

Crisis in What Exactly?

Olivia Guest¹

¹University of Oxford

April 17, 2023



Crisis in What Exactly?

OLIVIA GUEST¹

1. University of Oxford

READ REVIEWS

WRITE A REVIEW

CORRESPONDENCE:

olivia.guest@psy.ox.ac.uk

DATE RECEIVED:

June 14, 2016

DOI:

10.15200/winn.146590.01538

ARCHIVED:

June 14, 2016

KEYWORDS:

#LJAFreproducibility, scholarly communication, reproducibility, open science

CITATION:

Olivia Guest, Crisis in What Exactly?, *The Winnower* 3:e146590.01538, 2016, DOI: 10.15200/winn.146590.01538

© Guest 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.



THE SAME THING OVER AND OVER AGAIN

When scientists fail to reproduce the results of multiple experiments the support previously provided for theories breaks down. This is a modern phenomenon in science dubbed by some a “replication crisis” (Earp and Trafimow, 2015; Ioannidis, 2005; Maxwell, Lau, and Howard, 2015; Spellman, 2015; Stroebe and Strack, 2014). Few fields have been left unscathed, from medicine (Freedman, 2015) to computational science (Peng, 2011), and from psychology (Maxwell, Lau, and Howard, 2015) to political science. Replication crisis-like events, however, have been documented throughout history known as scientific revolutions or paradigm shifts. While they may appear similar, some important contrasts can be drawn out between replication crises and paradigm shifts.

Paradigm shifts occur when a mainstream theory ceases to predict or explain the data as well as originally thought (Kuhn, 1970). The data is instead accounted for by a marginal theory, which comes to be accepted as the new mainstream account. Replication crises are due to methodological failures, unlike scientific revolutions which are triggered by theoretical failures. The data is brought under question due to failure to reproduce the same results, which in turn causes the leading theories to be brought under scrutiny. In the case of replication crises, the presence of a new theoretical account to take over after the loss of evidence for the mainstream view is not guaranteed.

Is science now “progressing” via replication crises? If so, why and what does this mean for the future of the affected fields? In this article, I plan to address these issues by expounding further on the contrasts between paradigm shifts and replication crises and on the changes that have taken place within science as a community and industry. Can the current bout of replication crises culminate in a revolution within the cultural values scientists adhere to? Are fundamental changes required in the normative practice of science or in the descriptive attitudes of scientists toward their work or both?

I propose that replication crises are not the product of ineffective data gathering and insufficient dissemination of experimental paradigms. Instead, these are merely the side-effects caused by problems at the true core of the scientific enterprise: that science does not always provide environments appropriate for researchers to behave properly.

BURNING AND DECAPITATION



The Lavoisiers (Marie-Anne and Antoine) contemplating revolution. This science power-couple revolutionised their field — facilitating the transition from alchemy to chemistry. Together they showed that phlogiston, the mainstream view, did not actually predict the data at all. Instead, they demonstrated that oxygen, a substance they discovered, takes part in combustion.

Phlogiston theory was a 17th-18th century attempt to explain combustion, rusting and related phenomena — all of which are now described by the theory of oxidation. It was proposed that a material known as phlogiston existed in all combustible materials which when burned released it and hence lost weight. We now know, as they also discovered at the time, that not all materials lose weight when burned. Experiments, which showed these exceptional materials gaining weight after combustion, undermined the theory. As a result of this and other issues, phlogiston theory was slowly losing credibility. For phlogiston what is coming is clear, especially to those familiar with this story: its demise — attributed to a French nobleman named Lavoisier. The Lavoisiers, see , demonstrated that oxidation is due to oxygen, which reacts with the fuel to release energy. Phlogiston does not exist. The oxygen theory provides a more parsimonious account.

What can be learned from the events that took place during the chemical, and just before the French, revolutions? And what can be gleaned from other paradigm shifts and the history of science more generally? Firstly, most science, especially in the past and in fields with lower noise-to-signal ratios (e.g., physics, chemistry), progressed and continues to progress by the proposal of new theories. This is not the case when we examine replication crises. The paradigm shift that occurred due to the discovery of oxygen did not bring into question the phenomena of breathing or rusting (Trafimow and Earp, in press). It was in fact the trust in the data itself that sealed the fate of phlogiston: materials were measured that gained weight when burned. In the chemical and other scientific revolutions data was dependable.

