Confounding Variable / Third Variable [1]

Martyn Shuttleworth [2], Lyndsay T Wilson [3]

Confounding variables (aka third variables) are variables that the researcher failed to control, or eliminate, damaging the internal validity [4] of an experiment.

What are Confounding Variables?

A confounding variable, also known as a third variable or a mediator variable, influences *both* the independent variable [5] and dependent variable [6]. Being unaware of or failing to control for confounding variables may cause the researcher to analyze the results [7] incorrectly. The results may show a false correlation [8] between the dependent and independent variables, leading to an incorrect rejection [9] of the null hypothesis [10].

## The Problem with Confounding Variables

For example, a research group might design [11] a study to determine if heavy drinkers die at a younger age.

They proceed to design a study, and set about gathering data. Their results, and a battery of statistical tests [12], indeed show that people who drink excessively are likely to die younger.

Unfortunately, when the researchers gather data from their subjects' *non-drinking peers*, they discover that they, too, die earlier than average. Maybe there is another factor, not measured [13], that influences both drinking and longevity?

The weakness in the experimental design [14] was that they failed to take into account confounding variables [15], and did not try to eliminate or control [16] any other factors.

Imagine that in this case, there is in fact no relationship between drinking and longevity. But there may be other variables which bring about *both* heavy drinking and decreased longevity. If they are unaware of these variables, the researchers may assume that heavy drinking is causing reduced longevity, i.e. they'll make what's called a "spurious association." In reality, decreased longevity may be better explained by a third, confounding variable.

CONFOUNDING VARIABLE (THIRD VARIABLE)

INDEPENDENT VARIABLE

DEPENDENT VARIABLE

Spurious correlation
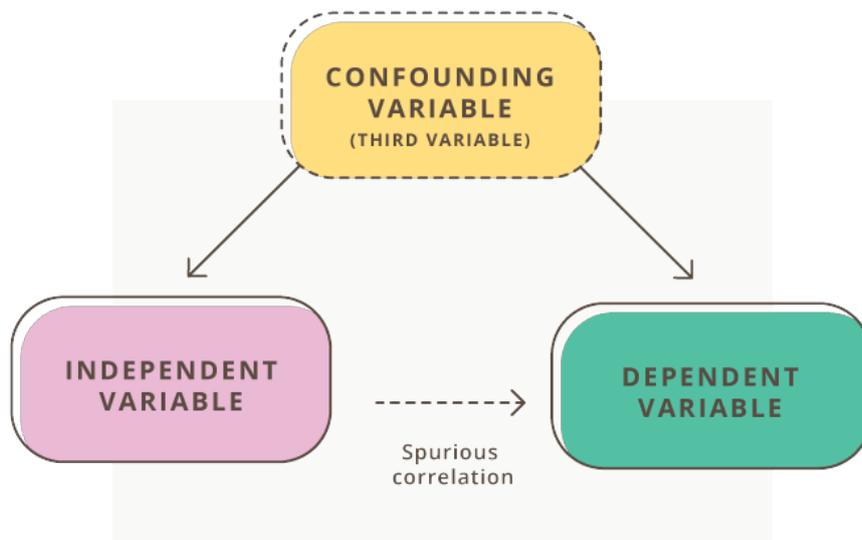
What at first looks like a causal relationship between IV and DV is ultimately spurious. The confounding variable is the hidden explanation.

For example, it is quite possible that the heaviest drinkers hailed from a different background or social group. This group might be, for unrelated reasons, shorter lived than other groups. Heavy drinkers may be more likely to smoke, or eat junk food, all of which could be factors in reducing longevity. In any case, it is the fact they belong to this group that is responsible for their decreased longevity, and not heavy drinking.

Without controlling for potential confounding variables, the internal validity of the experiment is undermined.

Extraneous Variables

Any variable that researchers are not deliberately studying in an experiment is an extraneous (outside) variable that could threaten the validity of the results. In the example above, these could include age and gender, junk food consumption or marital status.

