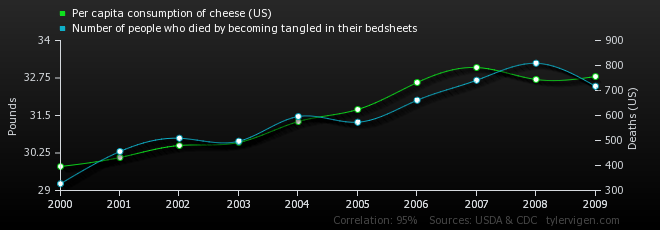
The Latin here translates to “not the cause for the effect.” Just like with post hoc, I have completely misidentified a causal link. However, this time it’s not because one event happened before another one, it’s because the two things happen at the same time, or in the same place, or I notice some other correlation between them, and use this correlation to assume there’s a causal link. As the saying goes, correlation is not causation. Here are a few examples.

Every time the governor of California gives a speech, a natural disaster happens somewhere in the world. He needs to stop giving speeches; he’s putting people in danger!

Notice how this is different from the example in the last section. The post hoc had two events, that happened one right after the other. The non causa finds a correlation – these two things keep happening at the same time.

Now, correlation *can* give you a sense that two things might be causally linked. If every venue Beyonce ever played in suffered a power outage, I’d start to suspect her show did have something to do with it. But simply noticing a correlation is not proof.

Here’s another example:



This is from Tyler Vigen’s web page; the link is here (it opens a new page): [Spurious Correlations](https://tylervigen.com/spurious-correlations). If we drew from this chart the conclusion that eating cheese causes people to die by becoming tangled in their bedsheets (or that people dying in their bedsheets is causing more people to eat cheese), we are making the non causa fallacy. And, a famous example:

Every time ice cream sales go up, so do shark attacks! Sharks must really like the taste of ice-cream stuffed humans.

This is an interesting one, because there is actually sort of an indirect a causal relationship between ice cream sales and shark attacks—they both share a cause. Buying (and presumably eating) ice cream does not cause sharks to attack; shark attacks do not cause a rise in ice cream sales. Rather, hot weather is responsible for both ice cream sales rising and for humans going swimming in shark-infested water. [[1]](#footnote-1)

### 3. Oversimplified cause

The fallacy of oversimplified cause is where you’re dealing with a complex phenomenon with many moving parts, but you pick a partial cause and pin the full weight of blame (or praise) on that one thing. It doesn’t completely misidentify a causal connection, like post hoc and non causa; the thing you’ve picked out does play a causal role, but you’ve oversimplified by ignoring all of the other parts of the equation. Here’s an example:

The economy of California is thriving. It must be because of the governor’s wise economic policies.

Or

The economy of California is crashing. It must be because of the governor’s idiotic economic policies.

The truth is that the governor’s economic policies do have an effect on the economy of the state. However, one person is never solely responsible for an entire state’s economy. Further, economic policies never have an immediate effect. If government policies do partially cause a change in the economy, we’re unlikely to see that change for several years – usually during the next governor’s term of office. This is the fallacy of oversimplified cause.

## III. Necessary and Sufficient Conditions

Now that we’ve reviewed the fallacies involved in causation we’re trying to avoid, we need to look at what we mean by “causation” in any particular case. A cause can be a necessary condition for the effect, a sufficient condition, both necessary and sufficient, or neither necessary nor sufficient.

### 1. Necessary conditions

A necessary condition is something that needs to happen in order to get the result. If you don’t have the condition, you’re not going to get the result. This does not mean that you’re guaranteed the result every time you have the condition in place, however. It just means that if you don’t have the condition, you’re not going to get the result. For example: water is a necessary condition of plant growth. The appropriate amount of water is needed to get your plants to grow. It’s not ALL you need though; the plant is also going to need sunlight, nutrients, access to air, and so forth.

Another example: getting enough sleep is necessary to have energy. If you don’t get any sleep, you’re not going to have a very good day the next day, and you’ll struggle to pay attention. The human body absolutely needs sleep. It’s not the only thing we need to give us energy, however; we also need to eat enough food, and to not have the flu, among other things.

