



# Replication celebration: rewarding replication to improve reproducibility

HANNAH HOBSON<sup>1</sup>

1. MRC Social, Genetic and Developmental Psychiatry Centre, Institute of Psychiatry, Psychology and Neuroscience, King's College London, 16 De Crespigny Park, London, SE5 8AF

READ REVIEWS

WRITE A REVIEW

CORRESPONDENCE:

[hannah.hobson@kcl.ac.uk](mailto:hannah.hobson@kcl.ac.uk)

DATE RECEIVED:

June 02, 2016

DOI:

10.15200/winn.146487.79727

ARCHIVED:

June 02, 2016

KEYWORDS:

#LJAReproducibility,  
replication, reproducibility

CITATION:

Hannah Hobson, Replication celebration: rewarding replication to improve reproducibility, *The Winnower* 3:e146487.79727, 2016, DOI: 10.15200/winn.146487.79727

© Hobson This article is distributed under the terms of the [Creative Commons Attribution 4.0 International License](#), which permits unrestricted use, distribution, and redistribution in any medium, provided that the original author and source are credited.



Biomedical science is recovering from an existential crisis. Recent reports have highlighted major problems with the reproducibility and reliability of a diverse range of scientific fields, including my own discipline, psychology and neuroscience. These issues have arisen due to a range of factors, some of which are specific to certain fields, while others apply across biomedical science (the Academy of Medical Science's 2015 report "Reproducibility and reliability of biomedical research: improving research practice," provides an excellent overview of the key issues involved). One factor is a lack of replication studies, which stems from an incentive structure built into academia that discourages scientists from replicating.

Replication should be the foundation for our confidence in scientific literature. In an ideal world, science would be self-correcting; even if one study finds an erroneous effect by chance, subsequent studies that do not produce the original effect will lead the field to accept that the original findings were incorrect. Replications are thus one of science's important homeostatic mechanisms – they keep scientists pursuing reliable lines of enquiry, and can save both research time and money. Yet scientists that replicate are arguably not rewarded for their efforts. A scientist who has conducted a replication may struggle to get their work published in high-impact journals, as many have a policy to *only* publish original research. In my own field, a recent case documented a journal editor refusing to send out a replication of a high-profile study on pre-cognition by Bem (2011b) to peer review. This was because "This journal does not publish replication studies, whether successful or unsuccessful" and "We don't want to be the Journal of Bem Replication" (Aldous, 2011). Having fewer publications, or having to settle for lower-impact journals, will negatively affect a scientist's career prospects. They may win fewer grants, or miss the opportunity to secure permanent positions. Ultimately they may be driven out of science, or leave quietly to pursue a vocation that appreciates their careful and thorough efforts. By undervaluing replication, we miss out on the publication of important work, and risk losing our most diligent colleagues.

Clearly science needs originality and creativity – we need researchers that come up with innovative study designs, new questions, different approaches. But having a work force that disproportionately engages in novel experiments and a publishing culture over-obsessed with originality risks building bodies of scientific literature on sand.

## WHAT MAKES A GOOD REPLICATION?

Of course, it is vital to note that not all replications are inherently excellent, simply by virtue of being

replications. Firstly, researchers must carefully consider whether their replication study should take the form of a “direct” replication or a “conceptual” replication. Direct replications are exact copies of the original research, while conceptual replications attempt to reproduce the original effect with a slightly different paradigm. For example, in a psychology experiment, this may mean using slightly different stimuli, rather than the exact set used in the original study. Using a direct or a conceptual replication may be more or less appropriate depending on whether the researcher wants to test the *reliability* or the *validity* of the effect in question. Directly replicating an effect may suggest such an effect is reliable, but it does not ensure that the interpretation is valid.

In addition, two further considerations for making a replication excellent are statistical power and pre-registration. The low statistical power of many studies has been another much discussed issue in forums on the reproducibility crisis. Smaller samples sizes mean that significant effects are more likely to be false-positives than if an appropriate amount of data-points had been collected (Button et al., 2013). Replications that are badly powered thus do not function properly as science’s self-correcting mechanism. Therefore, a study attempting to replicate must be sufficiently powered, in order not to add to the already cacophonous noise of a literature swimming with false-positives. Scientists embarking on a sufficiently-powered replication study should also strongly consider pre-registering their work, either through interfaces such as the Open Science Framework, or formally with a scientific journal offering publication via pre-registration. Replications (both direct and conceptual) are ideal candidates for pre-registration practices, and formal pre-registration with a journal will offer researchers surety that their results will be published.

