

NeuroChambers

Monday, 26 March 2012

You can't replicate a concept

Recent argy-bargy about a failed replication has exposed a disturbing belief in some corners of psychological research: that one experiment can be said to "conceptually replicate" another, even if it uses a completely different methodology.

John Bargh, a professor of psychology at Yale, made waves recently with a [stinging attack on virtually everyone](#) associated with a failed attempt to replicate one of his previous findings. The specifics of this particular *tango de la muerte* can be [found elsewhere](#), and I won't repeat them here, except to say that I thought Bargh's misrepresentation of the journal PLoS One was outrageous, offensive, and an extraordinary own goal.

That aside, Bargh-gate has drawn out a more important issue on the idea of "conceptual replication". [Ed Yong's article, and the comments beneath](#), exposed an unusual disagreement, with some (including [Bargh himself](#)) claiming that Bargh et al.'s original findings had been replicated at length, while others claimed that they had never been replicated.

How is this possible? Clearly something is awry.

All scientists, and many non-scientists, will be familiar with the basic idea of replication: that the best way to tell whether a scientific discovery is *real* is to repeat the experiment that originally found it. Replication is one of the bedrocks of science. It helps scientists achieve consensus and it acts like an immune system, eliminating findings that are irreproducible due to methodological error, statistical error or fraud.

It also goes without saying that the most important aspect of replication is to repeat the original experiment as closely as possible. This is why scientific journal articles contain a Method section, so that other scientists can precisely reproduce your experimental conditions.

Enter the notion of "conceptual replication". If you are a scientist and you've never heard of this term, you are not alone. The other day I did a straw poll of my colleagues – who are mostly experimental psychologists and neuroscientists – and got blank looks in response.

The basic idea is this: that if an experiment shows evidence for a particular phenomenon, you can "conceptually" replicate it by doing a completely different experiment that someone – the experimenter, presumably – believes measures a broadly similar phenomenon. Add a pinch of assumption and a healthy dose of subjectivity, and *viola*, you've just replicated the original 'concept'.

I must admit that when I first heard the term "conceptual replication", I felt like shining a spotlight on the clouds and calling for [Karl Pilkington](#). Psychology is already well known for devaluing replication and we do ourselves no favours by attempting to twist the notion of replication into something it isn't, and shouldn't be.

Here are four reasons why.

1. Conceptual replication is assumption-bound and subjective

From a logical point of view, a conceptual replication can only hold if the different methods used in two different studies are measuring the *same* phenomenon. For this to be the case, definitive evidence must exist that they are. But how often does such evidence exist?

Even if we meet this standard (and the bar seems high), how similar must the methods be for a study to qualify as being conceptually replicated? Who decides and by what objective criteria?

2. Conceptual replications can be "unreplicated"

A reliance on conceptual replications can be easily shown to produce absurd conclusions.

Consider the following scenario. We have three researchers, Smith, Jones, and Brown, who publish three scientific papers in a sequence.

Smith gets the ball rolling by showing evidence for a particular phenomenon.

Jones then comes along and uses a different method to show evidence for a phenomenon that *looks a bit like* the one that Smith discovered. The wider research community decide that the similarity crosses some subjective threshold (oof!) and so conclude that Jones conceptually replicates Smith.

Enter Brown. Brown isn't convinced that Smith and Jones are measuring the same phenomenon and hypothesises that they could actually be describing different phenomena. Brown does an experiment and obtains evidence suggesting that this is indeed the case.

We now enter the ridiculous, and frankly embarrassing, situation where a finding that was previously replicated can become *unreplicated*. Why? Because we assumed *without evidence* that Smith and Jones were measuring the same phenomenon when they were not. It's odd to think that a community of scientists would actively engage in this kind of muddled thinking.