### 2. Sufficient conditions

A sufficient condition is a condition that is enough to bring about the result. If the sufficient conditions are in place, you will get the result. It doesn’t mean the condition is needed, though. For example: beheading is a sufficient condition for death. That’s enough, that will get the job done. It’s not at all necessary, though; there are other ways to die. You know it’s not necessary because you *can* get the result without this condition being present.

Another example: throwing a brick through a glass window is sufficient to break it. That will get the job done. It’s not necessary though, because there are other ways to get the same result. You can throw a rock through it, you can throw a person through it (I’m thinking of those old Western movies with bar fights, where they usually throw someone right through a window out onto the street), an earthquake could break it, and so on.

### 3. Both necessary and sufficient conditions

Sometimes a condition is both necessary and sufficient. To identify these, you just ask “is it necessary?” and “is it sufficient?” If you say “yes” both times, it’s necessary and sufficient. It’s not a third category, really, just a conjunction of the two types of conditions.

It is a special sort of condition that is both necessary and sufficient, though. If you’ve identified one of these connections, you’ve identified the total cause. Scientific phenomenon, thoroughly explored, can turn up necessary and sufficient conditions. Really good definitions will give you the necessary and sufficient conditions of when that term can be used.

For example, being between the ages of 13 and 19 (inclusive) is necessary and sufficient for being a teenager. You have to be one of those ages to be a teenager, so it’s necessary. And, turning 13 is all you need to do. You don’t need to register someplace or sign a contract; turning 13 is sufficient for being a teenager until you’re 20. So, being between 13 and 19 is a great definition of teenager; I have provided necessary and sufficient conditions for when the word applies.

Here’s an example of a bad definition, because it fails to be necessary and sufficient. Consider this: “A human is a tool-using mammal.” Using tools is not necessary to be a human. Babies count as human, long before they are able to use tools. And, using tools is not sufficient for being a human – many primates and even some birds use tools. If tool-using mammal was sufficient to count as human, those primates would be human (but not the birds).

Let’s return to the plant example from the necessary condition section. Water is necessary, but not sufficient, for plant growth. What if I thoroughly explored what plants need to grow, though, and put them all together? What if I found all of the necessary conditions? I’m thinking about water, sunlight, nutrients, air, and anything else that plants might need. Once you lay all of this out, you’d have the necessary and sufficient conditions for plant growth. Each individual element is necessary; the group together is sufficient. This is called “individually necessary and jointly sufficient.” And then you’d understand how to make plants grow.

### 4. Contributing causes

Some causes are actual causes of a phenomenon without being necessary or sufficient. We can call these “contributing causes.” Think of the governor of California’s economic policies, from the false cause section above. The governor of California’s policies are not necessary for the state’s economy to do what it’s going to do; and they’re by no means sufficient. They do add something, though; they will eventually contribute to how well the state’s economy does.

Or, thinking back to the example of what I need in order to have energy – I absolutely need enough sleep, and to consume enough calories, and to avoid having an illness that saps my energy. What about exercise, though? Exercise is absolutely not sufficient for having energy – you still need sleep and food. Is it necessary? Do you need to exercise regularly to have energy? Nope, but it does help. Exercise is a contributing cause. So is caffeine. You don’t have to drink coffee to have energy – caffeine is not necessary. It’s also not sufficient. If you’re seriously sleep deprived, no amount of coffee can fix that.

## IV. Mill’s Methods

Mill’s methods were laid out by British philosopher John Stuart Mill in 1843, in his book *A System of Logic*. He was specifically looking into scientific inquiry of causes, but his methods can be applied outside of laboratories. You’ve probably used similar techniques to try to investigate causes yourself. He by no means invented these methods; he did, however, formalize them. Taking a more formal approach to providing evidence for (or against) a causal hypothesis can help you avoid making a leap in logic that is unwarranted, and committing a false cause fallacy.

