

# What is a confounding variable and can you provide an example?

Authored by  
**stats writer**

April 26, 2024

## RECOMMENDED CITATION

stats writer (2024). *What is a confounding variable and can you provide an example?*. PSYCHOLOGICAL SCALES. Retrieved from <https://scales.arabpsychology.com/?p=139529>

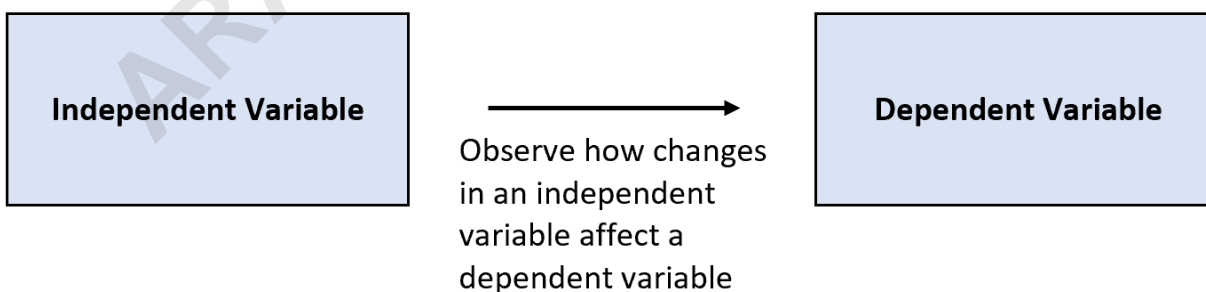
A confounding variable is an extraneous factor that affects the relationship between the independent and dependent variables in a study, making it difficult to accurately determine the true effect of the independent variable on the dependent variable. This variable can often lead to misleading or inaccurate results. For example, in a study examining the relationship between coffee consumption and heart disease, age may act as a confounding variable as older individuals may be more likely to both drink more coffee and have a higher risk of heart disease. Therefore, the true relationship between coffee consumption and heart disease may be unclear due to the influence of age.

## What is a Confounding Variable? (Definition & Example)

**In any experiment, there are two main variables:**

**The independent variable:** the variable that an experimenter changes or controls so that they can observe the effects on the dependent variable.

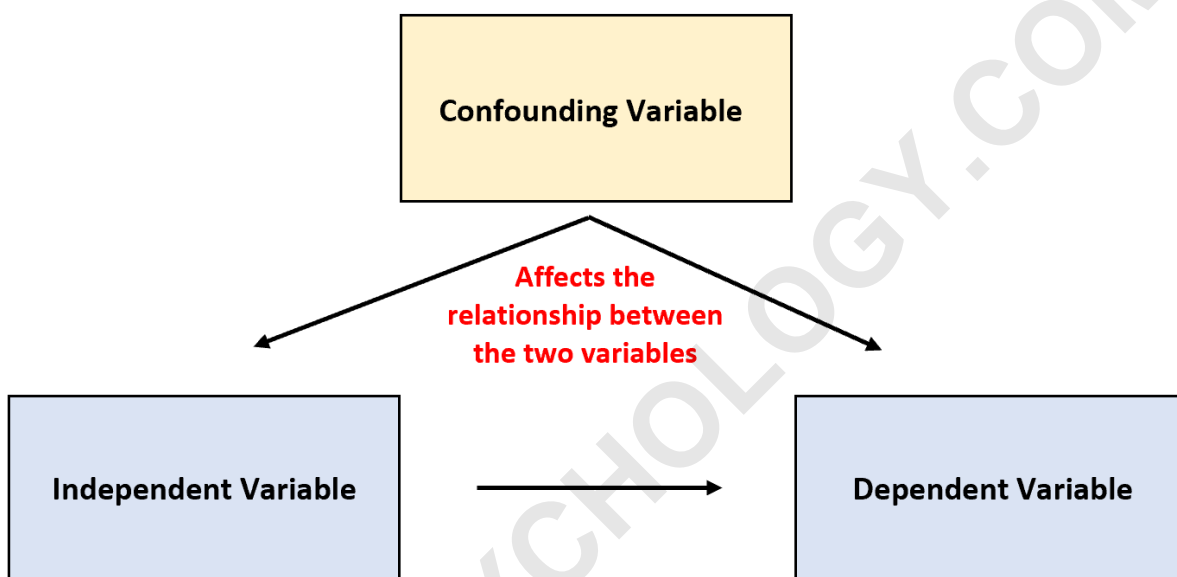
**The dependent variable:** the variable being measured in an experiment that is "dependent" on the independent variable.