## ENCOURAGING REPLICATIONS IN THE FUTURE

Attitudes towards replication may be improving. Considering the field of psychology, Makel, Plucker, and Hegarty (2012) examined the rate of replications (direct and conceptual) published in the top 100 psychology journals, and calculated that replications make up roughly 1.07% of published studies. This may sound quite low, but the authors note that replication has been on the rise, especially since 2010 (perhaps arguably when the extent of the reproducibility crisis in psychology started to become clear).

A number of suggestions have been made for how to continue to increase the number of replications conducted and published. Koole and Lakens (2012) suggest a three-pronged approach of co-publication, co-citation and collaboration. They argue that journals should have dedicated spaces for replications (rather than ostracizing all replications to their own extremely general “journals of replication” that no one will ever read), and that replications should be automatically cited with the original studies. Finally, they describe rare but admirable examples of “adversarial collaboration”, in which labs with contradictory results team up to design critical experiments. They note that replications are often associated with hostility – a failed replication is like throwing down a glove to the original research team, implying that their published effect is “wrong”. This need not be so. Scientists could enter into collaborations with a spirit of wanting to understand *why* groups found differing effects. Doing so may uncover important insights in their own right. Instead of feeling attacked by replication, the reproduction of a study by an independent, non-overlapping team of researchers is important to guarding against experimenter bias – the study by Makel et al. (2012) noted that when at least one author was on both the original and the replication article, only 3 out of 167 failed to replicate the initial findings of the first paper. Replication need not be seen as a challenge to one’s scientific pride, but as the gold-standard process by which research findings are confirmed.

Supporting and encouraging replication will require rebalancing incentives in academic culture. I propose that one seemingly simple means to demonstrate that the scientific community needs and appreciates replication would be rewarding scientists that undertake high-quality replications with appropriate prizes. Good replications (including those that are sufficiently powered and pre-registered) should receive awards that carry weight and recognition. Prestigious awards from major funding bodies or high-impact journals will help to elevate the profile of replication studies to the level of importance they surely deserve. In addition, smaller, field-specific awards would help neuroscientists, cell biologists, geneticists and so on who have conducted high-level replications become known within their

own disciplines for doing so, creating torch-bearers for promoting this crucial scientific endeavour. Being able to list such awards on CVs will help scientists that undertake this important work to benefit from doing so when attempting to move up the next stage of the career ladder. Such recognition should be available to graduate students right through to senior academics, to encourage replication at every stage of the scientific career.

Of course, a small number of awards is unlikely to radically shift the behaviour of most scientists – the odds of winning such a prize do not negate the current drawbacks associated with choosing to do a thorough replication over a novel study. You cannot win an award for a study that you cannot get published, if journals refuse to accept replication work. That is why these prizes must come in the context of other changes that both support and encourage scientists to embark on replication studies. That includes changes in journals' policies to only publish novel experiments, opening up new sources of funding to enable scientists to undertake replications, and building replication into the scientific curriculum to ensure the next generation of researchers appreciate the importance of replications and are equipped to execute them. We must stop treating replications as dull, second-fiddle to the "real" scientific endeavours of novel experiments. Let us celebrate our good examples of replication, and the careful scientists that undertake them.

## REFERENCES

- Academy of Medical Sciences. (October, 2015). Reproducibility and reliability of biomedical research: improving research practice.
- Aldhous, P. (5 May, 2011) Journal rejects studies contradicting precognition. New Scientist. Retrieved from: <https://www.newscientist.com/article/dn20447-journal-rejects-studies-contradicting-precognition/>
- Bem, D.J. (2011b). Feeling the future: Experimental evidence for anomalous retroactive influences on cognition and affect. *Journal of Personality and Social Psychology*, 100 (2011): 407–425. doi: [10.1037/a0021524](https://doi.org/10.1037/a0021524)
- Button, K. S., Ioannidis, J. P. a, Mokrysz, C., Nosek, B. a, Flint, J., Robinson, E. S. J., & Munafò, M. R. (2013). Power failure: why small sample size undermines the reliability of neuroscience. *Nature Reviews. Neuroscience*, 14(5): 365–76. <http://doi.org/10.1038/nrn3475>
- Koole, S. L., & Lakens, D. (2012) Rewarding Replications: A Sure and Simple Way to Improve Psychological Science. *Perspectives on Psychological Science*, 7(6),. 608–614. <http://doi.org/10.1177/1745691612462586>
- Makel, M. C., Plucker, J. A., & Hegarty, B. (2012)"Replications in Psychology Research: How Often Do They Really Occur? *Perspectives on Psychological Science*, 7(6): 537–542. <http://doi.org/10.1177/1745691612460688>