They do have some drawbacks that you need to keep in mind:

* This is solidly inductive logic. We can never prove with absolutely certainty what, in particular, caused an effect. Each of these methods has a different level of strength, and the more research we put into it can make our arguments stronger, but we can never be guaranteed we are correct.
* Mill’s methods are not great at telling you what the cause is if you have absolutely no idea. You need to come up with some hypotheses, some theories, and then Mill’s methods can help you test those theories.
* The methods can help you establish a causal connection, but can’t tell you if you’ve found the *full* cause or not. Beware of committing the fallacy of oversimplified cause!

That said, these methods help you systematically organize your data and testing of a hypothesis, to help you avoid mistakes. They can absolutely add evidence to your theory. And, they’re really good at eliminating causal suspects – they’re better at refuting your hypothesis than they are at proving it.

There are five methods: the method of agreement, the method of difference, the joint method of agreement and difference, concomitant variations, and the method of residues.

### 1. The Method of Agreement

Suppose you went to a picnic with a bunch of friends, and everyone brought a different dish to share. You had a great time, tried some of everyone’s food, and about twelve hours later, you got hit with a nasty bout of food poisoning. Once you recover, you’re curious exactly what caused your food poisoning – maybe we don’t want to let the friend who made it cook for you anymore. How would you go about figuring out what poisoned you?

You’d call your friends and ask them if they got sick, right? And then you’d probably ask them what they ate.

The method of agreement does exactly this, only a little more systematically. When we use the method of agreement, we want to focus on cases where the effect is present, and see if they all have something in common. So in this case, we want to talk to all the friends who also got sick, and ask them what they ate, to see if they all ate at least some of the same food.

Mill described the Method of Agreement as a way to gather evidence to explain a phenomenon (an effect) “by comparing together different instances in which the phenomenon occurs.”[[2]](#footnote-2) We see something has happened multiple times, so we gather instances where we know the effect occurred, and look to see if there’s something they all had in common.

Let’s break the method down into steps.

1. Identify the *effect* you’re trying to explain.
2. List all potential *causes* for that effect.
3. Identify which case experienced which of the causes.
4. If you find one potential cause that all cases had in common, you have evidence that this is the cause.

For our poisoned picnic: (1) the effect we’re trying to explain is “food poisoning,” (2) all potential causes are all the foods present at the picnic. For step (3) we want to ask our friends what they ate. (4) If we can find one and only one food that everyone ate, we have evidence that this is the food causing people to get sick.

A good chart can help us track our data. I’m going to put the friends who got sick up at the top of the chart, and the foods present down the side. Then I’m going to check off who ate what. Here’s my chart:

|  | Andrew | Betty | Cesar | Dominique | Elaine |
| --- | --- | --- | --- | --- | --- |
|  | Sick | Sick | Sick | Sick | Sick |
| Fried chicken | **\*** | **\*** |  | **\*** | **\*** |
| Potato salad |  | **\*** | **\*** | **\*** |  |
| Coleslaw | **\*** | **\*** | **\*** | **\*** | **\*** |
| Rolls | **\*** |  |  | **\*** |  |
| Fruit salad | **\*** | **\*** | **\*** |  |  |
| Cake |  |  | **\*** | **\*** | **\*** |

Notice that I’ve specified that each of these cases has the effect present – all five of these people got sick. It should stick out to you that the only thing everyone who got sick ate is the coleslaw. Coleslaw is now our main suspect. I haven’t proved it with certainty – it’s possible I’m leaving off a potential cause, or that someone misremembered what they ate, or my whole “food poison” hypothesis is wrong and we really have the stomach flu, or something else. But it’s a pretty good start. Also, I’ve got great evidence that the rolls are safe. Three people avoided the rolls and still got sick.

Except someone misremembered what they ate. Cesar calls back and says “whoops, I also had fried chicken. I forgot.” So we amend our chart:

|  | Andrew | Betty | Cesar | Dominique | Elaine |
| --- | --- | --- | --- | --- | --- |
|  | Sick | Sick | Sick | Sick | Sick |
| Fried chicken | **\*** | **\*** | **\*** | **\*** | **\*** |
| Potato salad |  | **\*** | **\*** | **\*** |  |
| Coleslaw | **\*** | **\*** | **\*** | **\*** | **\*** |
| Rolls | **\*** |  |  | **\*** |  |
| Fruit salad | **\*** | **\*** | **\*** |  |  |
| Cake |  |  | **\*** | **\*** | **\*** |

Now everyone who got sick had two things in common. They all ate the chicken, and they all ate the coleslaw. (The rolls are still looking pretty safe, though). This leaves us with a few different possibilities: (1) the chicken was bad. (2) the coleslaw was bad. (3) BOTH of them were bad. (4) some strange reaction between these two foods is what made people sick. (5) Neither of them were part of the cause; the real cause is something we didn’t look for (like the stomach flu). Fortunately, we have other methods we can use to try to figure out what exactly was making people sick.