Secondly, the Lavoisiers were French aristocrats: most highly prestigious science was carried out by privileged people who chose science because they loved it. Their background also protected them for undue stress due to, e.g., job loss — decapitation being a notable exception. In contrast, some of today's scientists, like people in the present more generally, seek to attain status through their careers, rather than by birth, often by means not compatible with the aims of science. Arguably, Lavoisier and contemporaries, from lowly research assistants to prominent figures, had better job security than many of today's PhD students and postdoctoral researchers (Garrison, Stith, and Gerbi, 2005; Powell, 2015). Prestige- and recognition-seeking while perhaps not inherently misplaced in and of themselves, can run counter to science's aims of truth-seeking. Most researchers are primarily intrinsically motivated, but there are enough exceptions to cause problems.

SCIENCE HACKING



Fabiola Gianotti and Peter Higgs embracing after she presented evidence for the existence of the Higgs boson (Credit for photograph: Denis Balibouse, Associated Press). Higgs was not always supported throughout his career by institutions and peers. Yet he believes he would have an even harder time in the present: “Today I wouldn’t get an academic job. It’s as simple as that. I don’t think I would be regarded as productive enough.”(Aitkenhead, 2013)

P-hacking is the practise of getting the results required regardless of the data by means of academic misconduct and rigged statistical testing. The mildest forms of p-hacking, from the author's personal experiences and others', are extremely widespread (Head et al., 2015; Lakens, 2015; Leggett et al., 2013; Masicampo and Lalande, 2012). This brings us to the crux of the present problems. Many scientists — enough to cause replication crises! — do not behave in a way that ensures their experiments can be reproduced (Ioannidis, 2005). The collection and dissemination of data and

experimental paradigms often is done under duress. The causes of such limitations in scientific rigour are manifold: too much pressure (see Figure 2); lack of guidance, training, and motivation (Borlee, 2011; Taylor, 2011; Woolston, 2015); and little job security (Alvarez, 2007).

Undoubtedly, all scientists start off extremely passionate, but some due to toxic environments spiral into depression (Gewin, 2012), while others become corrupted by a suboptimal system (McNutt, 2015). This has serious repercussions; the success of the whole enterprise depends on passion (Vallerand, 2012). In the same way that a hospital cannot prioritise saving money over saving patients without being branded a failure, science that gives higher priority to things other than producing and testing theories, is doomed to fall into further crises.

Nonetheless, progress cannot purely be made by replicating previous research. Scientific progress requires new theories which fuel revolutions in understanding. However, this idealised form of progress cannot be achieved without improving the foundations: experimental design and data gathering and analysis. So while progress cannot be made purely on the back of replication, it is a necessary part of rectifying: the problems we inherit from other researchers; the issues inherent in disciplines with high noise-to-signal ratios — false positives are often inevitable (Nuzzo and others, 2014); and the biases in publication and the file drawer effect (Simonsohn, Nelson, and Simmons, 2014).

Can a passion for science be maintained in PhD students and beyond? If not, no amount of data sharing (since if it is bad data it needs a passionate person to re-analyse and debunk, motivated by uncovering the truth) or prize-giving (working towards a prize like working towards an exam is not the same as doing good science) will bring better theories and better practice to the fore (Stoeber et al., 2011). Top down regulation cannot always drive systemic changes, especially if those working at the bottom are cunning enough to bypass controls, which depend on mutual trust U. (Frith and Frith, 2014).

“Questionable research practices” (Spellman, 2015; Stroebe and Strack, 2014) include: tricking or bypassing journal peer review systems, such as providing fake email addresses for reviewers or publishing a book instead of an article; and serious academic misconduct and outright fraud (McNutt, 2015; Yong, 2012), such as deleting participants before making data public. In the same way that the use of closed-circuit television systems does not reduce crime levels to zero, the use of fully open science cannot deter a motivated cheat from manipulating their data before making it accessible (Gigerenzer and Marewski, 2015). Such cases of misconduct or negligence sap other researchers’ time and waste everybody’s money (Freedman, 2015). Something needs to change.

REVOLUTION AND PROGRESS



The above diagram shows a distinction between preventative and curative ways of treating replication crises reproduced from figure 1 in: Leek and Peng 2015). The authors propose that “the replication crisis needs to be considered from the perspective of primary prevention” (pp.

1645-6 Leek and Peng, 2015); drawing attention to improved science education as a long-term solution.

