

Cosmic Habituation or Statistical Variation?

Rob Gandy and Denis Adams

William Ashton gave an excellent description of how the replication of scientific studies is seeing a fading away of original results, described as the *decline effect*, and considers Jonathan Schooler's suggestion that the cause might be *cosmic habituation*, i.e. after repeated replications of a study, *the Universe is getting tired of responding* [FT299:52-53]. To study the decline effect Schooler suggests an open and online registry for all research studies, which is actively being pursued by some psychologists with the Open Science Collaboration (OSC)¹, but the fear is that this might inadvertently lead to psychology being stigmatised.

As long-in-the-tooth statisticians we thought that we should add our tenpenn'uth, given that invariably the results of scientific experiments are measured statistically. We cannot provide conclusive answers, but we want to contribute to the debate.

Misapplication of statistics

Because a paper is published does not necessarily mean the statistics are valid, even with a statistician involved! Repeating studies may not yield significant results – because significance did not exist in the first place. Porter² screened all papers and letters in the weekly British Medical Journal, Lancet and the New England Journal of Medicine in 1997 for examples where correlation and bi-variate linear regression were used. This yielded 15 categories of errors – 8 important or common. These included failure to clearly state the number of cases involved, not quoting confidence intervals when appropriate, and attaching undue importance to a significant outcome in the context of correlation². Even where a study is replicated is there certainty that identical statistical analyses are used?

Biases in psychological research

Analyses of which people were studied in six sub-disciplines of psychology, in top behavioural science research journals from 2003 to 2007³, revealed that 68 per cent of subjects came from the United States, with 96 per cent of subjects from Western industrialized countries (specifically North America, Europe, Australia and Israel)⁴. Also, 73 per cent of first authors were at American universities, and 99 per cent were at universities in Western countries. Therefore 96 per cent of psychological samples came from countries with only 12 per cent of the world's population. Fortean will love the inference that most published research in top behavioural science journals is based on samples drawn entirely from Western, Educated, Industrialized, Rich, and Democratic (WEIRD) societies³.

Furthermore, in the Journal of Personality and Social Psychology, the premier journal in social psychology, 67 per cent of American samples (and 80 per cent of samples from other countries) were composed solely of undergraduates in psychology courses⁴. Hence, a randomly selected American undergraduate was over 4,000 times more likely to be a research participant than a randomly selected person from outside the West.

The dominance of American authors in psychology publications may simply reflect American universities being able to attract the best international researchers, and that similar tendencies exist in other fields. However, psychology is a distinct outlier: 70 per cent of all psychology citations come from the United States – a larger percentage than any of the other 19 sciences⁵.

Consequently, psychology research has been largely studying the nature of WEIRD people; a narrow and potentially peculiar subpopulation, rather than the full breadth of human diversity. Unfortunately, despite their narrow samples, behavioural scientists often are interested in drawing inferences about the human mind and human behaviour, something that is rarely challenged or defended³. Leading scientific journals and university textbooks routinely publish research findings claiming to generalize to “humans” or “people” based on research done entirely with WEIRD undergraduates. Commonly, there is no demographic information about the participants, aside from their age and gender³.

A genuine danger that needs to be addressed is that if study samples continue to be dominated by undergraduates in psychology courses⁴ then replicated studies might be affected by some of those participating having (some) prior knowledge of the published studies and results through their research and teaching.

Logistical pressures

It is inferred that William Ashton refers to the replication of published experiments where the findings were statistically significant, because if the findings were *not* statistically significant the authors would be unlikely to write up the research and submit for publication, the journals would be unlikely to publish and no-one would consider replicating the experiment. Potentially, researchers could have undertaken similar (if not identical) experiments before, but without significant findings, and without seeking publication. A published paper might only reflect the first time an experiment has produced statistically significant findings.

If a statistical analysis is “significant with 95 per cent confidence” (the usual yardstick adopted) then this means the probability of those statistical results occurring is less than 0.05; in broad terms, if an experiment is undertaken 20 times then such an outcome is likely to happen once. What is usually inferred by researchers is that if they get statistically significant results in one experiment then they should always get statistically significant results if the experiment is repeated – this is not necessarily a valid assumption.

Conclusions

Published work is biased towards statistically significant findings (where relevant) but it might only reflect that proportion of experimental research which would be expected to yield statistically significant findings in line with probability for a given experiment. The OSC¹ will be invaluable by providing data on all experiments: but its data might not help measure a *decline effect* but evidence the ranges of results that would be expected in line with the laws of probability. Less *cosmic habituation*, more *statistical variation*.

References

- 1 S Carpenter “Psychology’s bold initiative”, *Science*, 335, 2012, pp1558-1561
- 2 A Porter “Misuse of correlation and regression in three medical journals”, *Journal of the Royal Society of Medicine*, 92, 1999, pp 123-128
- 3 J Henrich, S Heine and A Norenzayan “The weirdest people in the world?”, *Behavioural and Brain Sciences*, 33, 2010, pp 61-135
- 4 J Arnett “The neglected 95%: Why American psychology needs to become less American” *American Psychologist*, 63(7), 2008, pp 602–614
- 5 R May “The scientific wealth of nations”, *Science*, 275, 1997, pp 793–796

899 words

Pen Pics

ROB GANDY is a visiting professor at the Liverpool Business School, in Liverpool John Moores University. He has written previously for *FT* on an eclectic range of fortean topics including Merseyside doppelgangers and ghostlore.

DENIS ADAMS is a retired business analyst from the Liverpool Business School, and has long witnessed the *decline effect* in his special interest of horse racing!