REVIEW

REVISED The academic, economic and societal impacts of Open Access: an evidence-based review [version 3; referees: 3 approved, 2 approved with reservations]

Jonathan P. Tennant[1], François Waldner[2], Damien C. Jacques[2], Paola Masuzzo[3,4], Lauren B. Collister[5], Chris. H. J. Hartgerink[6]

[1]Department of Earth Science and Engineering, Imperial College London, London, UK
[2]Earth and Life Institute, Université catholique de Louvain, Louvain-la-Neuve, Belgium
[3]Medical Biotechnology Center, VIB, Ghent, Belgium
[4]Department of Biochemistry, Ghent University, Ghent, Belgium
[5]University Library System, University of Pittsburgh, Pittsburgh, PA, USA
[6]Department of Methodology and Statistics, Tilburg University, Tilburg, Netherlands

v3 **First published:** 11 Apr 2016, **5**:632 (doi: 10.12688/f1000research.8460.1)
Second version: 09 Jun 2016, **5**:632 (doi: 10.12688/f1000research.8460.2)
Latest published: 21 Sep 2016, **5**:632 (doi: 10.12688/f1000research.8460.3)

Abstract

Ongoing debates surrounding Open Access to the scholarly literature are multifaceted and complicated by disparate and often polarised viewpoints from engaged stakeholders. At the current stage, Open Access has become such a global issue that it is critical for all involved in scholarly publishing, including policymakers, publishers, research funders, governments, learned societies, librarians, and academic communities, to be well-informed on the history, benefits, and pitfalls of Open Access. In spite of this, there is a general lack of consensus regarding the potential pros and cons of Open Access at multiple levels. This review aims to be a resource for current knowledge on the impacts of Open Access by synthesizing important research in three major areas: academic, economic and societal. While there is clearly much scope for additional research, several key trends are identified, including a broad citation advantage for researchers who publish openly, as well as additional benefits to the non-academic dissemination of their work. The economic impact of Open Access is less well-understood, although it is clear that access to the research literature is key for innovative enterprises, and a range of governmental and non-governmental services. Furthermore, Open Access has the potential to save both publishers and research funders considerable amounts of financial resources, and can provide some economic benefits to traditionally subscription-based journals. The societal impact of Open Access is strong, in particular for advancing citizen science initiatives, and leveling the playing field for researchers in developing countries. Open Access supersedes all potential alternative modes of access to the scholarly literature through enabling unrestricted re-use, and long-term stability independent of financial constraints of traditional publishers that impede knowledge sharing. However, Open Access has the potential to become unsustainable for research communities if high-cost options are allowed to continue to prevail in a widely unregulated scholarly publishing market. Open Access remains only one of the multiple challenges that the scholarly publishing system is currently facing. Yet, it

Open Peer Review

Referee Status: ✓ ? ? ✓ ✓

Invited Referees

provides one foundation for increasing engagement with researchers regarding ethical standards of publishing and the broader implications of 'Open Research'.

This article is included in the The Future of Scholarly Publishing channel.

1 **Gwilym Lockwood**, Max Planck Institute for Psycholinguistics Netherlands

2 **Peter Suber**, Harvard University USA

3 **Paige Brown Jarreau**, Louisiana State University USA

4 **Anne Tierney**, Edinburgh Napier University UK

5 Chris Chambers, Cardiff University UK

Discuss this article

Comments (3)

Corresponding author: Jonathan P. Tennant (jon.tennant.2@gmail.com)

How to cite this article: Tennant JP, Waldner F, Jacques DC *et al.* **The academic, economic and societal impacts of Open Access: an evidence-based review [version 3; referees: 3 approved, 2 approved with reservations]** *F1000Research* 2016, **5**:632 (doi: 10.12688/f1000research.8460.3)

Grant information: This research was partly funded by the Belgian National Fund for Scientific Research through a FRIA grant. PM acknowledges support from the European Commission Horizon 2020 Programme under Grant Agreement 634107 (PHC32-2014) 'MULTIMOT'.
The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests: JPT currently blogs for the PLOS Paleo Community, and works for ScienceOpen. CHJH is a Center for Open Science ambassador. DCJ and FW are members of the Open Access Working Group of EURODOC. PM is a Research Data Alliance member. LBC works for the University of Pittsburgh, which has an Open Access library publishing department. All views presented here are strictly personal.

First published: 11 Apr 2016, **5**:632 (doi: 10.12688/f1000research.8460.1)

REVISED **Amendments from Version 2**

The final version of this manuscript includes minor edits reflected in the last review by Peter Suber, as well as stylistic edits noted by other commenters. We hope that this paper will continue to be shared and discussed, and that it helps with future developments in Open Access.

See referee reports

Introduction

Open Access (OA) refers to the removal of major obstacles to accessing, sharing and re-using the outputs of scholarly research. The rationale is that the research process is facilitated by ensuring rapid and widespread access to research findings such that all communities have the opportunity to build upon them and participate in scholarly conversations. As such, the major drivers behind OA relate to within- and between-community equality (Veletsianos & Kimmons, 2012), as well as bridging the global North-South research divide (Adcock & Fottrell, 2008). Reflecting this ambition, there are currently over 700 OA policies and mandates recorded worldwide from a range of research institutes and funding bodies (roarmap.eprints.org). OA pertains to documents made available via two main pathways: the Gold route and the Green route (Harnad *et al.*, 2008). The Gold route refers to freely accessible research articles at the point of publication. This route is often, although not always, accompanied by article processing charges (APCs). The Green route refers to author self-archiving, in which peer-reviewed articles and/or not peer-reviewed pre-prints are posted online to an institutional and/or subject repository, or to a personal website. This route is often dependent on journal or publisher policies on self-archiving (sherpa.ac.uk/romeo). Some publishers require an embargo period before deposition in public repositories is allowed. These embargoes are applied in order to avoid putative reductions in subscription income due to such self-archiving, although there is little evidence to support the existence of such embargoes (Berners-Lee *et al.*, 2005; Bernius *et al.*, 2013; Henneken *et al.*, 2006; Houghton & Oppenheim, 2010; Swan & Brown, 2005). The Green route is also enabled through author rights retention, in which authors pre-emptively grant non-exclusive rights to their institutions before publishing any works. The institution then has the ability to make articles by these authors OA without seeking permission from the publishers (e.g., this is the case of the Dutch Taverne amendment that has declared self-archival of research after 'a reasonable period of time' a legal right (Open Access NL, 2015)). Through these dual pathways, almost 25% of all scholarly documents archived on the Web are now obtainable via OA somewhere on the Internet (Khabsa & Giles, 2014).