We must create an atmosphere where young researchers flourish as opposed to lose their spark, where passionate people care about generating and testing theories and do not worry disproportionately about impact factors or losing their jobs. Stigma needs to be removed from carrying out replications, publishing null results, and speaking out against academic misconduct. Instead we must actively discourage those who engage in misconduct. We must dissuade systems and people who care about quantity and prestige over quality (Fanelli, 2010; Neill, 2008; Reich and others, 2013; Schekman, 2013). Replication crises indicate that we need to provide scientists with an environment in which they can carry out their job to the best of their abilities and uncorrupted by ulterior motives. We need a culture that supports scientists to do good science. This culture is created by us, the individual scientists within it. We can choose collectively and individually which practices we wish to shame and which to praise, and which principles we want our contemporaries and successors to uphold.

If we want revolution, we will have it (Spellman, 2015). The machinery is in place for improvement, including: better pay (Kaiser, 2016; Smaglik, 2016); preventative measures as well as reactive “crises” (e.g., as shown in Figure 4; Leek & Peng, 2015); proposals for improving training, providing realistic career expectations for students and postdoctoral researchers, and severe criticisms of the current system in which PhD students are treated like cheap disposable labour (Borlee, 2011; Seeliger, 2012; Taylor, 2011; Woolston, 2015); a debate with respect to the retirement of baby-boomers (Scudellari, 2015); the open science movement (Barnes, 2010; R. D. Morey et al., 2016); calls for more clarity in journal articles (Casadevall and Fang, 2010; Cooper and Guest, 2014; Mesirov, 2010); understanding the need for better theories (Klein, 2014); the shaming of flagrant p-hackers, scammers, and bullies (Bohannon, 2014; McNutt, 2015); acknowledging the value of conceptual replications (Crandall and Sherman, 2016); and more. None of these can single-handedly change the institutional values of science, in the same way that a single individual cannot. But if we tackle the issues the replication crises have uncovered, from the top downwards and the grass-roots upwards, positive change is inevitable. Ultimately: only more honest people create more honest science — and only an honest culture breeds more honest people.

REFERENCES

Alvarez, Michael. 2007. “A Question of Supply and Demand.” *Nature* 124 (445): 124.
[doi:10.1038/nj7123-124a](https://doi.org/10.1038/nj7123-124a).

Barnes, Nick. 2010. “Publish Your Computer Code: It Is Good Enough.” *Nature* 467 (7317). Nature Publishing Group: 753–53. [doi:10.1038/467753a](https://doi.org/10.1038/467753a).

Bohannon, John. 2014. “Replication Effort Provokes Praise—and ‘bullying’ charges.” *Science* 344 (6186). American Association for the Advancement of Science: 788–89.
[doi:10.1126/science.344.6186.788](https://doi.org/10.1126/science.344.6186.788).

Borlee, Grace. 2011. “Where Do All the Postdocs Go?” *DNA and Cell Biology* 30 (8): 537.
[doi:10.1089/dna.2011.2506](https://doi.org/10.1089/dna.2011.2506).

Casadevall, Arturo, and Ferric C Fang. 2010. “Reproducible Science.” *Infection and Immunity* 78 (12). Am Soc Microbiol: 4972–75. [doi:10.1128/IAI.00908-10](https://doi.org/10.1128/IAI.00908-10).

Cooper, Richard P, and Olivia Guest. 2014. "Implementations Are Not Specifications: Specification, Replication and Experimentation in Computational Cognitive Modeling." *Cognitive Systems Research* 27. Elsevier: 42–49. doi:[10.1016/j.cogsys.2013.05.001](https://doi.org/10.1016/j.cogsys.2013.05.001).

Crandall, Christian S., and Jeffrey W. Sherman. 2016. "On the Scientific Superiority of Conceptual Replications for Scientific Progress." *Journal of Experimental Social Psychology*. doi:[10.1016/j.jesp.2015.10.002](https://doi.org/10.1016/j.jesp.2015.10.002).

Earp, Brian D., and David Trafimow. 2015. "Replication, Falsification, and the Crisis of Confidence in Social Psychology." *Frontiers in Psychology* 6 (621). doi:[10.3389/fpsyg.2015.00621](https://doi.org/10.3389/fpsyg.2015.00621).

Fanelli, Daniele. 2010. "Do Pressures to Publish Increase Scientists' Bias? An Empirical Support from US States Data." *PloS One* 5 (4). Public Library of Science: e10271. doi:[10.1371/journal.pone.0010271](https://doi.org/10.1371/journal.pone.0010271).

