

Nudge theory doesn't work after all, says new evidence review – but it could still have a future

At the end of last year (2021), there was lots of excitement about the first comprehensive analysis of past research on techniques designed to change people's behaviour (known as “nudging”), confidently showing that they work. This was great news for researchers, but also for governments across the world who have invested in “nudge units” that use such methods.

Nudges aim to influence people to make better decisions. For example, authorities may set a “better” choice, such as donating your organs, as a default. Or they could make a healthy food option more attractive through labelling.

But new research reviewing this paper – which had looked at 212 published papers involving more than 2 million participants – and others now warns nudges may not have any effect on behaviour at all.

To understand why, we need to go into some details about statistics, and how experimental findings are analysed and interpreted. Researchers start off with a hypothesis that there is no effect (null hypothesis). They then ask, what is the probability of getting an actual effect by chance?

Our mission is to share knowledge and inform decisions.

About us

So, if in my experiment there is a group of people who are exposed to a specific nudge technique, and a control group that isn't nudged, my starting point is that the two groups won't differ. If I then find a difference, I use statistics to work out how probable it is that this would have happened by chance alone. This is called the P-value, and the lower it is the better. A big p-

value would mean that the differences between the two groups can largely be explained by chance.

The opposite is true for effect sizes. It is also important to measure the size of the effect to assess the practical value of an experiment. Imagine I am testing a nudge that is supposed to help obese people reduce their weight, and I observe that people in the nudged group lose a pound over the course of six months. While this difference may be significant (I obtain a low p-value), I might rightly ask whether this effect is big enough for any practical purposes.

So whereas p-values provide us with an indication of how likely an observed difference is by chance alone, effect sizes tell us how big – and therefore how relevant – the effect is.

A good study needs to show a moderate or large effect size, but it also needs to set out how much of it was the result of “publication bias”. This is the cherry-picking of results to show a win for nudge, meaning that studies finding that nudges don't work aren't included or even published in the first place. This may be because editors and reviewers at scientific journals want to see findings showing that an experiment worked – it makes for more interesting reading, after all.



Nudges were used to encourage social distancing. [Travelpixs/Shutterstock](#)

The authors of the original 2021 study, which reported a moderate effect size of nudging on behaviour, ruled out publication bias that was severe enough to have a major influence on the reasonable effect size they found.

Trouble for nudge

Two things have happened since though. This year, a colleague and I highlighted that, regardless of the 2021 results, there are still general issues with nudge science. For example, scientists overly rely on certain types of experiments. And they often don't consider the benefits relative to the actual costs of using nudges, or work out whether nudges are in fact the actual reason for positive effects on behaviour.

Many researchers also started becoming increasingly suspicious about the reported effect size of the 2021 study. Some called for the paper to be retracted after finding out the data analysed appeared to include studies that had used faked data.

And now a new study, published in PNAS, has re-examined the estimated impact of publication bias in the 2021 study. The authors of the new paper used their own statistical methods and

assessed the severity of publication bias as well as its impact on the actual effect size. They showed that the original effect size of all 212 studies wasn't actually moderate – it was zero.

How bad is all this? From a scientific perspective, this is excellent. Researchers start a process of gathering data to inform general assumptions about the effectiveness of nudges. Other researchers inspect the same data and analyses, and then propose a revision of the conclusions. Everything advances in the way science should.

How bad is this for nudge? Investment in it is huge. Researchers, governments, as well as organisations such as the World Health Organisation use nudges as a standard method for behavioural change. So, an enormous burden has been placed on the shoulders of nudgers. This may also have resulted in the serious publication bias, because so many were invested in showing it to work.

Right now, the best science we have is seriously questioning the effectiveness of nudging. But many, including myself, have long known this – spending many years carefully commenting on the various ways research on nudging needs to improve, and have been largely ignored.

That said, efforts to use behavioural interventions need not be abandoned. A better way forward would be to focus on building an evidence base showing which combinations of nudges and other approaches work together. For example, as I have shown, combinations of nudging methods together with changes in taxation and subsidies have a stronger effect on sustainable consumption than either being implemented alone.

This takes the burden off nudge being solely responsible for behavioural change, especially since alone it doesn't do much. In fact, how could it? Given how complex human behaviour is, how could one single approach ever hope to change it? There's not a single example of this being successfully done in history, at least not without impinging on human rights.

As I have shown before, if we are honest about the possibility of failure, then we can use it to learn what to do better.

Author



1. **Magda Osman**
Principal Research Associate in Basic and Applied Decision Making, Cambridge Judge Business School

Disclosure statement

ESRC, Research England, EPSRC, British Academy, NIHR