A core issue remains: universal or even marginal access to approximately 75% of articles is not directly possible unless one either is in a privileged position to work at an institute that has subscription access to these articles, or has enough money to pay on a per-article basis (given that journals provide this feature; some do not). Subscriptions to all peer-reviewed journals is not affordable for any single individual, research institute or university (Odlyzko, 2006; Suber, 2012). Consequently, the potential impact of research articles is never fully realized, impeding scientific progress by a lack of use, while simultaneously negatively affecting the recognition of individual researchers (Hitchcock, 2013) and the funders who support their work.

Because of these issues, free and unrestricted access to primary research literature has become a global goal of the OA movement. The steady increase in OA over the past two decades has required careful negotiations between a range of stakeholders (e.g., librarians, funders, academics). Much of the driving force behind this global change has been through a combination of direct, grassroots advocacy initiatives and policy reforms from universities, funders and governments. The debates regarding the benefits of OA over subscription-based access often hinge on the increased value to academics. However, increased access has broader benefits to research through enhanced visibility, facilitating innovation by businesses and decreasing financial pressure on academic/research libraries (known more broadly as the 'serials crisis' (McGuigan & Russel, 2008)). Additionally, increased access to scholarly outputs might help foster a culture of greater scientific education and literacy, which in turn could have a direct impact on public policy (European Commission, 2012; Zuccala, 2010), particularly in domains such as climate change and global health, as well as increasing public engagement in scientific research (Stodden, 2010). OA also includes a moral aspect, where access to scientific knowledge and information is regarded as a fundamental feature of global human equality. For example, Article 27 of the United Nations Declaration of Human Rights states that "*Everyone has the right to freely participate in the cultural life of the community, to enjoy the arts and to share in scientific advancement and its benefits.*" (United Nations, 1948).

This review aims to provide information on the various impacts of OA to scholarly research. We consider the impact of OA from the academic, economic, and societal perspective. In addition, we shortly consider the broader implications of OA on Open Data, a closely related aspect united under a general theme of Open Research or Open Science. By aggregating evidence from a range of primary sources, this review should be useful to those broadly interested in the impact of open scholarly research, as well as policymakers and others involved in implementing OA policies and strategies. We refrain from making predictions about the future of OA publishing or policy recommendations, as these are both beyond the scope of this work.

A brief history of Open Access

The OA movement is intrinsically tied to the development of the Internet and how it redefined communication and publishing (Laakso *et al.*, 2011). With increased availability of Internet bandwidth, print articles have become virtually redundant, and sharing of information has never been cheaper. As a consequence, the costs per research article should have potentially decreased as a result of not investing material resources in publications printing and distribution. Therefore, widespread dissatisfaction with the expensive traditional publishing model has increased, resulting in the OA movement and concomitant innovations in scholarly publishing. A comprehensive timeline of the OA movement is provided as part of the Open Access Directory (oad.simmons.edu/oadwiki/Timeline).

Interest in using the Internet for facilitating access to scientific research coalesced throughout the 1990s, culminating with the 2001 conference on "Free Online Scholarship" by the Open Society Institute in Budapest. The result of this conference was the release of the Budapest Open Access Initiative (BOAI), which is recognized as one of the defining points of the OA movement. The BOAI was the first initiative to use the term "Open Access" and articulated the following definition:

> *By "open access" to [peer-reviewed research literature], we mean its free availability on the public internet, permitting any users to read, download, copy, distribute, print, search, or link to the full texts of these articles, crawl them for indexing, pass them as data to software, or use them for any other lawful purpose, without financial, legal, or technical barriers other than those inseparable from gaining access to the internet itself. The only constraint on reproduction and distribution, and the only role for copyright in this domain, should be to give authors control over the integrity of their work and the right to be properly acknowledged and cited.*

This definition is broadly equivalent to the Creative Commons Attribution license (CC-BY), which is widely considered to be a standard for OA (creativecommons.org/licenses/). One result of the growing OA movement is the rise of OA-only publishers, who publish exclusively digital content and have demonstrated that such a business model is financially feasible (but does not necessarily sustain the current journal ecosystem). Some of these publishers are for-profit and some are non-profit. For example, pioneer OA publishers BioMed Central (for-profit) and the Public Library of Science (PLOS) (non-profit) were founded in the early 2000s and remain successful OA publishing businesses to date. More recently, OA publishing has gained increasing momentum among researchers, funders, and governments. This has led to a proliferation of innovative approaches to publishing (e.g., *PeerJ*, peerj.com; *F1000Research*, f1000research.com; *Open Library of Humanities*, openlibhums.org) and a range of different policies from research funders and institutes mandating OA. All of these different policies and new business models, combined with traditional publishers launching their own OA titles and programs, have made the overall OA ecosystem quite complex.

Even with this growing prevalence of publishers that facilitate OA to the scholarly literature, OA is still hardly ubiquitous. Bjork *et al.* (2009) estimated that the total number of published articles in 2006 was approximately 1,350,000. Of these, 4.6% became immediately accessible and an additional 3.5% became accessible after an embargo period of typically one year. Furthermore, usable copies of 11.3% could be found in repositories or on the author's home pages. Since the U.S. National Institutes of Health (NIH) mandated archival of articles in the public PubMed Central repository in 2008, the cumulative number of OA articles in PMC has increased more than the number of non-OA articles (see Figure 1). In 2013, the total percentage of OA articles available was estimated at 24% of English-language scholarly documents accessible on the Web (Khabsa & Giles, 2014).