Freedman, Iain M. AND Simcoe, Leonard P. AND Cockburn. 2015. "The Economics of Reproducibility in Preclinical Research." *PLoS Biol* 13 (6). Public Library of Science: 1–9. doi:[10.1371/journal.pbio.1002165](https://doi.org/10.1371/journal.pbio.1002165).

Frith, Uta, and Chris Frith. 2014. "A Question of Supply and Demand." *The Guardian*.

Garrison, Howard H., Andrea L. Stith, and Susan A. Gerbi. 2005. "Foreign Postdocs: The Changing Face of Biomedical Science in the U.S." *The FASEB Journal* 19 (14): 1938–42. doi:[10.1096/fj.05-1203ufm](https://doi.org/10.1096/fj.05-1203ufm).

Gewin, Virginia. 2012. "Mental health: Under a cloud." *Nature* 490 (7419). Nature Publishing Group: 299–301. doi:[10.1038/nj7419-299a](https://doi.org/10.1038/nj7419-299a).

Gigerenzer, Gerd, and Julian N Marewski. 2015. "Surrogate Science the Idol of a Universal Method for Scientific Inference." *Journal of Management* 41 (2). SAGE Publications: 421–40. doi:[10.1177/0149206314547522](https://doi.org/10.1177/0149206314547522).

Head, Megan L, Luke Holman, Rob Lanfear, Andrew T Kahn, and Michael D Jennions. 2015. "The Extent and Consequences of P-Hacking in Science." *PLoS Biol* 13 (3). Public Library of Science: e1002106. doi:[10.1371/journal.pbio.1002106](https://doi.org/10.1371/journal.pbio.1002106).

Ioannidis, John P. A. 2005. "Why Most Published Research Findings Are False." *PLoS Med* 2 (8).

Public Library of Science. doi:10.1371/journal.pmed.0020124.

Kaiser, Jocelyn. 2016. "New U.S. overtime rules will bump up postdoc pay, but could hurt research budgets." *Science*. doi:10.1126/science.aaf5735.

Klein, Stanley B. 2014. "What Can Recent Replication Failures Tell Us About the Theoretical Commitments of Psychology?" *Theory & Psychology*. SAGE Publications, 0959354314529616. doi:0.1177/0959354314529616.

Kuhn, Thomas S. 1970. *The Structure of Scientific Revolutions*. University of Chicago Press.

Lakens, Daniël. 2015. "What P-Hacking Really Looks Like: A Comment on Masicampo and LaLonde (2012)." *The Quarterly Journal of Experimental Psychology* 68 (4). Taylor & Francis: 829–32. doi: 10.1080/17470218.2014.982664.

Leek, J. T., & Peng, R. D. (2015). Opinion: Reproducible research can still be wrong: Adopting a prevention approach. *Proceedings of the National Academy of Sciences*, 112 (6), 1645–1646. doi: 10.1073/pnas.1421412111

Leggett, Nathan C, Nicole A Thomas, Tobias Loetscher, and Michael ER Nicholls. 2013. "The Life of P: 'Just Significant' Results Are on the Rise." *The Quarterly Journal of Experimental Psychology* 66 (12). Taylor & Francis: 2303–9. doi: 10.1080/17470218.2013.863371.

Masicampo, EJ, and Daniel R Lalonde. 2012. "A Peculiar Prevalence of P Values Just Below .05." *The Quarterly Journal of Experimental Psychology* 65 (11). Taylor & Francis: 2271–79. doi: 10.1080/17470218.2012.711335.

Maxwell, Scott E, Michael Y Lau, and George S Howard. 2015. "Is psychology suffering from a replication crisis? What does 'failure to replicate' really mean?" *American Psychologist* 70 (6). American Psychological Association: 487. doi: 10.1037/a0039400.

McNutt, Marcia. 2015. "Editorial Retraction." *Science* 348: 1100–1100. doi: 10.1126/science.351.6273.569-a.

Mesirov, Jill P. 2010. "Accessible Reproducible Research." *Science* 327 (5964). American Association for the Advancement of Science: 415–16. doi: 10.1126/science.1179653.

Morey, Richard D., Christopher D. Chambers, Peter J. Etchells, Christine R. Harris, Rink Hoekstra,

Daniël Lakens, Stephan Lewandowsky, et al. 2016. "The Peer Reviewers Openness Initiative: Incentivizing Open Research Practices Through Peer Review." *Open Science* 3 (1). The Royal Society. doi: [10.1098/rsos.150547](https://doi.org/10.1098/rsos.150547).