To sum up: the method of agreement gathers instances where the effect is present, and looks for one (and hopefully only one) thing they all have in common. It is most often used where the data is already available, so you can, for example, go through a database and retrieve only examples where the effect happened to analyze further. Or, after the picnic happened, you can call your friends, find out who else got sick, and find out if there is something they all ate in common.

It is also important that you brainstorm as many possible causes as you can; you want to be thorough with this part of the procedure to make sure you didn’t overlook the actual cause.

So, the method of agreement: (1) gathers instances with the effect present; (2) lists all elements which are potential causes; (3) looks for one element which was present every time the effect is present. If you find one, that is likely the cause.

### 2. The Method of Difference

The method of difference is very different indeed from the method of agreement. The method of difference cannot be performed by analyzing existing data; you need to set up an experiment to test your hypothesis. This is because you need to control for as many factors as you can, to make sure you’re *only* testing your hypothesis.

Mill described the method of difference as a means of providing evidence to explain a phenomenon “by comparing instances in which the phenomenon does occur, with instances in other respects similar in which it does not.”[[3]](#footnote-3) So, you want to compare a case where the effect happened, and a case that is quite similar in which the effect did not happen. Or, to put it another way, you want to try to *make* the effect happen – and then take away the cause you’re testing and see if the effect also goes away.

How the method of difference works: After identifying the effect you want to explain, and the element you want to test for a causal relationship, you want to set up two trials, one with the element present, and one with the element absent. You want to see if you get the effect when the element is present, and fail to get the effect with the element absent. When you set up your two trials, you want to keep everything else as similar in both instances as you can, to make sure you’re only testing your hypothesis, and not accidentally testing something else. To break it down into steps:

1. Identify the effect you want to explain.
2. Identify the causal hypothesis you want to test.
3. Set up an experiment, with two trials. Control as many variables as you can. Vary only the element you want to test as the cause.
4. One trial should receive that element, one trial should not receive it.
5. If you get the effect when that element is present, and do not get it when that effect is absent, you have evidence that it is playing a causal role.

If we go back to our poisoned picnic where we were left with two suspects, fried chicken and coleslaw, we can now set up a trial. We can’t test both of these at once, so we have to choose one to test at a time. If we choose the fried chicken, we want to take two friends, and try to poison one of them. (Don’t try this at home. Or in the lab. There are ethical regulations on what kinds of tests you can do on humans.)

To formalize it: find two friends who are as similar as possible, especially in relevant ways. You want them to be in similar health, and have the same food allergies. They should be around the same age, and so forth. Make sure neither of them have the stomach flu. Then, control everything they eat for 24 hours, to make sure they’re eating the same amount of the same thing at the same times. Once you have your two trials set up, perform the test: give one friend the potentially bad chicken and let the other friend get off without eating it. Examine the results:

* If the friend who ate the chicken gets sick, and the one who avoided it does not get sick, that’s some evidence that the chicken was bad.
* If both of them get sick, that’s evidence that something you controlled for might be a suspect. Did they both eat that coleslaw? Did you feed them both milk, and both are lactose intolerant? Note, the chicken still could be bad in this case.
* If neither of them get sick, that’s evidence that the chicken is fine, and the cause is something different that they haven’t consumed in the last 24 hours.