**Researchers are often interested in understanding how changes in the independent variable affect the**

**dependent variable.**

**However, sometimes there is a third variable that is not accounted for that can affect the relationship between the two variables under study.**



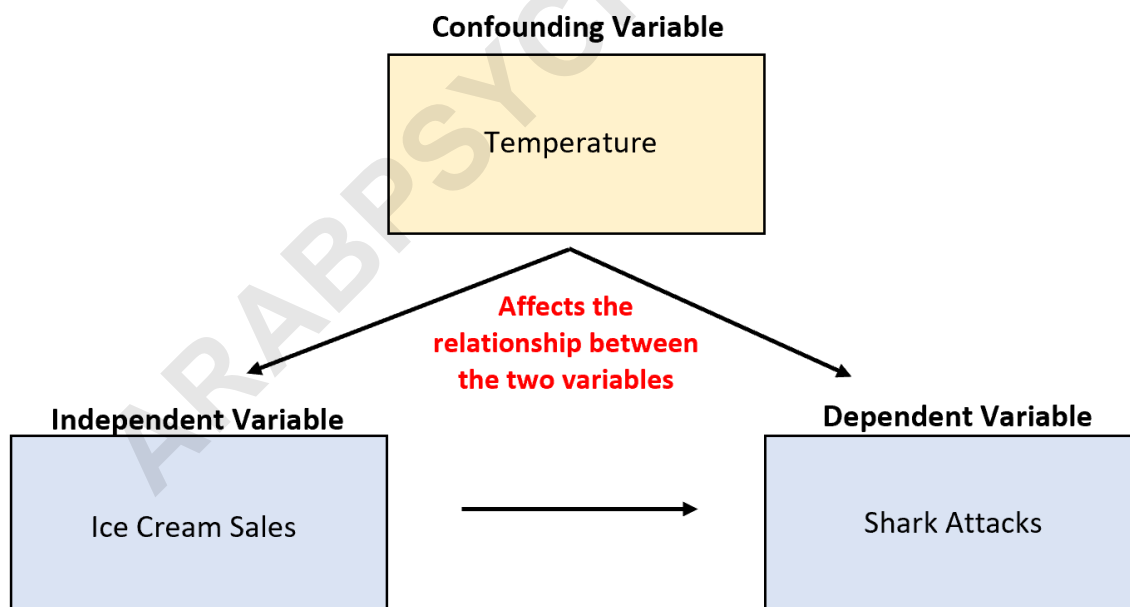
**This type of variable is known as a confounding variable and it can *confound* the results of a study and make it appear that there exists some type of cause-and-effect relationship between two variables that doesn't actually exist.**

**Confounding variable: A variable that is not included in an experiment, yet affects the relationship between the two variables in an experiment.**

This type of variable can *confound* the results of an experiment and lead to unreliable findings.

For example, suppose a researcher collects data on ice cream sales and shark attacks and finds that the two variables are highly correlated. Does this mean that increased ice cream sales cause more shark attacks?

That's unlikely. The more likely cause is the confounding variable temperature. When it is warmer outside, more people buy ice cream and more people go in the ocean.



### Requirements for Confounding Variables

In order for a variable to be a confounding variable, it

**must meet the following requirements:**

**1. It must be correlated with the independent variable.**

**In the previous example, temperature was correlated with the independent variable of ice cream sales. In particular, warmer temperatures are associated with higher ice cream sales and cooler temperatures are associated with lower sales.**

**2. It must have a causal relationship with the dependent variable.**

**Why Are Confounding Variables Problematic?**

**Confounding variables are problematic for two reasons:**

**1. Confounding variables can make it seem that cause-and-effect relationships exist when they don't.**

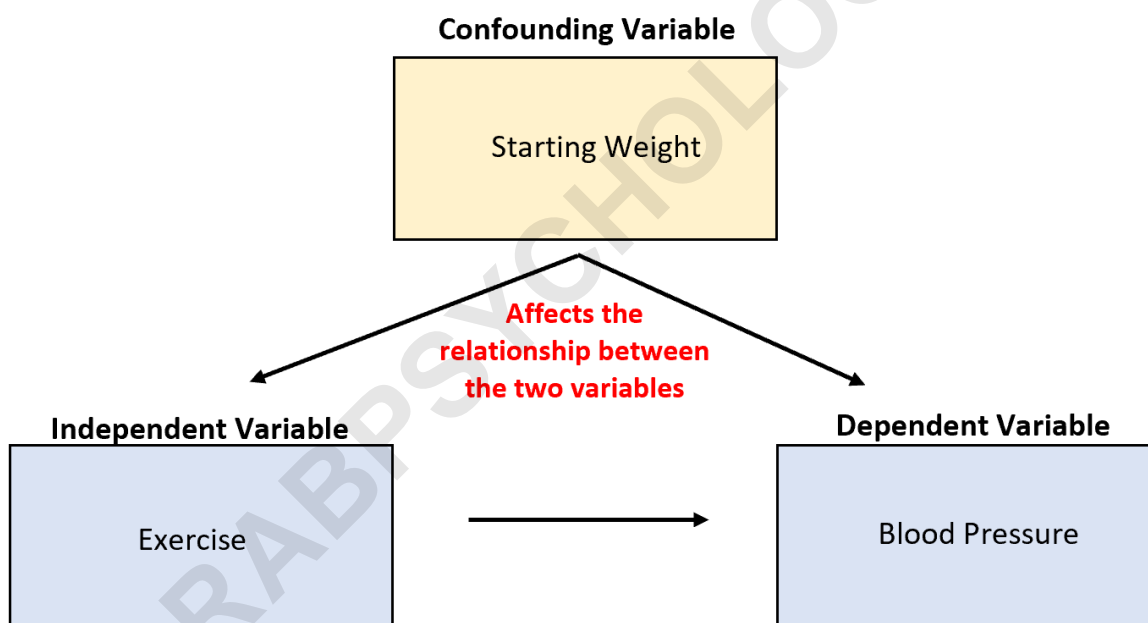
**In our previous example, the confounding variable of temperature made it seem like there existed a cause-and-effect relationship between ice cream sales and shark attacks.**