Blog Archive

- [2022](#) (2)
- [2020](#) (1)
- [2019](#) (2)
- [2018](#) (3)
- [2017](#) (5)
- [2016](#) (9)
- [2015](#) (4)
- [2014](#) (3)
- [2013](#) (7)
- ▼ [2012](#) (17)
 - [December](#) (2)
 - [October](#) (2)
 - [September](#) (2)
 - [August](#) (1)
 - [July](#) (1)
 - [June](#) (1)
 - [May](#) (3)
 - [April](#) (2)
 - ▼ [March](#) (3)
 - [Brain stimulation to solve impossible problems?](#)
 - [You can't replicate a concept](#)
 - [It's not a job.](#)

About Me



Chris Chambers

I'm a psychologist and neuroscientist at the School of Psychology, Cardiff University. I

created this blog after taking part in a debate about science journalism at the Royal Institution in March 2012. The aim of my blog is give you some insights from the trenches of science. I'll talk about a range of science-related issues and may even give up a trade secret or two. You can follow me on Mastodon: <https://scicomm.xyz/@chrisdc77>

[View my complete profile](#)

3. Conceptual replications exacerbate confirmation bias

Conceptual replications are vulnerable to a troubling confirmation bias and a logical double-standard.

Suppose two studies draw similar conclusions using very different methods. The second study could then be argued to "conceptually replicate" the first.

But suppose the second study drew a very different conclusion. Would it be seen to conceptually *falsify* the first study? Not in a million years. Researchers would immediately point to the multitude of differences in methodology as the reason for the different results. And while we are all busily congratulating ourselves for being so clever, Karl Popper is doing somersaults in his grave.

4. Conceptual replication substitutes and devalues direct replication

I find it depressing and mystifying that direct replication of specific experiments in psychology and neuroscience is so vital yet so grossly undervalued. Like many cognitive neuroscientists, I have received numerous rejection decisions over the years from journals, explaining in reasonable-sounding boilerplate that their decision "on this occasion" was due to the *lack of a sufficiently novel contribution*.

Replication has no place because it is considered boring. Even incremental research is difficult to publish. Instead, reproducibility has been trumped by novelty and the quest for breakthroughs. Certainty has given way to the X factor. At dark moments, I wonder if we should just hand over the business of science to Simon Cowell and be done with it.

The upshot

First, we must jettison the flawed notion of conceptual replication. It is vital to seek converging evidence for particular phenomena using different methodologies. But *this isn't replication*, and it should never be regarded as a substitute for replication. Only experiments can be replicated, not concepts.

Second, journals should specifically ask reviewers to assess the likely reproducibility of findings, not just their significance, novelty and methodological rigor. As reviewers of papers, we should be vocal in praising, rather than criticising, a manuscript if it directly replicates a previous result. Journal editors should show some spine and actively value replication when reaching decisions about manuscripts. It is not acceptable for the psychological community to shrug it's shoulders and complain that "*that's the way it is*" when the policies of journal editing and manuscript reviewing are entirely in our own hands.

Psychology can ill afford the kind of muddled thinking that gives rise to the notion of conceptual replication. The field has taken big hits lately, with prominent fraud cases such as Diederik Stapel producing very bad publicity. The irony of the Stapel case is that if we truly valued *actual* replication, rather than Krusty-brand replication, his fraud could have been exposed years sooner and before he had made such a damaging impact. This makes me wonder how many researchers have, with the very best of intentions, fallen prey to the above problems and 'conceptually' replicated Stapel's fraudulent discoveries?

Posted by [Chris Chambers](#) at [14:33](#)

7 comments:



Joseph Brooks 26 March 2012 at 17:52

I can say that I also had never heard of the idea of "conceptual replication" and agree that the idea is fatally flawed for the reasons that you clearly spell out.

Science clearly has good reason to give incentives for novel, ground-breaking findings that push the boundaries of science. Thus, it makes sense to reward scientists who produce such findings. However, a flood of novel findings needs a further process to select which ones actually hold up to long-term scrutiny and comparison with other evidence. This is the role of replication (and replicate and extend). But this can only happen when researchers are given incentives which are aligned with that goal. The balance of incentives has tipped to far in the direction of novelty. There is room for both incremental, replication work as well as broadly creative and novel work within science. We need both in a balance. But in order to get both, we need to fund and reward both!