Let’s look at another example, which doesn’t violate obvious ethical regulations. Suppose I want to test to see if yeast actually has a causal role in helping bread rise. Going through steps 1 through 5 above, here’s what I’ll do:

1. Identify the effect: bread rising.
2. Identify the element you want to test as a potential cause: yeast.
3. Set up your trials, controlling for what you can: this means you want to bake two loaves of bread. You want to make sure that other than the yeast, both breads have exactly the same amounts of all the same ingredients. You want to make sure you prepare the doughs in exactly the same ways. You want to make sure you bake them at the same temperature for the same amount of time in the same type/size/shape of pan.
4. The only thing that should differ between them is that one gets yeast, and one does not.
5. The result is going to be the one with yeast shows the effect (it rises), and the one that lacks yeast did not show the effect (it did not rise).

The method of agreement is called that because you find cases that agree in that they all have the effect, and you look to see if they also all agree on an element that is a potential cause. The method of difference is called that because you want two cases to differ on whether they include the potential cause, and you hope they will also differ in whether they show the effect. With the method of agreement, you’re analyzing cases for similarities; with the method of difference, you’re purposefully trying to make the effect happen, and having a control case to make sure the thing you tested is really the cause.

### 3. Joint Method of Agreement and Difference

The joint method combines some elements from each of the above two methods. But really, it’s an expansion of the method of agreement. John Stuart Mill said this method could also be called the indirect method of difference, but also said it is essentially applying the method of agreement, but twice.

Here’s how to lay it out:

1. Identify the effect you want to explain.
2. Collect several examples with the effect present, and several examples with the effect absent.
3. Brainstorm all elements which could be potential causes for that effect.
4. If you find one element which is present every time the effect is present, and absent every time the effect is absent, you have evidence that this is the cause.

Informally, if we go back to the poisoned picnic attendees, we already have data on what everyone who got sick ate, and what they have in common. Now we want to phone the rest of the people who attended the picnic, and ask them what *they* ate. You’re looking for something the people who got sick ate, and the people who dodged food poisoning all avoided. So, let’s put up our last chart from the method of agreement, where we ended with a mixed result:

|  | Andrew | Betty | Cesar | Dominique | Elaine |
| --- | --- | --- | --- | --- | --- |
|  | Sick | Sick | Sick | Sick | Sick |
| Fried chicken | **\*** | **\*** | **\*** | **\*** | **\*** |
| Potato salad |  | **\*** | **\*** | **\*** |  |
| Coleslaw | **\*** | **\*** | **\*** | **\*** | **\*** |
| Rolls | **\*** |  |  | **\*** |  |
| Fruit salad | **\*** | **\*** | **\*** |  |  |
| Cake |  |  | **\*** | **\*** | **\*** |

Now, let’s call the people who did not get sick back, and ask what they ate:

|  | Andrew | Betty | Cesar | Dominique | Elaine | Franco | George | Haily | Isaac | Jacob |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | Sick | Sick | Sick | Sick | Sick | Not sick | Not sick | Not sick | Not sick | Not sick |
| Fried chicken | **\*** | **\*** | **\*** | **\*** | **\*** |  | **\*** |  | **\*** |  |
| Potato salad |  | **\*** | **\*** | **\*** |  | **\*** |  | **\*** | **\*** |  |
| Coleslaw | **\*** | **\*** | **\*** | **\*** | **\*** |  |  |  |  |  |
| Rolls | **\*** |  |  | **\*** |  |  | **\*** | **\*** | **\*** | **\*** |
| Fruit salad | **\*** | **\*** | **\*** |  |  | **\*** |  |  | **\*** |  |
| Cake |  |  | **\*** | **\*** | **\*** |  |  |  |  |  |

I’ve put a dark line to divide the sick from the not sick people. Notice that everyone who was sick ate the coleslaw and the chicken; everyone who did not get sick avoided the coleslaw. I’ve ruled out the chicken; George and Isaac ate chicken and remained healthy.

Mill thought this was doing the method of agreement twice – once on the group who was sick, and once on the group who did not get sick. For the group who was sick, we want to see agreement on what element is present – one food they all ate. For the group who did not get sick, we want to see agreement on what element is *absent*. One food they all managed to avoid.

### Summary of the First Three Methods

The methods of agreement, difference, and the joint method are the three main methods for attempting to prove a causal link, and to rule out suspects that have no causal effect. Let’s take a minute and summarize them, highlighting their various features, before moving on to the last two methods.

