Science for All

Liam J. Messin[1], John C. Meadows[1,2,*]

[1]Division of Biomedical Sciences, Warwick Medical School, University of Warwick, Gibbet Hill Road, Coventry, CV4 7AL, UK

[2]Institute of Advanced Study, University of Warwick, Gibbet Hill Road, Coventry, CV4 7AL, UK

*Correspondence: j.c.meadows@warwick.ac.uk

Abstract *Science is the best method humanity has for understanding the universe and our place in it. However, despite its astonishing achievements, the implementation and communication of science is not without problem. In this article the development of scientific publishing is discussed, as is its subsequent impact on scientific discovery and on the profession itself. The ways in which a move towards a more open science framework might alter how science is reported and ultimately performed is also addressed.*

Keywords: Open Science, Science Publishing, Open Access

Peer review: This article has been subject to a double blind peer review process

Introduction

The scientific method, when performed correctly and given enough time, will eventually uncover all that can ever be known. Through its application humanity has discovered more about itself and the universe than ever before and the impact of these findings on society are immeasurable. However, all is not perfect with science — there are entrenched problems both in how it is communicated and how it is performed. To more fully understand why contemporary science faces these challenges and how these can be addressed it is instructive to look at the development of science from an historical perspective.

The rise and rise of the scientific journal

The earliest written dissemination of scientific information, excluding the publication of academic books, was between scientific peers via letters outlining their discoveries. Data were frequently not included overtly in these missives and were instead encoded as anagrams that could be unscrambled retroactively if discoveries were contested. The first scientific journal, *Philosophical Transactions of the Royal Society of*

London, published continuously since 1665, was launched with a mandate to allow scientists '*to search, try, and find out new things, impart their knowledge to one another, and contribute what they can to the Grand design of improving Natural knowledge*' (**Oldenburg, 1665**). It was, like many early journals, a collection of the type of transcribed letters that scientists were previously sending only to each other and whilst this broadened the audience somewhat it was still restricted to members of learned bodies.

During the 1850s more and more journals were established (**Barton, 1998**), coincident with rampant industrialisation and a Victorian fascination for discovery. In parallel there was a move away from epistolary dissemination toward a format that would be recognised today, whereby scientists generate data to test a hypothesis, finally pulling all the relevant data together and forming a narrative which reads abstract, methodology, results and conclusions. This is then submitted to a journal for evaluation by scientific peers and/or editors and following potential revisions, published or not. This system allows for anonymous feedback from experts/competitors only at the very end of the process and has remained largely unchanged to the present day. The immense impact of peer-reviewed high-level scientific publishing cannot be overstated in its facilitation of scientific discovery.

Also in the late nineteenth century the demand for popular science writing for non-expert audiences increased. This was met with the publication of anthologies such as *Science For All* (**Brown, 1877**), dealing with topics thought to be of interest to a lay audience and covering a wide spectrum of disciplines. This format of closed-access primary scientific publishing and popular science writing set the tone for science communication and dissemination until the advent of the Internet.

Science commoditised

Journal-based science communication has given a distinct flavour not only to how science is presented but also fundamentally to how it is conducted and financed. Scientific journals are in competition with each other to publish the best research they can and to increase their readership. Indeed, journals are assigned impact factors based on the average number of citations papers that they publish receive in a given period. These impact factors are widely available, in some cases actively promoted, and stand as a shortcut for the perceived quality of the journal. That they say little about the quality of individual articles and do not reflect whether a citation is in agreement or opposition to the original article are only some of the limitations to this metric. In the most part this quantification means that journals favour publications that will garner more citations. At first glance this appears reasonable, as good science should be referenced frequently, however, this has unintended ramifications for both the science that they publish and for scientific discovery in general.

Journals have preferentially published positive data stories over negative ones (**Matosin et al., 2014**). This means a paper reporting data on, for example, an anti-cancer agent is rightly published, however, it is much harder, if not impossible, to publish a story showing that a potential drug has no effect. This bias is, in part, due to negative data receiving fewer citations than positive data (**Fanelli, 2010**). Initially this seems perfectly intuitive, however, these negative data are frequently lost to the scientific record, whereas if available they might have informed others' experiments and/or methodologies. It also prevents verification of negative data between groups if no one is aware of its existence. Both negative and positive data are generated using the same expertise, technology and time-commitment from researchers, however, positive data are preferentially rewarded with publications. This leads to the so-called file drawer problem, in which negative data are discarded prior to publication with scientists instead beginning new experiments (**Scargle, 1999**). Whilst negative data continue to be undervalued and underrepresented only a small fraction of the total work conducted is ever visible to all the stakeholders. This problem, referred to as dark data, is one that requires innovative solutions — journals are not vast archival repositories and nor should they necessarily be.