Although these estimates show OA is on the rise, the full potential of OA is far from achieved. Björk *et al.* (2014) pointed out that 62% of journals (from the top 100 journal publishers indexed by SCOPUS) endorse immediate Green OA self-archiving by their authors, 4% impose a 6-month embargo, and 13% impose a 12-month embargo. As such, 79% of articles published in any recent year could already be OA within 12 months after publication via Green OA, 62% of them immediately if authors were actually self-archiving properly (Gargouri *et al.*, 2010; Gargouri *et al.*, 2012). The disconnect between practice and what is allowed has three potential explanations: (i) researchers are unsure whether they have the legal right

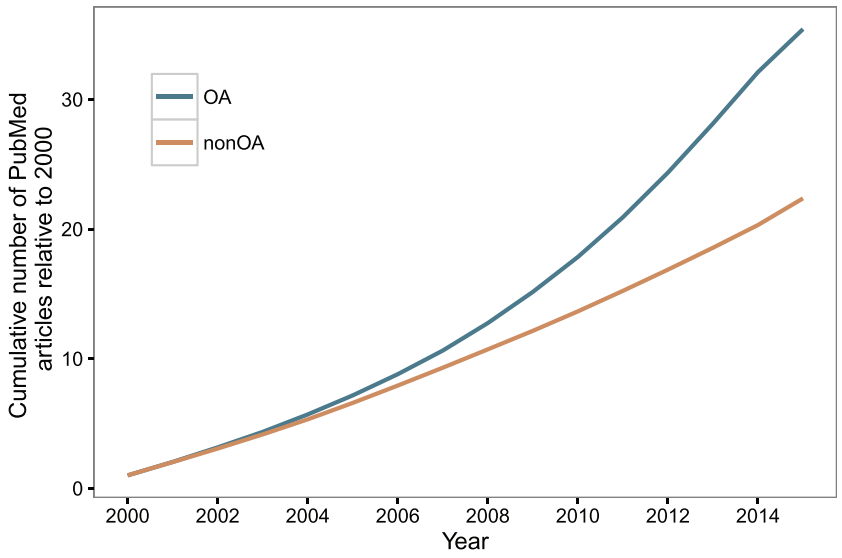

Figure 1. Percentage increase in research articles in PubMed Central, relative to 2000. Since 2004, the growth rate of OA articles is significantly higher than that of non-OA articles.

to self-archive, (ii) they fear that it might put their article's acceptance for publication at risk, and (iii) they believe that self-archiving may be a lot of work (Harnad, 2006). Research funders and institutions worldwide are now beginning to realize that they need to alter their conditions to make OA *mandatory* (Vincent-Lamarre *et al.*, 2016) in order to counteract these misconceptions of self-archiving (Carr *et al.*, 2007; Swan & Brown, 2005). Swan & Brown (2005) have indicated that the vast majority of researchers (81%) would comply with mandatory OA if it were a condition of funding. On the other hand, it is worth mentioning that ensuring compliance with OA policies set by research institutions is rather difficult. Some tools, such as the *Open Access Monitor* (http://symplectic.co.uk/elements-updates/introducing-open-access-monitor), help institutions to track compliance with their OA policy.

Table 1 shows a non-exhaustive summary of the developments in the advancement of scholarly publishing and the OA movement. Included are the founding of major institutions in the movement as well as policy and legal developments. Several controversial moments are included, because they have spurred action or generated awareness for the movement. One of them is the suicide of Aaron Swartz, who was arrested for downloading JSTOR articles on the grounds that he allegedly intended to make these publicly available. Another ongoing controversy is scholarly piracy; this includes the Sci-Hub and LibGen projects, which have created an online repository of pirated scholarly papers (around 50 million at the time of writing). Both projects gained increased attention after becoming the target of a lawsuit by the publisher Elsevier. There have been mixed responses to these kinds of activities, polarising the view that illegal acts regress or weaken the case for OA, while some hail the development as the 'Napster moment' (i.e., a change inducing disruption; Rosenwald, 2016) for the OA movement, which will force the established industry to change. Regardless of the legality of it, Sci-Hub is used by a large number of people from all over the world to access research articles (Bohannon, 2016; Elbakyan & Bohannon, 2016).

The effect of Open Access upon academia

The two main ways in which OA affects academia are (i) through association with a higher documented impact of scholarly articles, as a result of availability and re-use, and (ii) through the possibility of non-restrictively allowing researchers to use automated tools to mine the scholarly literature. For the former, major arguments in favor of OA include the evidence that work that is openly available generates more academic citations, but also has more societal impact. In addition, appropriately-licensed OA works play a major role in academic education, including re-use in classes and for dissertations. The latter major argument involves non-restrictive access to the scholarly literature through appropriate licensing, making it possible to use automated tools to collect and analyze the entire body of scholarly literature in a legally sound framework and irrespective of copyright laws. The following sections cover these two effects of OA.

The potential impact advantage

Academic impact. Academic impact is frequently measured through citation counts, and these remain fundamental as the 'currency units' for researchers, research groups, institutes and universities. Lawrence (2001) was the first to propose that OA would have a citation advantage. The utility and consistency of the citation advantage across different research fields has been intensively debated because its magnitude substantially varies depending on the discipline (Table 2). However, the general tendency identified by studies to date indicates that there is at least some association between OA publishing and increased citation counts across most disciplines (Antelman, 2004; Hajjem *et al.*, 2006) (Figure 2 and Table 2). A comprehensive and annotated bibliography of studies documenting potential citation impacts was created by Steve Hitchcock (eprints.soton.ac.uk/354006/1/oacitation-biblio-snapshot0613.html) and has been managed by SPARC Europe since 2013 (sparceurope.org/oaca/).