#### *Agreement*:

* Identify the effect you want to explain. Collect several cases with the effect present. List all of your suspected causes for this effect. If there is one element that is present in all of these cases, you have evidence that it could be a cause.
* Usually done by analyzing data you already have, so you can pick out the cases that have the effect.
* Example: if I want to use agreement to find evidence for the yeast/rising cause and effect, I would collect several loafs of bread that have the desired effect, and see if they all share any ingredients.

#### Difference:

* Identify the effect you want to explain, and the cause you want to test. You will set up two trials; one will have the potential cause present, one will not. Everything else must be kept as identical as possible, so that you’re only testing the one factor you are interested in. If you get the effect where the element is present, and do *not* get the effect when the element is absent, you have evidence that this could be a cause.
* Cannot be done by gathering pre-existing data; you must set up a controlled experiment.
* The example we saw above with the yeast had us bake two identical loaves of bread, only one left out the yeast, and one made sure it had yeast in it.

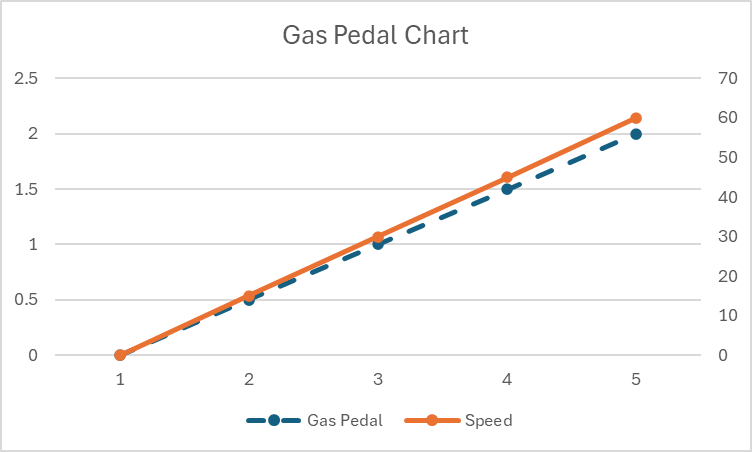
#### Joint:

* Identify the effect you want to explain, and all of your suspected causes for this effect. Collect several cases where the effect is present, and several cases where the effect is absent. If you find one element that is present wherever the effect is, and absent wherever the effect is absent, you have evidence that this is your cause.
* This method can be used to analyze pre-existing data, where you sort the cases into those with the effect and those without the effect. It also can be done with a series of experiments, where you see what combinations get you the effect and what combinations don’t.
* Example: suppose my method of agreement gave me a long list of ingredients, and four of them were present in every loaf: flour, sugar, salt, and yeast. I can bake several loaves of bread using a variety of combinations of ingredients, making sure I leave one of these out for at least one loaf, and playing with the other ingredients listed as well. I’ll get a variety of results; some will rise, and some will not. Then I can see which ingredients were present wherever it rose and absent wherever the loaf did not rise.

### 4. Concomitant Variations

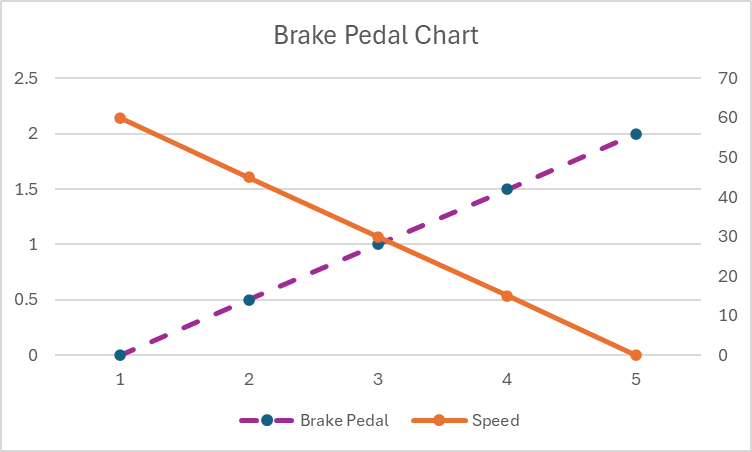
This method can be very helpful in gathering further evidence for a suspected causal connection. And, unlike the other methods, it can sometimes be used to identify causal connections you hadn’t suspected. However, *it must never be your only method of proof*. “Variations” means changes, and “concomitant” roughly means “together.” You’re looking for two things that change together.