A related issue is that of Ockham's broom — removal of data that do not exactly fit your hypothesis so as to make a complex story appear simpler (**Robertson, 2009**). This cherry picking is contrary to the testing of a scientific hypothesis to destruction and is anti-transparency at best and fraudulent at worst. When scientists are driven by the narrative and space constraints of publishing in a journal they may potentially choose only the data that fit. This is termed confirmation bias — finding supporting evidence for a hypothesis instead of conducting experimentation that could disprove your theory. Similarly, large datasets can be mined using only those criteria that give statistically significant outputs, in a process termed p-hacking, with the other criteria remaining unreported. This is especially problematic if researchers measure a wide range of outcomes until they find one that produces positive data rather than setting out to measure a specific set of metrics. The more outcomes a scientist measures the higher the probability that any difference measured is the result of stochasticity. Ideally measurement of positive data generated in this fashion should be repeated, but this is not always the case. Removal of data, either negative or complex, means that it is excluded from the peer-review process, the scientific gold standard to ensuring accuracy and promoting standards.

Lastly, novelty is a highly prized attribute for publication in higher impact journals (**Kravitz and Baker, 2011**), meaning that fewer findings are independently verified, as there is scant reward for doing so. This has led to what is termed the reproducibility crisis, affecting both basic and clinical research — a 2012 review found that only six out of 53 clinical oncology papers were reproducible (**Begley and Ellis, 2012**). Unfortunately it is not the case that all science is done so well, and published methodologies so thorough, that it is reproducible by anyone with the competency to do so. The importance of novelty, being the first to publish something, also leads to the phenomenon of 'scooping' whereby if work that broadly agrees with your own is published prior to your own the perceived worth of your findings are diminished leading to publication in a journal with a lower impact factor if at all - again reducing verification rates. The novelty of a result does not intrinsically mean that it is important or will better inform either scientific understanding or clinical relevance. It also relies on scientists having an unlearnable intuition as to what experiments will produce novelty as the final outcome. Scientists may also come under pressure to shift focus into research fields that are currently fashionable in the hopes of publication, regardless of how important that trend may eventually be (**Horton, 2015**).

Publish and/or perish

The aforementioned issues matter for reasons of scientific transparency and integrity, however, human beings conduct science and it is necessary to look at the pressures that the current system imposes on them to understand how these issues might be circumvented. Science funding generally requires grants from state-funded or charitable bodies and securing these is extremely competitive. As journal articles are a scientist's main output it is understandable that amongst the criteria that determine success is a demonstrable track record of publication in high impact journals (**Schekman, 2013**). Ensuring continued success in an ever more competitive landscape means not just publishing but also publishing in the 'right' journals. Funders and employers obviously require metrics to be able to distinguish the best-suited candidates, however, this does mean that any pre-existing issues with journal-published science become ever more intractable.

In addition to the issues of chasing novel, positive data to achieve credible metrics at the expense of verification there are some other issues affecting the type and rate of scientific discovery that are becoming ever more apparent. The current model favours risk-mitigation strategies leading to a trend towards conservatism in grant proposals, papers and science discovery. This conservatism is backed up by short-termism due to funding durations meaning larger scale expensive projects are disfavoured due to associated or perceived risk. A study of 6.4 million chemistry and biomedicine publications (published 1934–2008) found that answering established questions resulted in more publications and securer funding for scientists whereas riskier research was less likely to produce funding (**Foster et al., 2015**). If published, however, riskier projects are rewarded with more citations and major rewards such as Nobel Prizes — with the study's author concluding that innovative research is '*a gamble whose payoff, on average, does not justify the risk*' (**Hampton, 2015**).