Estimates for the open citation advantage range from +36% (Biology) to +600% (Agricultural Sciences) (Swan, 2010; Wagner, 2010). In a longitudinal study, Eysenbach (2006) compared the bibliometric impact of a cohort of articles from a multidisciplinary journal (*Proceedings of the National Academy of Sciences*) that offers both OA and non-OA publishing options. After adjusting for potentially confounding variables, the results indicated that non-OA papers were twice as likely to remain uncited six months after publication when compared to OA articles. Additionally, the average number of citations for OA articles was more than double than that of the non-OA articles. The study also differentiated the type of OA article, namely the self-archived (i.e., Green OA) and the publisher version of record (VOR) that is freely available (i.e., Gold OA). Gold OA was found to have a higher overall academic impact than Green OA.

Despite strong evidence for a citation advantage, the magnitude of this advantage remains variable. The substantial heterogeneity in observed citation advantages can be due to different academic cultures or could simply be spurious. For example, self-archiving prior to publication is a community standard in fields such as high energy physics or mathematics, but has yet to be widely adopted among the life sciences. Such 'pre-prints' have also been associated with an overall increase in the average number of citations, the total number of citations, and the speed of citation accumulation (Aman, 2014; Gentil-Beccot *et al.*, 2010). Other studies could only replicate immense citation advantages (+600%) if relevant predictors were omitted (McCabe & Snyder, 2014), which indicates a potential spurious effect. When taking into account these relevant predictors, the citation advantage became much smaller (+8%). When the citation advantage is low or non-existent, this could suggest that in those research fields there is a sufficient level of access to the literature such that OA confers no localised access advantage, or that adoption of OA has not yet reached a level where any such advantage has become statistically evident.

One alternative explanation for the existence of citation advantages could be that researchers choose to publish OA when a finding is more impactful, but empirical evidence contradicts this selection effect. Gargouri *et al.* (2010) compared citation counts for articles which were self-selected as OA or mandated as OA (e.g., by funders). The study concluded that both were cited significantly more than non-OA articles and showed no differences in citation rates. As such, these findings rule out a selection bias

Table 1. Major historical milestones in the progress of Open Access publishing.

Year	Milestone
1454	Invention of **printing**
1665	January 5: First issue of The ***Journal des sçavans*** (later spelled *Journal des savants*), the earliest academic journal published in Europe and established by Denis de Sallo.
1807	25-year-old **Charles Wiley** opens a small printing shop at 6 Reade Street in lower Manhattan.
1842	May 10: **Julius Springer** founded what is now Springer Science+Business Media in Berlin.
1848	**John Wiley** (son of Charles Wiley) gradually started shifting his focus away from literature toward scientific, technical, medical, and other types of nonfiction publishing.
1880	Foundation of **Elsevier**.
1936	First scientific book published by **Elsevier**.
1990	First **web page**.
1991	An online repository of electronic preprints, known as e-prints, of scientific papers is founded in Los Alamos by the American physicist Paul Ginsparg. It was renamed to **ArXiv.org** in 1999. The total number of submissions by May 11st, 2016 (after 24.8 years) is 1,143,129 (arxiv.org/stats/monthly_submissions).
1993	Creation of the **Open Society Institute** (renamed to the Open Society Foundations [OSF] since 2001) by the progressive liberal business magnate George Soros. The OSF financially supports civil society groups around the world, with a stated aim of advancing justice, education, public health and independent media.
1997	Launch of **SciELO** in Brazil. There are currently 14 countries in the SciELO network and its journal collections: Argentina, Bolivia, Brazil, Chile, Colombia, Costa Rica, Cuba, Mexico, Peru, Portugal, South Africa, Spain, Uruguay, and Venezuela.
1998	**Public Knowledge Project** (PKP) is founded by John Willinsky in the Faculty of Education at UBC, with Pacific Press Professorship endowment, dedicated to improving the scholarly and public quality of research.
	PKP has created the **Open Conference Systems** (2000), **Open Journal Systems** (2001), **Open Harvester Systems** (2002) and the **Open Monograph Press** (2013).
2000	**BioMed Central**, the self-described first and largest OA science publisher and **PubMed Central**, a free digital repository for biomedical and life sciences journal, is founded. In 2008, Springer announces the acquisition of BioMed Central, making it, in effect, the world's largest open access publisher.
2001	An online petition calling for all scientists to pledge that from September 2001 they would discontinue submission of papers to journals which did not make the full-text of their papers available to all, free and unfettered, either immediately or after a delay of several months is released. The petition collected 34,000 signatures but publishers took no strong response to the demands. Shortly thereafter, the **Public Library of Science** (PLOS) was founded as an alternative to traditional publishing. *PLOS ONE* is currently the world's largest journal by number of papers published (about 30,000 a year in 2015).
	December 1–2: **Conference convened in Budapest** by the Open Society Institute to promote open access – at the time also known as Free Online Scholarship. Where the Budapest Open Access Initiative (BOAI) was born.
2002	February 14th: Release of the **Budapest Open Access Initiative** (BOAI), a public statement of principles relating to OA to the research literature. This small gathering of individuals is recognised as one of the major defining events of the OA movement. On the occasion of the 10th anniversary of the initiative, it was reaffirmed in 2012 and supplemented with a set of concrete recommendations for achieving "the new goal that within the next ten years, Open Access will become the default method for distributing new peer-reviewed research in every field and country."
	Start of the Research in Health - **HINARI** programme of the World Health Organization and major publishers to enable developing countries to access collections of biomedical and health literature online at reduced subscription costs. Together with Research in Agriculture - **AGORA**, Research in the Environment - **OARE** and Research for Development and Innovation - **ARDI** programmes, it currently forms **Research4Life** that provides developing countries with free or low cost access to academic and professional peer-reviewed content online.
2008	The **National Institutes of Health** (NIH) Public Access Policy, an OA mandate requiring that research papers resulting from NIH funding must be freely and publicly available through PubMed Central within 12 months of publication, is officially recorded.
2009	The **Fair Copyright in Research Works Act** (Bill H.R 801 IH, also known as the "Conyers Bill") is submitted as a direct response to the National Institutes of Health (NIH) Public Access Policy; intending to reverse it. The bill's alternate name relates it to U.S Representative John Conyers (D-MI), who introduced it at the 111th United States Congress on February 3, 2009.
2011	Arrest of **Aaron Swartz** after he systematically downloaded articles from JSTOR, for alleged copyright infringement.
	In reaction to the high cost of research papers behind paywalls, **Sci-Hub**, the first known website to provide automatic and free, but illegal, access to paywalled academic papers on a massive scale, is founded by Alexandra Elbakyan from Kazakhstan.