[Reply](#)

[Replies](#)



Chris Chambers 26 March 2012 at 18:04

Thanks for commenting, Joe.

I agree that we need incentives to pursue direct replications because at the moment all we have are powerful disincentives.

So how can we encourage replication...What about a replication index for specific journals (or even specific researchers) that reports the extent to which specific experiments and key findings are replicated? Or, because scientists flock to any publication with the word "Nature" in the title like moths to a flame, how about a new journal Nature Replications? Or a funding agency specifically set up to fund direct replications of existing work?

[Reply](#)

Cedar Riener 29 March 2012 at 13:13

I hadn't hear of conceptual replication until this recent kerfuffle either. I had always referred to this process as called "converging evidence" which I think is a better term for it. The evidence relates, it converges on a concept, but it doesn't try to leverage the word "replication" for increased credibility.

Thanks for this post, I am going to add it to [my link roundup and storify](#) of the whole Bargh/Doyen business. In retrospect, I found myself defending conceptual replication by simply reframing it in my head as the value of converging evidence in science (whatever name we use for it). When Ed Yong had his doubts about weak studies propping each other up, he could have just as easily been worrying about converging evidence (using different measures, or different manipulations). But I agree that calling it "replication" is not kosher, and devalues direct replication (which already has a hard time anyways with current publication biases).

[Reply](#)



Chris Chambers 31 March 2012 at 10:40

Thanks Cedar, I enjoyed your very thorough round-up piece - and thanks for including a link back to my blog.

[Reply](#)

Gabi Jiga-Boy 5 April 2012 at 17:07

Bargh's comments are sometimes extremely arrogant and nasty personal attacks on Doyen et al. - and I think this does him a huge disservice by focusing us folks on how personally threatened he is, instead of on the content of his remarks. But set aside that, I'm not at all convinced that (1) he's wrong about what he criticizes in Doyen et al.'s paper and (2) he misrepresents the journal PLoS one. Which means I think you're wrong (-)) when you're finding him "outrageous, offensive, and an extraordinary own goal" about PLoS one.

How do you think we should protect our science from journals that publish stuff that's not reviewed by professionals in our fields? I know there are endless power and political issues with publishing work that's against the tide, and the Doyen work possibly ended up here because it didn't have enough influence to pass through the gates of the political interests of powerful people in the field of semantic priming. But still, I don't trust PLoS one. I don't trust it because I expect more than a "test of scientific soundness" from a journal that expects to be taken seriously. I expect it to have reviewers that know what I mean when I say "social cognition", for example, not reviewers who look at my 2 x 2 design and say ok, she's done it right. As a scientist, I need reviews from my peers, preferably people who are not offended when I show their work was crap, to sieve my work and set aside my own biases and the junk I can infuse my work with. So only from this point of view (there may be tons of others), I need reliable peer-review. Then when someone points at a lack of peer-review, why does it make more sound than it should??

[Reply](#)



nmmichalak 12 August 2015 at 15:59

Convergent evidence and conceptual replication are the same thing. Sure, replication is synonymous with duplication, cloning, copying, mirroring, etc., but the term conceptual replication doesn't hide that it's not an exact replication. If you haven't heard the term before, you might be puzzled about its definition, but you know it's something along the lines of, "replication, but not quite."

You're right to criticize the paucity of replicatons of specific effects (e.g., exact replications), but conceptual replications (i.e., convergent evidence) are necessary for scientific progress.

[Reply](#)



nmmichalak 12 August 2015 at 16:00

Convergent evidence and conceptual replication are the same thing. Sure, replication is synonymous with duplication, cloning, copying, mirroring, etc., but the term conceptual replication doesn't hide that it's not an exact replication. If you haven't heard the term before, you might be puzzled about its definition, but you know it's something along the lines of, "replication, but not quite."

You're right to criticize the paucity of replicatons of specific effects (e.g., exact replications), but conceptual replications (i.e., convergent evidence) are necessary for scientific progress.

[Reply](#)



Enter comment

[Newer Post](#)

[Home](#)

[Older Post](#)

Subscribe to: [Post Comments \(Atom\)](#)