However, perhaps the greatest risk to the incentivisation of scientists via these metrics is that it could lead to avoidable errors, either inadvertent or fraudulent, creeping into methodologies or the scientific record (**Nature Editorial, 2012**). The pressure applied by funding agencies and employers — judged on the number of high impact papers an author has published — with the admirable intention of improving quality could instead be leading to a decrease in the accuracy of scientific findings. Richard Horton, Editor-in-Chief of the *Lancet* recently wrote that '*…much of the scientific literature, perhaps half, may simply be untrue*' (**Horton, 2015**). Cases of papers being retracted for a myriad of reasons including plagiarism and deliberate falsification of data are documented on the *Retraction Watch* blog (http://retractionwatch.com). Several such retractions have been initiated by users of *PubPeer* (http://pubpeer.com), a web resource that allows scientists to anonymously comment on others' articles in an example of post-publication peer review that itself has been subject to criticism (**Blatt, 2015**). Science has operated primarily on the presumption of honesty, scientists peer-reviewing papers assume that the data presented are truthful, but as stakes continue to rise and funding becomes more competitive can this continue?

Open access publishing and beyond

Until recently the majority of scientific journals operated in similar ways. Institutions or individuals paid for subscriptions to individual publications or suites of titles from the same publisher and it was this, together with advertisements and page fees, that allowed the journals to be funded. Of course this meant that only those with the means to view the publications could do so — as taxpayers and charitable institutions fund the majority of scientific research this meant that large swathes of stakeholders were being excluded. This, and other drivers, led the push towards an open-access publication model (**Suber, 2012**).

Open access publication democratises the dissemination of scientific information as these journals can be read free of charge because their funding is independent of subscriptions. Instead the typical funding model is that contributing authors pay an upfront article processing charge (APC) to ensure that the journal is free at the point of delivery. A significant landmark in open access publishing occurred in 2006 when the Public Library of Science (PLoS) created their exclusively online journal *PLoS ONE* (http://journals.plos.org/plosone/s/journal-information). A key difference from most traditional journals is that its criteria for publishing are that the science has been conducted rigorously and that the findings are correct. Work is therefore peer-reviewed for accuracy and not importance or novelty. Once published, *PLoS* articles are open to post-publication peer-review where the scientific community can inform judgement on importance through online comments and annotations in addition to citations. Criticism of open access publication focuses on the need for an APC suggesting that only those with the means to do so can publish; however, PLoS will waive or lower fees for scientists without the means to pay them. Similarly funders and host institutions, recognising the importance of open access publishing, have made funds available exclusively for this purpose. The *PLoS ONE* journal has proven successful allowing PLoS to subsidise its other, more traditional, journals. More recently the Howard Hughes Medical Institute, Max Planck Society and the Wellcome Trust established the open access journal *eLife*. A large, and recently renewed (**Callaway, 2016**), endowment means that, at least initially, there is no APC for publishing in *eLife*. Set up as an alternative to the journals in the field with the highest impact factors, *eLife* differs as it uses working scientists in lieu of professional editors, has a more open style of peer-review with reviewers' comments and authors' responses published alongside articles, offers post peer-review comment, has taken steps to actively promote reviewing early career researchers' manuscripts, does not promote its impact factor (**Kaiser, 2015**) and crucially does not put its content behind a paywall. Importantly, *eLife* publishes a non-technical summary for the non-specialist thus facilitating a much broader scientific discussion between all stakeholders.

The introduction of open access publishing has irreversibly altered the scientific landscape. Funders, through grant requirements, and institutions, through the research excellence framework (REF), strongly support publishing in this format. Therefore formerly exclusively subscription-based journals have had to adopt embargo periods whereby articles can be deposited into repositories or viewed for free after a set period has elapsed since publication. Likewise traditional publishers have established open-access only journals into their suite of publications or, somewhat controversially, allowed authors to pay for specific articles to be made open access in subscription-based journals — so-called hybrid open access (**Pinfield et al., 2015**). Subscription-based publications have historically provided other important activities to promote science and scientific communities such as conference funding or travel bursaries only possible due to their subscription model and commitment to greater scientific communication. It will be interesting to observe how the publishing ecosystem develops and responds to the challenges inherent in supporting multiple formats each with their own benefits.