Year	Milestone
2012	Start of the **Academic Spring**, a trend wherein academics and researchers began to oppose restrictive copyright in traditional academic journals and to promote free online access to scholarly articles.
	Start of the **Cost of Knowledge** campaign which specifically targeted Elsevier. It was initiated by a group of prominent mathematicians who each made a commitment to not participate in publishing in Elsevier's journals, and currently has over 15,933 co-signatories.
	Start of the United States-based campaign **Access2Research** in which open access advocates (Michael W. Carroll, Heather Joseph, Mike Rossner, and John Wilbanks) appealed to the United States government to require that taxpayer-funded research be made available to the public under open licensing. This campaign was widely successful, and the directive and FASTR (the Fair Access to Science and Technology Research Act) have become defining pieces in the progress of OA in the USA at the federal level.
	Launch of ***PeerJ***, an OA journal that charges publication fees through researcher memberships, not on a per-article basis, resulting in what has been called "a flat fee for 'all you can publish'". Note that as of October 2015 *PeerJ* also have a flat rate APC of $695.
2013	January: The suicide of **Aaron Swartz** draws new international attention for the Open Access movement.
	November: **Berlin 11 Satellite Conference** for students and early career researchers, which brought together more than 70 participants from 35 countries to engage on Open Access to scientific and scholarly research.
2014	First **OpenCon** in Washington DC, an annual conference for students and early career researchers on Open Access, Open Data, and Open Educational resources.
	Open Access is embedded the European Commission's **Horizon 2020** Research and Innovation programme.
2015	Academic publisher **Elsevier** makes a complaint in New York City for copyright infringement by **Sci-Hub.** Sci-Hub is found guilty and ordered to shut down. The website re-emerges under a different domain name as a consequence. A second hearing in March 2016 is delayed due to failure of the defendant to appear in court, and to gather more evidence for the prosecution.

from authors as the cause for the citation advantage (Gargouri *et al.*, 2010). However, research that is selected to merit funding by funding agencies may, in itself, be perceived to be more impactful than research that is not funded. Additionally, as no single OA mandate is ever 100% effective, it might be the simple case that authors are more likely to comply with a mandate for the research they perceive to be of higher impact. In a study of articles in the field of psychology, Anderson (2013a) found that publications with funding sources reported in the text were found to be more highly cited and connected to other highly-cited publications (this type of publication is called "generative" in the study) than publications with no reported funding sources. Furthermore, research that was privately funded was found to be more generative than publicly funded research. In a similar study in the Library and Information Sciences field done by Zhao (2010), the citation counts for grant-funded publications were "substantially higher" than publications without grant funding. Although these studies indicate that grant funding is correlated with increased citation rates, the openness of articles was not addressed in either study. Future research will be required to demarcate the potential causality and to determine the conditions under which we could see whether or not OA has an effect on citation counts. For example, this could be conducted through a randomised controlled trial in which research articles from a particular funder are randomly assigned to OA and non-OA routes, with the citation counts assessed after a certain time.

In sum, evidence indicates that OA is broadly related to increased academic impact in terms of citations (Figure 2; see also McKiernan *et al.* (2016)), but given the large variability in results, further research should aim to synthesize these findings in a meta-analysis and try to explain the cause of this variability.

Broader societal impact. Scholarly articles also have a societal impact, such as when they are covered in news media or are discussed in social media channels; alternative metrics, or altmetrics, can be used as a guide to measure this mode of impact (Liang *et al.*, 2014). Information such as social media usage, Mendeley readership, and media attention (Piwowar, 2013) can be tracked by various altmetrics providers (e.g. ImpactStory, Plum Analytics, and Altmetric.com). As such, when an article generates discussions outside of the academic literature, altmetrics are capable of tracking this. Despite limitations (such as academics discussing their own research on platforms like Twitter), altmetrics provide a general view of the wider societal impact of research articles. Considering the increased pressure on researchers and research institutes to communicate research findings to the public, altmetrics can provide additional insight into which research drives public interest. A working group established by NISO is investigating the future role of altmetrics in research communication and assessment (www.niso.org/topics/tl/altmetrics_initiative/).

OA articles would be expected to have an altmetrics advantage compared to the non-OA literature; if an article has fewer restrictions for journalists, citizens, businesses, and policy-makers, it seems logical that this would enable the research to be publicly re-used. Furthermore, those parties may be more likely to promote articles which are publicly accessible into different communication channels. In other words, increased access removes barriers to widespread societal engagement, whereas a relative lack of article access discourages engagement.

There is research showing evidence for an altmetrics advantage for OA articles, but this does not reflect itself in the most

Table 2. Main scientific papers that have investigated and quantified the citation advantage as well as its origin.