For example, you may have noticed that when you push the gas pedal when you’re driving, the car speeds up. The more you push the pedal, the more speed you get. If you back off on the gas, the car will slow down. These two things – how much you’re pushing the gas pedal, and how fast the car is going – are changing together. We can represent this with a line graph. Below, the solid orange line will be the speed of the car in miles per hour, and the blue line with dashes will be how much I pushed the gas pedal, in centimeters (because this is science, and centimeters sound more “sciencey” than inches).



Concomitant Variations usually never end up with such a smoothly correlated data set, but I just made this data up and I wanted the chart to be pretty.

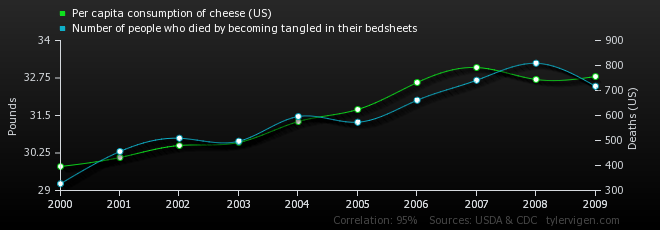
When the two things you’re measuring on your chart rise at the same time, and fall at the same time, we call this a “positive correlation.” Negative, or inverse, correlations exist as well; this is where the two things change together, but in opposite directions from each other. For example: the brake pedal has an effect on the speed of the car too, but in this case, the *more* I push the pedal, the *less* speed I get. Let’s graph this out too, assuming I start by coasting along at 60 miles an hour and not touching the brake pedal at all.



If I don’t touch the brake at all, the speed remains at 60; the more I push it, the less speed I get, and if I floor the brake pedal, the car will come to a full stop. Notice the two lines mirror each other, even though they head in opposite directions.

This method can add data to our attempt to prove causation, but it also can help us discover unsuspected cause/effect relationships. If we find a correlation we never knew existed, we might have discovered a new cause/effect relationship.

Ok cool, two things change together so they might be causally related. But why only “might be?” This is because concomitant variations establishes a correlation. And, as we learned in the false cause fallacy section above, *correlation does not prove causation*. Let me repeat a chart from that section.



Tyler Vigen has used the method of concomitant variations to establish a correlation between the amount of cheese Americans eat and the number of people who died by becoming tangled in their bedsheets. If we assume that one is causing the other, we’re committing the non causa pro causa false cause fallacy.

Our other example of a non causa fallacy was:

Every time ice cream sales go up, so do shark attacks. Sharks must really like ice-cream stuffed humans!

Here, I’ve used the method of concomitant variations as well, but in this case, the fallacy occurred because I didn’t use the method correctly. Here’s what John Stuart Mill had to say about this method:

*Whatever phenomenon varies in any manner whenever another phenomenon varies in some particular manner, is either a cause or an effect of that phenomenon, or is connected with it through some fact of causation.[[4]](#footnote-4)*

In other words, if I find a correlation between A and B, that suggests there could be a causal relationship, but there are three possible causal relationships here: A causes B, B causes A, or A and B are in some other way causally connected. So, let’s rephrase our original argument to be less fallacious.

Every time ice cream sales go up, so do shark attacks. Shark attacks and ice cream sales must be causally related!

And this conclusion is much closer to the truth! Ice cream sales are not causing the sharks to attack; shark attacks are not causing humans to buy more ice cream. Rather, these two phenomenon *share a common cause*: hotter weather causes humans to eat more ice cream, and it causes us to go swimming in shark-infested water more often. They are causally related.

Bed sheet deaths and cheese consumption are not causally related (to my knowledge), however, even though they’re well correlated. So, here’s the take-away lesson for this section:

The method of concomitant variations establishes that there is a correlation between two things, A and B. One of four possibilities is responsible for this correlation:

1. A causes B.
2. B causes A.
3. A and B are otherwise causally connected (such as sharing a common cause).
4. The correlation is a complete coincidence and there is no causal connection.