For some researchers however, open access publishing is just the start of the open science revolution (**Bartling and Friesike, 2014**), described as '… *the idea that scientific knowledge of all kinds should be openly shared as early as is practical in the discovery process*' (**Nielsen, 2011**). This promotes the idea that access should be possible at all points in the scientific process not just at the point of publication. Open methodologies would allow for feedback at the experimental design stage. They would also require the use of open source software wherever possible to democratise verification. In the current system publishing your experimental methodologies and proposed experiments could be detrimental, as it would inform competitors of your plans. Archiving of all data generated would abolish the bias toward novel, positive data in the scientific record. It would also make science more transparent if the raw data could be interrogated directly by independent practitioners prior to journal publication. Indeed, it could nucleate collaboration, perhaps even between seemingly disparate fields (**Nielsen, 2012**). The preprint movement uses electronic repositories, such as *ArXiv* (http://arxiv.org), to allow researchers to collaborate and critique their data and ideas prior to journal publication. Publishing articles in this changed environment, however, could prove problematic due to non-trivial copyright issues over ownership of the data between authors, host institutions, funders and journals and this would need to be overcome. However, accreditation of contributions from researchers and groups should become more transparent. Of course, wholesale archiving, and making this archive open, would require a massive investment to infrastructure, not only to archive data but just as importantly to curate them (**Attard et al., 2015; The Royal Society, 2012**). Guidelines would need to be considered and implemented to maintain consistency and integrity of the archive.

The role of journals in this ecosystem would likely evolve further with narratives written around the archived datasets. Indeed, large archives of data would undoubtedly be unwieldy to navigate with a low signal to noise ratio. Therefore new roles in quality control and gatekeeping of datasets may well appear with bespoke interfaces to data servers individually tailored to specific interests. Describing science for the non-specialist audience may also take on ever more importance. The rise of the citizen science movement tells us that there are large groups of people looking to engage with science. Likewise science 2.0 practitioners who seek to self-publish their research through blogs, wikis and by conducting open (lab) book science shows that part of the science community is ready to interact openly with immediacy as a priority. Of course moving forward the crucial role of peer-review must be maintained – increased transparency must not be allowed to promote inaccuracy. In the interests of space this article has focused on journal-based communication throughout, however, long form publishing in monographs is similarly subject to the same ongoing debate over open access (**HEFCE, 2015a**).

Towards open science

A move towards more inclusive, accountable and transparent science will require significant changes not only in the metrics used to determine success (**San Francisco DORA, 2012**; **HEFCE, 2015b**) but also in the dissemination and practice of science necessitating a radical overhaul of the current infrastructure. The recent 'Amsterdam Call for Action on Open Science' in the EU advocates full open access publishing and mandatory data sharing for publicly funded research (**The Netherlands EU Presidency, 2016**). Issues of data archiving, copyright, accreditation, separating quality science from incorrectly performed science, building interdisciplinary networks, while increasing transparency and reproducibility will require new roles to be developed and funded. The way science is currently undertaken is a result of each player acting in the most logical way for their own circumstances and if real change is to be affected all aspects of science culture must be dealt with holistically. Communication, as ever, will be critical to this transition. As society is the major funder and beneficiary of science, researchers must engage with and enthuse the public about science. This in itself will require greater transparency and will in turn drive policy discussions and decisions about the best way to present and promote research. Funders are already promoting public engagement activities and schools outreach. Finding greater provision for these activities in the busy daily lives of active researchers is just one of the challenges facing the community.

Organisations such as the Voice of Young Science (VOYS) actively encourage early career researchers to engage with the public from the very start of their careers. Changing public perception of science and scientists away from breakthroughs and boffins is key and it is the responsibility of all scientists to actively engage in this on-going discourse.

Science is incremental and progressive, it can be difficult to comprehend or staggeringly beautiful in its intuitiveness, it is, at its best, powerful because it is true. Science is far too precious to be left exclusively to the scientists - the way it is practiced and reported is directly relevant to all of society and it should be embraced by all of us.

References

Attard, J., F. Orlandi, S. Scerri and S. Auer (2015), 'A systematic review of open government data initiatives', *Government Information Quarterly*, 32, 399–418.

Bartling, S. and S. Friesike (eds) (2014), *Opening Science: The Evolving Guide on How the Internet is Changing Research, Collaboration and Scholarly Publishing*, Springer International Publishing.

Barton, R. (1998), 'Just before Nature: The purposes of science and the purposes of popularization in some English popular science journals of the 1860s', *Annals of Science*, 55 (1), 1–33.

Begley, C. G. and L. M. Ellis (2012), 'Drug development: Raise standards for preclinical research', *Nature*, 483, 531–533.

Blatt M. R. (2015), 'Vigilante Science', *Plant Physiology*, 169, 907–909.

Brown, R. (ed.) (1877), *Science For All*, London: Cassell, Petter, Galpin & Co.

Callaway, E. (2016), 'Open-access journal gets £25-million boost', *Nature*, 534, 14–15.