Reference	Discipline	Citation advantage	Origin
Antelman (2004)	Mathematics, Electrical Engineering, Political Science, Philosophy	+91%, +51%, +86%, +45% per discipline respectively	NA
Atchison & Bull (2015)	Political Science	Statistically significant citation advantage	NA
Cheng & Ren (2008)	Medicine, Biology, Agricultural Sciences, Chemistry and University Journals	+200%	NA
Davis & Fromerth (2007)	Mathematics	+35%	Quality advantage, no evidence of early advantage
Davis *et al.* (2008)	Physiology	-5%	NA
Davis (2011)	Sciences, Social Sciences, and Humanities	+1% but statistically indistinguishable	No evidence of an early advantage
Evans & Reimer (2009)	All	+8% for newly published articles; +16% for citations from developing countries	NA
Eysenbach (2006)	Natural Sciences	+210 up to +290%	NA
Frandsen (2009)	Biology, Mathematics, Pharmacy and Pharmacology	No clear tendency towards an increase in impact	NA
Gargouri *et al.* (2010)	Engineering, Biology, Biomedicine, Chemistry, Psychology, Mathematics, Clinical Medicine, Health, Physics, Social Science, Earth Sciences	+?% to ?% depending on the discipline	Quality advantage is confirmed no evidence for selection bias
Gaule & Maystre (2011)	Biology	No evidence of citation advantage	NA
Gentil-Beccot *et al.* (2010)	High Energy Physics	+200%	Early advantage confirmed
Hajjem *et al.* (2006)	Biology, Psychology, Sociology, Health, Political Science, Economics, Education, Law, Business, Management	+36% to 172%	NA
Harnad & Brody (2004)	Physics	+250% to 580%	NA
Henneken *et al.* (2006)	Astronomy and Physics	+200%	NA
Kousha & Abdoli (2010)	Agricultural Science	+621% but not to every journal	NA
Kurtz *et al.* (2005)	Astronomy	None	Selection bias and early advantage
Kurtz & Henneken (2007)	Astronomy	+200%	Early advantage confirmed
Lansingh & Carter (2009)	Opthalmology	No	NA
Lawrence (2001)	Computer Science	+157% up to +284% for top publication	NA
McCabe & Snyder (2014)	Ecology, Botany, Multidisciplinary Science and Biology	+8%	NA
McVeigh (2004)	Natural Sciences	0-+50% in 2003 depending on field, negative citation advantage in 2000	NA
Metcalfe (2005)	Astronomy	+200%	NA
Metcalfe (2006)	Solar Physics	+170% and +260% depending on the online repository	No evidence for selection bias
Moed (2006)	Condensed Matter Physics	NA	Confirm early access advantage and selection bias but no OA effect

Reference	Discipline	Citation advantage	Origin
Norris *et al.* (2008)	Ecology, Applied Mathematics, Sociology and Economics	+157%	NA
Sahu *et al.* (2005)	Medicine	+300% up to +450%	NA
Schwarz & Kennicutt Jr (2004)	Astronomy	+200%	Early advantage
Vanclay (2013)	Environmental Science	Not significant	NA
Wang *et al.* (2015)	All	+111% up to 152%	NA
Wohlrabe & Birkmeier (2014)	Economics	+35% up to 64% depending on the database used	NA
Xu *et al.* (2011)	Humanities, Life Sciences, Mathematics & Physical Science, Medicine, Social Sciences	-49.24%-+87.73%	NA
Zhang (2006)	Communication Studies	+200%	NA

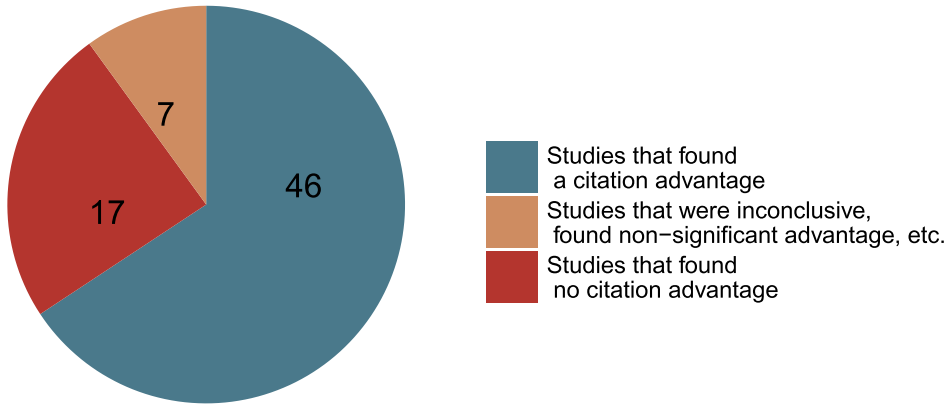

Figure 2. Studies that investigated the citation advantage grouped by their conclusion. The majority concluded that there is a significant citation advantage for Open Access articles. Source: Data from The Open Access Citation Advantage Service, SPARC Europe, accessed March 2016.

impactful articles. Wang *et al.* (2015) found evidence that OA articles receive more attention through social media. The authors compared social media attention (Twitter and Facebook) between OA and non-OA articles at *Nature Communications* and found that OA articles get 1.2–1.48 times as much social media attention as compared to non-OA articles (see also Adie, 2014). Nonetheless, of the top 100 articles of 2015 as presented by Altmetric.com, only 42 articles were OA (www.altmetric.com/top100/2015/). This 42% is larger than the overall proportion of OA articles in the literature, which indicates that OA contributes relatively more impact per paper. However, it also indicates that the open impact advantage can be overshadowed by the intrinsic nature of the research published or by the traditionally prestigious journals with a larger and dedicated media apparatus (e.g., Nature, Science; Brembs *et al.*, 2013).

Allen *et al.* (2013) found that a social media announcement of the release of a research article increases the number of users who view or download that article, but does not translate to increases in the citation count in the field of clinical pain research. Costas *et al.* (2015) found a relatively weak correlation between social media activity and citation counts for the articles in their sample (over 1.5 million article records), while Mohammadi *et al.* (2015) found that the number of Mendeley readers with a status of graduate student or faculty correlated with citation counts. When OA to the articles is factored into an analysis, there is a potential recursive relation between citation counts and altmetrics due to OA. Eysenbach (2011) indicated that there is a moderate correlation (0.42–0.72) between the tweets and citations of articles from an OA journal (*Journal of Medical Internet Research*). Highly tweeted articles were eleven times more likely to be highly cited

than less-tweeted articles, or vice versa (75% of highly tweeted articles were highly cited; 7% of less-tweeted articles were highly cited). However, it is difficult to assess causality in these cases: do research papers that have more academic impact make their way more frequently into societal discussions, or does online discourse increase their potential citation rates? Overall, this evidence implies that there is a general media advantage with OA (see also McKiernan *et al.* (2016)), which can be used as a proxy or pathway to indicate greater societal impact.

