Unpublished results hide the decline effect

Some effects diminish when tests are repeated. **Jonathan Schooler** *says being open about findings that don't make the scientific record could reveal why.*

Many scientifically discovered effects published in the literature seem to diminish with time. Dubbed the decline effect, this puzzling anomaly was first discovered in the 1930s in research into parapsychology, in which the statistical significance of purported evidence for psychic ability declined as studies were repeated. It has since been reported in a string of fields — both in individual labs (including my own) and in meta-analyses of findings in biology and medicine. The issue has been recognized in some circles within the scientific community, but rose to wider prominence last December when it was discussed in an article in the magazine *The New Yorker*.

Some scientists attribute the decline effect to statistical self-correction of initially exaggerated outcomes, also known as regression to the mean. But we cannot be sure of this interpretation, or even test it, because we do not generally have access to 'negative results': experimental outcomes that were not noteworthy or consistent enough to pass peer review and be published.

How could the availability of unpublished results be improved? I suggest an open-access repository for all research findings, which would let scientists log their hypotheses and methodologies before an experiment, and their results afterwards, regardless of outcome. Such a database would reveal how published studies fit into the larger set of conducted studies, and would help to answer many questions about the decline effect.

Availability of unpublished findings could also address other shortcomings of the current scientific process, including the regular failure of scientists to report experiments, conditions or observations that are inconsistent with hypotheses; the addition or removal of participants and variables to generate statistical significance; and the probable existence of numerous published findings whose non-replicability is shrouded because it is difficult to report null results.

To address the decline effect, such a database could pinpoint whether the phenomenon reflects how scientists design experiments, how they write them up or how journals decide what to publish. It could be used to explore whether genuine changes in studied phenomena could stem from conventional mechanisms; for example, in social sciences, decline effects could be the result of participants no longer being naive about the effect under investigation. Less likely, but not inconceivable, is an effect stemming from some unconventional process. Perhaps, just as the act of observation has been suggested to affect quantum measurements, scientific observation could subtly change some scientific effects. Although the laws of reality are usually understood to be immutable, some physicists, including Paul Davies, director of the BEYOND: Center for Fundamental Concepts in Science at Arizona State University in Tempe, have observed that this should be considered an assumption, not a foregone conclusion.

➲ NATURE.COM
Discuss this article online at:
go.nature.com/xbvqlt

More prosaic explanations for the decline effect include the previously mentioned regression to the mean. If early results are most likely to be reported when errors combine to magnify the apparent effect, then published studies will show systematic bias towards initially exaggerated findings, which are subsequently statistically self-corrected (although this would not account for the typically linear nature of the decline).

Publication bias could also be responsible. Researchers might only be able to publish initial findings on an effect when it is especially large, whereas follow-up studies might be more able to report smaller effects. Other potential answers include unreported aspects of methods, exclusive reporting of findings consistent with hypotheses, changes in researcher enthusiasm, more rigorous methodologies used in later studies, measurement error resulting from experimenter bias and the general difficulty of publishing failures of replication.

WE NEED A BETTER RECORD TO LEARN HOW WELL SCIENCE DISTINGUISHES TRUTH FROM FALLACY.

An open-access database of research methods and published and unpublished findings would go a long way towards testing these ideas. For example, both the regression to the mean and degradation of procedure explanations assume that early published studies benefit from being at one statistical end of a larger body of (unpublished) findings. Publication bias and selective reporting of data are similarly difficult to investigate without knowing about unpublished data.

An open-access repository of findings would be difficult to introduce. It would need an automated protocol to enable study methods and results to be entered and retrieved. Some way to assess the quality of the work would be required — perhaps through open-access commentaries moderated in a manner similar to Wikipedia. We would need to assure the qualifications of researchers who use it, and maintain a blackout period to protect hypotheses and findings prior to publication. Reluctant scientists would need incentives — and perhaps new rules from funders — to take part.

Such challenges would not be insurmountable. Similar, if more narrowly defined, databases have already been set up for clinical trials (http://clinicaltrials.gov) and educational research (http://pslcdatashop.web.cmu.edu). A good starting point might be to develop a host of subject-specific repositories. However it is implemented, we need a better record of unpublished research before we can know how well the current scientific process, based on peer review and experimental replication, succeeds in distinguishing grounded truth from unwarranted fallacy. ■

Jonathan Schooler *is a professor of psychology at the University of California, Santa Barbara.*
e-mail: schooler@psych.ucsb.edu