Fanelli, D. (2010), 'Do pressures to publish increase scientists' bias? An empirical support from US States Data', *PLoS ONE*, 5, e10271.

Foster, J. G., A. Rzhetsky and J. A. Evans (2015), 'Tradition and Innovation in Scientists' Research Strategies', *American Sociology Review*, 80, 875–908.

Hampton, P. (2015), *Pressure to 'publish or perish' may discourage innovative research, UCLA study suggests*, http://newsroom.ucla.edu/releases/pressure-to-publish-or-perish-may-discourage-innovative-research-ucla-study-suggests, accessed 1 August 2016.

HEFCE (2015a), *Monographs and Open Access,* http://www.hefce.ac.uk/media/hefce/content/pubs/indirreports/2015/Monographs,and,open,access/2014_monographs.pdf, accessed September 26.

HEFCE (2015b), *The Metric Tide: Report of the Independent Review of the Role of Metrics in Research Assessment and Management,* http://www.hefce.ac.uk/media/HEFCE,2014/Content/Pubs/Independentresearch/2015/The,Metric,Tide/2015_metric_tide.pdf, accessed September 26.

Horton, R. (2015), 'Offline: What is medicine's 5 sigma?', *Lancet*, 385, 9976, 1380.

Kaiser, J. (2015), *Q&A Outgoing HHMI chief reflects on leading $19 billion biomedical charity*, http://www.sciencemag.org/news/2015/08/qa-outgoing-hhmi-chief-reflects-leading-19-billion-biomedical-charity, accessed 1 August 2016.

Kravitz, D. J. and C. I. Baker (2011), 'Toward a New Model of Scientific Publishing: Discussion and a Proposal, *Frontiers Computational Neuroscience*, 5, 55.

Matosin, N., E. Frank, M. Engel, Sl. L. Lum and K. A. Newell (2014), 'Negativity towards negative results: a discussion of the disconnect between scientific worth and scientific culture', *Disease Models & Mechanisms*, 7, 171–173.

Nature Editorial (2012), 'Must try harder', *Nature*, 483, 509.

Nielsen, M. (2011), *Definitions of Open Science?*, http://lists.okfn.org/pipermail/open-science/2011-July/000907.html, accessed 6 October 2016.

Nielsen, M. (2012), *Reinventing Discovery: The New Era of Networked Science*, Princeton: Princeton University Press.

Oldenburg, H. (1665), 'The Introduction', *Philosophical Transactions of the Royal Society of London*, 1, 1–2.

Pinfield, S., J. Salter and P. A. Bath (2016), 'The "total cost of publication" in a hybrid open-access environment: Institutional approaches to funding journal article-processing charges in combination with subscriptions', *Journal of the Association for Information Science and Technology*, 67, 1751–1766.

Robertson, M. (2009), 'Ockham's broom: A new series', *Journal of Biology*, 8, 79.

San Francisco DORA (2012), *Declaration on Research Assessment,* http://www.ascb.org/dora, accessed 6 October 2016.

Scargle, J. D. (1999), Publication Bias (The "File-Drawer Problem") in Scientific Inference, arXiv:physics/9909033.

Schekman, R. (2013), 'How journals like Naure, Cell and Science are ruining science', *The Guardian*, 9 December 2013, http://www.theguardian.com/commentisfree/2013/dec/09/how-journals-nature-science-cell-damage-science, accessed 1 August 2016.

Suber, P. (2012), *Open Access*, Cambridge: The MIT Press, also available at http://mitpress.mit.edu/sites/default/files/9780262517638_Open_Access_PDF_Version.pdf, accessed 26 September 2016.

The Netherlands EU Presidency (2016), *Amsterdam Call for Action on Open Science*, http://english.eu2016.nl/binaries/eu2016-en/documents/reports/2016/04/04/amsterdam-call-for-action-on-open-science/amsterdam-call-for-action-on-open-science.pdf, accessed 1 August 2016.

The Royal Society (2012), *Science as an open enterprise*, http://royalsociety.org/~/media/Royal_Society_Content/policy/projects/sape/2012-06-20-SAOE.pdf, accessed 26 September 2016.

To cite this article:

Messin, L. J. and Meadows, J. C. (2016). Science for All. Exchanges: The Warwick Research Journal, 4(1), 147-157. Retrieved from: http://exchanges.warwick.ac.uk/index.php/exchanges/article/view/127