Neill, Ushma S. 2008. "Publish or Perish, but at What Cost?" *The Journal of Clinical Investigation* 118 (7). Am Soc Clin Investig: 2368–68. doi: [10.1172/JCI36371](https://doi.org/10.1172/JCI36371).

Nuzzo, Regina, and others. 2014. "Statistical Errors." *Nature* 506 (7487). Macmillan Publishers Ltd., London, England: 150–52. doi: [10.1038/506150a](https://doi.org/10.1038/506150a).

Peng, Roger D. 2011. "Reproducible Research in Computational Science." *Science* (New York, Ny) 334 (6060). NIH Public Access: 1226. doi: [10.1126/science.1213847](https://doi.org/10.1126/science.1213847).

Powell, Kendall. 2015. "The Future of the Postdoc." *Nature* 520 (7546): 144–47. doi: [10.1038/520144a](https://doi.org/10.1038/520144a).

Reich, Eugenie Samuel, and others. 2013. "The Golden Club." *Nature* 502 (7471). Macmillan Publishers Ltd., London, England: 291–93. doi: [10.1038/502291a](https://doi.org/10.1038/502291a).

Schekman, Randy. 2013. "How Journals Like Nature, Cell and Science Are Damaging Science." *The Guardian*.

Scudellari, Megan. 2015. "The Retirement Debate: Stay at the Bench, or Make Way for the Next Generation." *Nature* 521: 20–23. doi: [10.1038/521020a](https://doi.org/10.1038/521020a).

Seeliger, Jessica C. 2012. "Scientists Must Be Taught to Manage." *Nature* 483 (7391): 511.

Simonsohn, Uri, Leif D Nelson, and Joseph P Simmons. 2014. "P-Curve: A Key to the File-Drawer." *Journal of Experimental Psychology: General* 143 (2). American Psychological Association: 534. doi: [10.1037/a0033242](https://doi.org/10.1037/a0033242).

Smaglik, Paul. 2016. "Activism: Frustrated Postdocs Rise up." *Nature* 530 (7591). Nature Publishing Group: 505–6. doi: [10.1038/nj7591-505a](https://doi.org/10.1038/nj7591-505a).

Spellman, Barbara A. 2015. "A Short (Personal) Future History of Revolution 2.0." *Perspectives on Psychological Science* 10 (6). SAGE Publications: 886–99. doi: [10.1177/1745691615609918](https://doi.org/10.1177/1745691615609918).

Stoeber, Joachim, Julian H Childs, Jennifer A Hayward, and Alexandra R Feast. 2011. "Passion and Motivation for Studying: Predicting Academic Engagement and Burnout in University Students." *Educational Psychology* 31 (4). Taylor & Francis: 513–28. doi: [10.1080/01443410.2011.570251](https://doi.org/10.1080/01443410.2011.570251).

Stroebe, Wolfgang, and Fritz Strack. 2014. "The Alleged Crisis and the Illusion of Exact Replication." *Perspectives on Psychological Science* 9 (1): 59–71. doi: [10.1177/1745691613514450](https://doi.org/10.1177/1745691613514450).

Taylor, Mark. 2011. "Reform the PhD System or Close It down." *Nature* 472 (7343): 261–61. doi: [10.1038/472261a](https://doi.org/10.1038/472261a).

Trafimow, David, and Brian D. Earp. in press. "Badly Specified Theories Are Not Responsible for the Replication Crisis in Social Psychology: Comment on Klein (2014)." *Theory & Psychology*. doi: [10.1177/0959354316637136](https://doi.org/10.1177/0959354316637136).

Vallerand, Robert J. 2012. "From Motivation to Passion: In Search of the Motivational Processes Involved in a Meaningful Life." *Canadian Psychology/Psychologie Canadienne* 53 (1). Educational Publishing Foundation: 42. doi: [10.1037/a0026377](https://doi.org/10.1037/a0026377).

Woolston, Chris. 2015. "Graduate Survey: Uncertain Futures." *Nature* 526 (7574). Nature Publishing Group: 597–600. doi: [10.1038/nj7574-597a](https://doi.org/10.1038/nj7574-597a).

Yong, Ed. 2012. "In the Wake of High Profile Controversies, Psychologists Are Facing up to Problems with Replication." *Nature* 483: 298–300. doi: [10.1038/485298a](https://doi.org/10.1038/485298a).