Ok, we only have one more method to explore! Of all five methods, the last one gives us the least amount of evidence – except for in very particular circumstances.

### 5. Method of Residues

The word “residues” means, roughly, “leftovers.” If you bake a pan of brownies, and remove the brownies, those crumbs still sticking to the pan are the brownie residue.

Here, we’re looking for a left-over cause, and a left-over effect. For a complicated cause-effect relationship, with many elements, you want to eliminate the parts of the effect that you can explain, by identifying what caused each part. This ends up taking a large problem and making it smaller, but it also can give you some hints as to where to look next.

Here’s an example:

Suppose a restaurant manager is looking at the budget at the end of the month and realizes they have $1,000 less than they should have. Where did this extra money go? She goes through the accounts books and discovers the following causes for parts of the missing money:

$500: replacing a broken freezer

$100: replacing the food that went bad due to a broken freezer

$200: training a new employee

We have systematically accounted for the parts of the effect that we can explain, by hooking each part of the effect up with its corresponding cause. We have not, however, explained the entire phenomenon. We’ve only explained $800 worth; there’s a $200 residue effect – the portion we were unable to explain. This is useful, however, because we took a $1,000 problem and turned it into a $200 problem. But also, it can give us hints of where to look next.

The manager systematically went through the accounts books, and explained every part of the extra expense that she could. That means that whatever happened to that $200, it’s the sort of thing that didn’t get written in the books!

Theft, of course, is one explanation. No thief ever has pocketed $200, and then has gone back to the office to record the theft in the accounts books. However, while I want to investigate theft, I also want to look for other explanations. Human error is another thing that can account for that $200 gap. Maybe some of the math has been done wrong. Maybe one of the managers made a legitimate purchase for the restaurant, and simply forgot to record it; they’re still carrying the receipt around in their wallet.

So, the method of residues can make a large problem smaller, and can give you clues of where to look next, but it can not prove that your suspected cause is the actual cause of the residue effect – except in one particular circumstance.

Mill wants us to imagine that we’ve *established* a causal link between a complicated cause and a complicated effect. So suppose we’ve used our previous methods to get evidence that ABC together produce the effect abc. Now, if we can show that A by itself can produce the effect a, and B can produce the effect b, it follows from this that the remaining part of the cause, C, must be responsible for the remaining part of the effect, c. The restaurant example is using the method of residues to try to figure out a cause. Here, though, we already know the cause, we’re just trying to untangle which part of the cause is responsible for which part of the effect.

Either way you work it, though, the method of residues requires more work proving a causal link. If you’ve established the complete cause earlier, you can use the method of residues to link up each part of the cause with each part of the effect. If you’re searching for the cause, once you have hints of where the residue might lie (theft, human error, and so on), then you have to use other methods to prove that your newly suspected cause is in fact responsible for the residue effect.

There you have it, folks, all five of Mill’s Methods for proving causation. Now instead of leaping to causal conclusions and making false cause fallacies, you have some tools at your disposal to help you gather evidence for your suspected causes.

1. I taught with this example for years before I finally looked it up to see if the premise was true. It is not. Ice cream sales and shark attacks peak in different months. However, remember that we’re looking at the connection between premises and conclusion when we evaluate logic. IF the premise were true, how much evidence would it give for the conclusion? Answer: not very much at all. [↑](#footnote-ref-1)
2. John Stuart Mill, *A System of Logic*, Vol. 1, Project Gutenberg, 2008, p. 394. Available online at this link: [A System Of Logic, Ratiocinative And Inductive (Vol. 1 of 2)](https://www.gutenberg.org/files/26495/26495-pdf.pdf) [↑](#footnote-ref-2)
3. Ibid. [↑](#footnote-ref-3)
4. John Stuart Mill, *A System of Logic, Ratiocinative, and Inductive*, originally published in 1843. You can find it at the following link (opens in a new window): [A System of Logic](https://www.gutenberg.org/ebooks/27942) [↑](#footnote-ref-4)