



Neuroskeptic

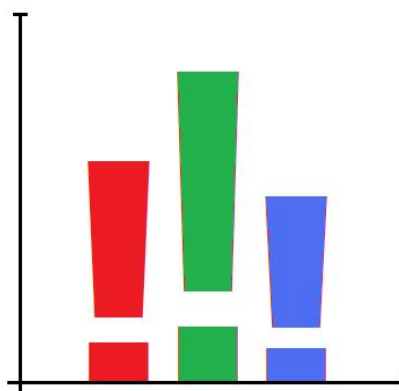
« Rogue Editors at a Psychiatry Journal?

Flexible Measures Are A Problem For Science »

Statistics: When Confounding Variables Are Out of Control

By Neuroskeptic | April 2, 2016 10:38 am

Does ice cream cause drownings? Let's think about this statistically. Consider that, in any given city, daily sales of ice cream are, most likely, positively correlated with daily rates of drownings.



Now, no matter how strong this correlation is, it doesn't really mean that ice cream is dangerous. Rather, the association exists because of a 'confound' variable. In this case it's temperature: on sunny days, people tend to eat more ice cream and they also tend to go swimming more often, thus risking drowning. The ice cream/drownings correlation would cease to exist once you take temperature into account. This means that ice cream has no 'incremental validity' over temperature: it doesn't add anything to our ability to predict drowning, above what we can predict from temperature.

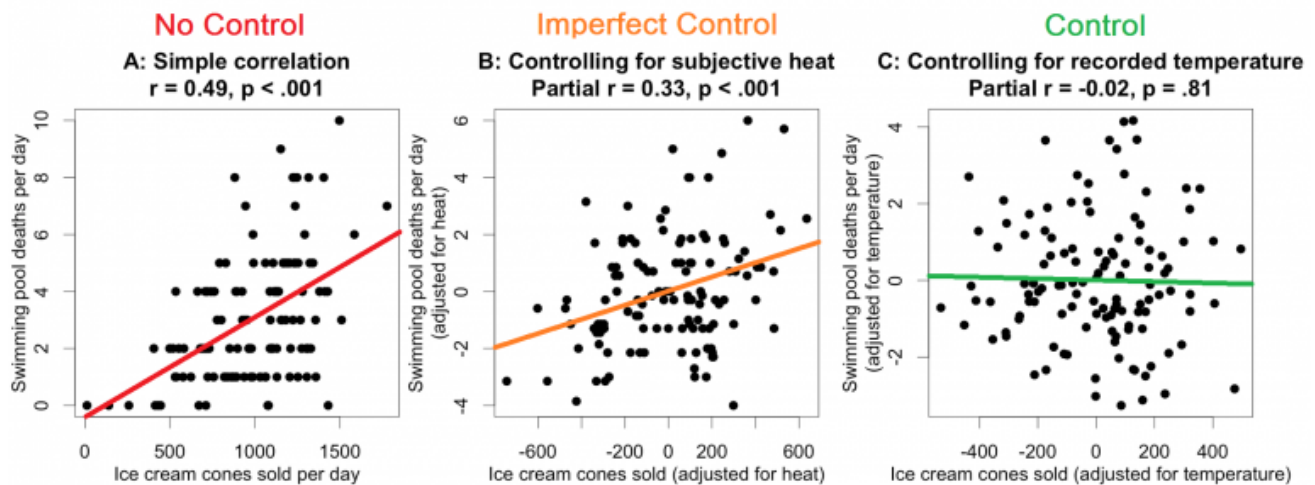
Controlling for confounds is a widely-used technique in science. However, according to researchers Jake Westfall and Tal Yarkoni, there's a major pitfall associated with the method. In a new paper, they warn that **Statistically Controlling for Confounding Constructs Is Harder than You Think**

Their argument is a simple one but the implications are rather serious: if you want to control for a certain variable, let's call it C , your ability to successfully correct for it is limited by the reliability of your *measure* of C . Let's call your experimental measure (or construct) C_m . If you find that a certain correlation holds true even controlling for C_m , this might be because C is not really a confound, but it might also be that C_m is a poor measure of C . "Controlling for C_m " is not the same as "Controlling for C ".

Staying with the ice cream/drowning example, for instance, suppose that we had a broken thermometer, meaning that our measure of temperature was noisy. Ice cream sales might well predict drownings even after controlling for our flawed 'temperature' variable. We might, then, conclude that ice cream and drownings have some deep connection beyond temperature.

Here's Westfall and Yarkoni's illustration of the problem. On the left we see the original icecream-drowning correlation, on the right the zero correlation after correcting for C , temperature. In the middle we see that the correlation remains (albeit smaller) with some hypothetical imperfect measure of

temperature, C_m .



Based on various analyses of real and generated data, Westfall and Yarkoni conclude that many scientists have been controlling for C_m and wrongly interpreting this as ‘controlling for C ’, thus wrongly concluding that incremental validity has been shown.

Literally hundreds of thousands of studies spanning numerous fields of science have historically relied on measurement-level incremental validity arguments to support strong conclusions about the relationships between theoretical constructs. The present findings inform and contribute to this literature – and to the general practice of “controlling for” potential confounds using multiple regression – in a number of ways.

First, we show that the traditional approach of using multiple regression to support incremental validity claims is associated with extremely high false positive rates under realistic parameter regimes. Researchers relying on such arguments will thus often conclude that one construct contributes incrementally to an outcome, or that two constructs are theoretically distinct, even when no such conclusion is warranted...

Taken as a whole, our results demonstrate that drawing construct-level inferences about incremental validity is considerably more difficult than most researchers recognize. We do not think it is alarmist to suggest that many, and perhaps most, incremental validity claims put forward in the social sciences to date have not been adequately supported by empirical evidence, and run a high risk of spuriousness.

The authors note that they were not the first to discuss the issue. In the past, this problem has been known as ‘**residual confounding**’, amongst other names.

So what’s the solution? Westfall and Yarkoni say that the answer is structural equation modelling (**SEM**), ideally drawing on multiple different measures (or indicators) to estimate the confounding variable better. However, even if only one confound measure is available, “*researchers can use an SEM approach to estimate what level of reliability must be assumed in order to support the validity of one’s inferences.*”

The point is that when a scientific argument rests on the failure of controlling for a confound to affect the results, a straightforward correlation analysis is not enough.

Westfall J, & Yarkoni T (2016). Statistically Controlling for Confounding Constructs Is Harder than You



Think. *PloS ONE*, 11 (3) PMID: **27031707**

CATEGORIZED UNDER: **PAPERS, SELECT, STATISTICS, TOP POSTS, UNCATEGORIZED**