Altmetrics themselves should not be conflated with citations when it comes to assessing impact, even though some providers such as Altmetric.com supply a single score that can be used to rank an article in a similar way to a journal's Impact Factor. Each measure of altmetrics tells a different story about the impact of research, and a careful understanding of the altmetrics landscape in conjunction with citation-based metrics can lead to a clearer picture of societal impact of scientific research.

Open Access and text- and data-mining

Traditionally, in order to publish a paper, researchers hand over their copyright via a Copyright Transfer Agreement. Copyright transfer as the default has far-reaching consequences on the ability of both the authors and others to re-use that published research, and many authors are not aware of the impact of these transfers on their ownership of the work. Academics frequently give the copyright to the publishers in exchange for the perceived prestige of publishing in one of their venues (e.g., Müller-Langer & Watt, 2010). In some cases, institutes adopt rights-retention OA policies that grant authors non-exclusive rights to their institutes before signing copyright agreements with publishers, which enables them to make articles OA without requiring permission from publishers (cyber.law.harvard.edu/hoap/Good_practices_for_university_open-access_policies). Essentially, copyright is a pre-digital tool wielded by traditional publishers to maintain revenues rather than fostering creativity, innovation, or protecting authors (Okerson, 1991; Willinsky, 2002). For example, the Author's Guild sued Google Books for copyright infringement because they provided freely available digital copies; the court rejected this suit in 2016, stating that Google Books served the public interest and that copyright's "primary intended beneficiary is the public" (EFF, 2015). In the digital age, copying is essential to perform necessary research tasks. These activities range from viewing the article (i.e., downloading requires copying) to re-using figures from an article in a book. The interaction of OA and copyright is complex and deserves extended research in itself (e.g., Scheufen, 2015). We will highlight how OA views copyright and relate this to its effects on text- and data-mining (TDM).

The majority of 'born OA' journals and publishers do not request or receive copyright from authors. Instead, publishers are granted non-exclusive rights to publish, and copyright is retained by authors through a Creative Commons license (typically CC-BY). Importantly, this represents a power shift from publisher-owned to author-owned rights to research. This model of author-retained copyright appears to be favoured by the majority (71%) of the research community (Hoorn & van der Graaf, 2006). Shifting copyright to stay with the author, combined with appropriate open licensing, allows for wider re-use, including TDM, and forms the basis for a robust scholarly ecosystem.

As such, copyright in OA publications is non-restrictive and also allows machines to freely access it. In traditional publishing, human reading and computer reading are seen as two separate things which require different agreements, whereas OA publishing views them both in the same, non-restrictive manner. In other words, in order to mine OA journals, one only needs the technical skills to do so. In order to mine closed access journals, one needs to sign or negotiate access conditions, even if legitimate access to the articles has already been bought (Bloudoff-Indelicato, 2015).

Automated extraction of information from scholarly research via TDM is a methodology that can be applied to investigate the scholarly literature at an enormous scale, creating new knowledge by combining individual findings. This has already proven to be useful for a large variety of applications (e.g., Glenisson *et al.*, 2005; Martone *et al.*, 2016; Swanson, 1987). Moreover, OA publishers facilitate TDM on a massive scale by allowing multiple options for collecting the literature needed. For example, PLOS is non-restrictive and allows users to scrape articles directly from the website or using its API. As a result, scraping tools can be used, such as `rplos`, an R package developed to search and download full-text scholarly papers (Chamberlain *et al.*, 2016).

TDM is not only a knowledge-generation tool; it also allows for automated screening for errors and automated literature searches that renew scientific discovery. With TDM it becomes possible to easily compare one's results with those of the published literature, identify convergence of evidence and enable knowledge discovery (Natarajan *et al.*, 2006) or discover frequent tentative hypotheses that can be used for new research (Malhotra *et al.*, 2013). It has already been used to make major advances in fields such as biomedicine (Gonzalez *et al.*, 2016). TDM also allows for computer applications that can download all scholarly literature given certain search terms (e.g., ContentMine's `getpapers` tool ; github.com/ContentMine/getpapers), simplifying and shortening the tedious literature search. TDM can also serve a screening purpose similar to plagiarism scanners, helping to detect statistical errors in the scholarly literature (e.g., Nuijten *et al.* (2015)). TDM can be used in various innovative ways and is an emerging and rapidly advancing field; non-restrictive licensing through OA certainly promotes its wider application.

Given the exponential increase in the number of scholarly publications, (semi-)automated methods to synthesize results have become increasingly important. TDM decreases the time dedicated to the search for relevant information in scholarly literature by categorizing information (Leitner & Valencia, 2008), highlighting and annotating relevant results to specific users (Shatkay *et al.*, 2008), and profiling research (Porter *et al.*, 2002). Furthermore, TDM also prevents researchers and readers from

wasting time on reinventing the wheel, simply because one can no longer keep up with the huge amount of published literature available (Harmston *et al.*, 2010).

Because of traditional copyright transfers, TDM has often been stymied by traditional, closed access publishers who frequently see it as a copyright infringement. Researchers using software that harvests data from online publications have been (threatened to be) cut off from accessing the articles. These researchers found themselves trapped in negotiations to resume their research (even though their universities had paid subscription fees for access (e.g., Bloudoff-Indelicato, 2015; Van Noorden, 2012)). Standard subscriptions do not permit systematic downloads because publishers fear that their content might be stolen and their revenue therefore lost (Van Noorden, 2012). In 2014, Elsevier opened its papers for TDM via a proprietary API (Van Noorden, 2014), but placed restrictions on the researchers using it; however, researchers are not legally required to comply with these restrictions in some countries (e.g., U.K., U.S.A., Handke *et al.*, 2015).