An extraneous variable becomes a confounding variable when it *varies along with* the factors you are actually interested in. In other words, it becomes difficult to separate out which effect belongs to which variable, complicating the data.

To return to the example, age might be an extraneous variable. The researchers could control for age by making sure that everyone in the experiment is the same age. If they didn't, age would become a confounding variable.

Any time there is another variable in an experiment that offers an alternative explanation for the outcome, it has the potential to become a confounding variable. Researchers must therefore control for these as much as possible.

Minimizing the Effects of Confounding Variables

A well-planned experimental design [17], and constant checks, will filter out the worst confounding variables.

For example, randomizing groups, utilizing strict controls, and sound operationalization [18] practice all contribute to eliminating potential third variables.

After research, when the results are discussed and assessed by a group of peers, this is the area that stimulates the most heated debate. When you read stories of different foods increasing your risk of cancer, or hear claims about the next super-food, assess these findings carefully.

Many media outlets jump on sensational results, but never pay any regard to the possibility of confounding variables.

Mini-quiz:

Imagine that a research project attempts to study the effect of a popular herbal antidepressant. They sample participants from an online alternative medicine group and ask them to take the remedy for a month. The participants complete a depression inventory before and after the month to measure whether they experience any improvement in their mood. The researchers do indeed find that the participants' moods are better after a month of treatment.

Can you identify any variables which may have confounded this result? The answer is at the bottom of the page.

Correlation and Causation

The principle is closely related to the problem of correlation and causation [8].

For example, a scientist performs statistical tests, sees a correlation and incorrectly announces that there is a causal link [19] between two variables.

The problem is that the research has not actually isolated a true cause and effect relationship. It is similar to a researcher who notices that the fewer storks there are in a country, the lower the birth rate is. They would be mistaken to assume that a decrease in storks *causes* a decrease in birth rate.

Though these factors might show some correlation, it doesn't mean that one is causing the other. In fact, two variables may move with one another purely by coincidence!

Constant monitoring, before, during and after an experiment [14], is the only way to ensure that any confounding variables are eliminated.

Statistical tests, whilst excellent for detecting correlations [20], can be almost too accurate.

Human judgment is always needed to eliminate any underlying problems, ensuring that researchers do not jump to conclusions [21].

Mini-quiz Answer

The fact that the participants were sampled from a group with an interest in alternative medicine may mean that a third variable, their belief in the effectiveness of the remedy, was responsible. You may have thought of other confoudning variables. For example their mood might have improved for a number of other unrelated reasons, like a change in weather, holidays, or an improvement in personal circumstances.

Links
[1] https://app.assistertselvhjelp.no/en/confounding-variables
[2] https://app.assistertselvhjelp.no/en/users/martyn
[3] https://app.assistertselvhjelp.no/en/users/Lyndsay%20T%20Wilson
[4] https://app.assistertselvhjelp.no/internal-validity
[5] https://app.assistertselvhjelp.no/independent-variable
[6] https://app.assistertselvhjelp.no/dependent-variable
[7] https://app.assistertselvhjelp.no/statistically-significant-results
[8] https://app.assistertselvhjelp.no/correlation-and-causation
[9] https://app.assistertselvhjelp.no/type-I-error
[10] https://app.assistertselvhjelp.no/null-hypothesis
[11] https://app.assistertselvhjelp.no/research-designs
[12] https://app.assistertselvhjelp.no/statistical-hypothesis-testing
[13] https://app.assistertselvhjelp.no/scientific-measurements
[14] https://app.assistertselvhjelp.no/experimental-research
[15] http://encyclopedia.thefreedictionary.com/Confounding+variable
[16] https://app.assistertselvhjelp.no/controlled-variables
[17] https://app.assistertselvhjelp.no/true-experimental-design
[18] https://app.assistertselvhjelp.no/operationalization
[19] https://app.assistertselvhjelp.no/causal-reasoning
[20] https://app.assistertselvhjelp.no/statistical-correlation
[21] https://app.assistertselvhjelp.no/drawing-conclusions