**However, we know that ice cream sales don't cause shark attacks. The confounding variable of temperature**

just made it seem this way.

**2. Confounding variables can mask the true cause-and-effect relationship between variables.**

Suppose we're studying the ability of exercise to reduce blood pressure. One potential confounding variable is starting weight, which is correlated with exercise and has a direct causal effect on blood pressure.



While increased exercise may lead to reduced blood pressure, an individual's starting weight also has a big impact on the relationship between these two variables.

**Confounding Variables & Internal Validity**

**In technical terms, confounding variables affect the internal validity of a study, which refers to how valid it is to attribute any changes in the dependent variable to changes in the independent variable.**

**When confounding variables are present, we can't always say with complete confidence that the changes we observe in the dependent variable are a direct result of changes in the independent variable.**

**How to Reduce the Effect of Confounding Variables**

**There are several ways to reduce the effect of confounding variables, including the following methods:**

### **1. Random Assignment**

**refers to the process of randomly assigning individuals in a study to either a treatment group or a control group.**

**For example, suppose we want to study the effect of a new pill on blood pressure. If we recruit 100 individuals to participate in the study then we might use a random number generator to randomly assign 50 individuals to**

a control group (no pill) and 50 individuals to a treatment group (new pill).

By using random assignment, we increase the chances that the two groups will have roughly similar characteristics, which means that any difference we observe between the two groups can be attributed to the treatment.

This means the study should have internal validity - it's valid to attribute any differences in blood pressure between the groups to the pill itself as opposed to differences between the individuals in the groups.

## 2. Blocking

refers to the practice of dividing individuals in a study into "blocks" based on some value of a confounding variable to eliminate the effect of the confounding variable.

For example, suppose researchers want to understand the effect that a new diet has on weight loss. The independent variable is the new diet and the dependent variable is the amount of weight loss.

However, a confounding variable that will likely cause variation in weight loss is gender. It's likely that the gender of an individual will effect the amount of weight they'll lose, regardless of whether the new diet works or not.

One way to handle this problem is to place individuals into one of two blocks:

**Male**  
**Female**

Then, within each block we would randomly assign individuals to one of two treatments:

**A new diet**  
**A standard diet**

By doing this, the variation within each block would be much lower compared to the variation among all individuals and we would be able to gain a better understanding of how the new diet affects weight loss while controlling for gender.

### 3. Matching

A is a type of experimental design in which we "match" individuals based on values of potential confounding

**variables.**

**For example, suppose researchers want to know how a new diet affects weight loss compared to a standard diet. Two potential confounding variables in this situation are age and gender.**

**To account for this, researchers recruit 100 subjects, then group the subjects into 50 pairs based on their age and gender. For example:**

**A 25-year-old male will be paired with another 25-year-old male, since they "match" in terms of age and gender. A 30-year-old female will be paired with another 30-year-old female since they also match on age and gender, and so on.**

**Then, within each pair, one subject will randomly be assigned to follow the new diet for 30 days and the other subject will be assigned to follow the standard diet for 30 days.**

**At the end of the 30 days, researchers will measure the total weight loss for each subject.**

Pair	Total weight loss (lbs)	
	New Diet	Standard Diet
1	3	1
2	6	4
3	7	2
4	5	5
...	...	...
99	3	4
100	8	5

By using this type of design, researchers can be confident that any differences in weight loss can be attributed to the type of diet used rather than the confounding variables age and gender.

This type of design does have a few drawbacks, including:

1. Losing two subjects if one drops out. If one subject decides to drop out of the study, you actually lose two subjects since you no longer have a complete pair.
2. Time-consuming to find matches. It can be quite time-consuming to find subjects who match on certain variables, such as gender and age.
3. Impossible to match subjects perfectly. No matter

**how hard you try, there will always be some variation within the subjects in each pair.**

**However, if a study has the resources available to implement this design it can be highly effective at eliminating the effects of confounding variables.**

ARABPSYCHOLOGY.COM