Credit: Super Formosa
Photography

Reward PhDs' high-quality, slow science

Publishing novel, eye-grabbing results is rewarded in academia; whether publishing robust replications will be rewarded by graduation committees and future employers is yet to be determined. Andrea Stoevenbelt calls on committees to consider how different publications are weighed on candidates' CVs.

The current crisis of confidence in psychology, wherein many prominent findings fail to replicate, can be partially attributed to the pressure to publish novel, eye-grabbing results rather than doing rigorous research and robust replications. During my PhD, I want to contribute to counteracting this crisis, and therefore I am currently working on a large-scale replication project. Whether dedicating my PhD to one large but robust project will make me competitive on the job market—where publications count—is entirely unclear.

My study investigates stereotype threat. This is a described effect whereby groups that are stereotypically—but unjustly—presumed to do worse in one task or another do indeed perform worse when these stereotypes are made salient to them. If stereotype threat has strong effects in real life, this could help to explain the lower performance of negatively stereotyped ethnic minority or female students in several academic fields. Like other fields in psychology, the stereotype threat literature is plagued by small sample sizes, suboptimal research methods and publication bias (with 'positive' results being more likely to be published, which means that established effects might be spurious, but that the evidence indicating this is suppressed). It is unclear whether stereotype threat is robust, and my project, a registered replication report (RRR), is set up to address these concerns.

The replication is pre-registered to make it impossible for us to cherry-pick positive results for publication. It will feature a large sample size (aiming at over 1,200 participants) and be subjected to extensive piloting and expert peer review before data collection. The procedure sets my RRR in sharp contrast to standard studies in the field, in which small samples, untested materials and analytical choices that are made when the data is available (leading researchers to trick themselves and others into believing in spurious results) are the norm.

However, hand-in-hand with these improved study-to-study differences comes

an existential disadvantage of the RRR in the current academic system: my time investment. Where standard studies can be readily conducted over much shorter time spans, my RRR will take over two years of full-time work to develop and to collect data. This is partially due to the extended review process and the fact that replications are frequently held to higher methodological standards than novel work. The pilot of our study has a larger sample size than most actual studies in the literature. In the meantime, I could have conducted multiple smaller status quo experiments and chosen not to pre-register them. And I probably would have been able to publish multiple novel significant results to advance my career. But such results would probably not have stood up to scrutiny in later replication research, thereby inadvertently contributing to the crisis of confidence. I am convinced that this standard approach will advance neither science nor my own résumé in the long run. Rather I choose to do an RRR to minimize flaws.

An RRR is an investment that will yield an important result for science as a whole. However, despite this desirable outcome, I worry about my publication record. In the end, I fear that hiring or grant committees will see the RRR as just a single line on my CV, of similar impact as a standard stereotype threat experiment that could have led to an impressive publication list. If this is the case, it seems that I, paradoxically, am frustrating my own scientific career by adopting practices that are better for science as a whole. However, it should be noted that an accepted registered (replication) report guarantees a publication, whereas null results from standard studies often remain unpublished. Considering the time investment of an RRR over a standard study, it is, unfortunately, still tempting to take quantity over quality and adopt best practices later, once I'm an established scientist.

To ensure that high-quality research projects, such as registered reports, will be conducted in the future, appreciation for these projects should be re-evaluated

until they replace the current status quo in psychological research. Large, robust studies require authors to go the extra mile to increase the quality of research in their field. For this reason, they should be weighted differently, doing justice to their scope. As I, and dozens of other researchers in similar situations, could have chosen quantity over quality and published more than one standard study rather than one registered (replication) report, hiring and grant committees should adjust their standards in evaluating academic CVs. Conducting high-quality research should not be a hindrance to one's career but a facilitator. A step in the right direction would be to replace the current focus on the length of academic resumes by selection, tenure and grant committees with focus on the quality, rigour and scope of individual publications. Special attention and rewards should be available to the researcher who aims to improve science, e.g., by valuing pre-registrations and open materials, and maybe temporarily considering registered (replication) reports as multiple publications (separately for the protocol and the final report). Luckily, in the Netherlands, some of this culture change is already present. My project was funded by the Netherlands Organisation for Scientific Research with a replication grant. This trend should continue and academics and funders should continue to recognize this 'new', robust way of conducting research. Only in this way can we ensure that publication formats that promote scientific rigour and credibility, such as registered (replication) reports, can continue to serve as backbone for robust research results. □

Andrea H. Stoevenbelt

Tilburg University, Tilburg, the Netherlands.
e-mail: a.h.stoevenbelt@tilburguniversity.edu

Published online: 10 October 2019
<https://doi.org/10.1038/s41562-019-0694-3>

Competing interests

The author declares no competing interests.