To make the enormous corpus of closed access papers retrospectively available to the public might be possible through legal action at an institutional or governmental level. The Dutch Government, for example, has recognized OA as a right, with Dutch citizens capable to make their scientific publications free to access after a 'reasonable period of time' (Open Access NL, 2015). Such steps are further supported by Shavell (2010) and Eger & Scheufen (2012) who ascertained that transition towards an OA model could not be smooth without first undertaking the necessary legislative steps. The position of institutes regarding copyright transfer remains generally unclear. While academics themselves may have little power in broader debates regarding copyright, institutes could claim ownership of the work by invoking their rights under the work made-for-hire doctrine (Denicola, 2006). However, OA policies at universities generally use a system of non-exclusive rights, presupposing that faculty are the owners of their work and can grant non-exclusive rights to the university for use (for examples of approaches and language used when drafting OA policies, see (Shieber & Suber, 2016)). Importantly, this means that OA through the 'Green' route does not always depend on permission from publishers, and increasingly is becoming dependent on rights retention by authors, through carefully-drafted and widely-supported university policies. While these are positive steps towards making research available for TDM, in light of the potential copyright problems for closed access articles and the fact that not all research is available through institutional Open Access policies, TDM will be easier and legally safer for OA journals. As a consequence, TDM is likely to be more readily applied to OA literature when compared to closed access literature.

The economic impact of Open Access

The effect on publishers

Any publisher has to cover operating costs, which are primarily made of (i) article processing charges (APCs), (ii) management and investment costs, and (iii) other costs. Article processing includes editing, proofreading and typesetting, among other things. Management and investment are instead the marginal costs needed to establish and keep the journal running. Other costs include promoting the journal, hosting and infrastructural services, sponsoring conferences, and other services that are extrinsic to research articles themselves. The average production cost for a single research article is estimated to be around $3500–$4000 (Van Noorden, 2013) but these costs are highly depending on the publisher. For example, Philip Campbell (Editor-in-Chief of *Nature*) stated that his journal's internal costs were at $20,000–$30,000 per paper (Van Noorden, 2013), due in part to the high selectivity and rejection rate at *Nature* (i.e., this is an average cost per published paper, and not the production costs associated with publishing a single accepted paper). However, these are at the high end of the cost spectrum, with other journals, such as the *Journal of Machine Learning Research* (JMLR) costing between $6.50–$10 per article (blogs.harvard.edu/pamphlet/2012/03/06/anefficient-journal/). Other publishers are completely transparent about their direct and indirect production costs, such as Ubiquity Press, which levies an APCs of $500 (ubiquitypress.com/site/publish/). One possible reason for such variation between journals and publishers is that it is generally unclear whether proposed costs relate to those directly involved in article processing or those required in order for a publisher to 'break even' if they receive zero subscription income for an article made OA.

In order to cover those costs and make a profit, closed access publishers charge for access via subscriptions, whereas many OA publishers or journals charge to publish. Due to increased subscription costs, closed access publishing is becoming an increasingly unsustainable business model (Odlyzko, 2013) with prices estimated to have outpaced inflation at 250% in the past thirty years (www.eff.org/issues/open-access). This will slowly but surely diminish the scope of access to the scholarly literature as fewer organisations are able to pay such high costs. Only recently has any transparency into the detailed costs of subscriptions been gained by using Freedom of Information Requests to bypass non-disclosure agreements between libraries and publishers (Bergstrom *et al.*, 2014; Lawson & Meghreblian, 2015). These requests provide the basis for understanding the economics of scholarly communication. For example, Bergstrom *et al.* (2014) found that commercial publishers, including Emerald, Sage, and Taylor and Francis, have prices of ten times the amount of non-profit publishers per citation for PhD-granting institutions. Two potential ways to prevent future retention of an unsustainable model is through decreasing the subscription prices, thereby lowering publishers' profit margins and the financial burden on subscribers, or through switching to new OA-oriented business models and creating new value. Either way, price transparency will be essential for future bargaining efforts between academic libraries and publishers, and will be of interest to those involved in public policy and scholarly publishing. The concept of transitioning from a subscription-based model to one driven by APCs will be financially appealing to journals that operate with minimal profits or at a loss, and can be a pathway to achieve financial security and long-term journal sustainability. As such, increasing revenues is a strong incentive for OA (osc.hul.harvard.edu/programs/journal-flipping/public-consultation/4/6/, accessed 26/04/2016).

OA publishing has become associated with a 'pay-to-publish' model, whereas around 70% of peer-reviewed OA journals do not

charge APCs, according to data from the Directory of Open Access Journals (DOAJ) (see blogs.harvard.edu/pamphlet/2009/05/29/what-percentage-of-open-access-journals-charge-publication-fees/ and citesandinsights.info/civ16i4.pdf). However, approximately 50% of all articles published in peer-reviewed OA journals are published in APC-based venues (Crawford, 2015; Laakso & Björk, 2012; Walters & Linvill, 2011). Authors paying to publish can be viewed as a fundamental conflict of interest for researchers. Nonetheless, this payment model has proven itself to function properly when editorial decisions are separated from the business-side of the publisher (i.e., editorial independence), removing the problem of 'publication-bribery'. Additionally, many journals have always levied charges for to cover the costs of publishing regardless of OA; for example, *PNAS* charges $1225 per regular research article (with an additional $1350 for OA; pnas.org/site/authors/fees.xhtml), and *Cell* charges $1000 for the first colour figure and $275 for each subsequent one (cell.com/cell/authors; as of April 2016). Therefore, equating OA with 'pay-to-publish' is actually a bit of a misnomer, as several closed journals charge to publish and many open journals do not. Furthermore, many publishers (e.g., *PLOS*, *PeerJ*), as well as many learned societies, operate fee waiver schemes for researchers unable to obtain funds to cover publication fees.

For those OA publishers implementing a pay-to-publish model, around 68.8% of publishers offer fee waivers to low- and middle-income countries (Lawson, 2015), while other journals offer fee discounts often given in lieu of total fee waivers. Solomon & Björk (2012) investigated the sources of funding used by authors for APCs, indicating that these are highly variable across academic disciplines. For example, while 45.5% of authors in Health Sciences, Biology and Life Sciences use grant or contract funding as source for APCs, only 10.4% use this in Business and Economics, with 45.8% coming from personal funds. Other sources include national funding bodies, and discretionary funds administered by institutions, as well as institutional funds specifically in place to support OA policies (see also Dallmeier-Tiessen *et al.*, 2011). Sources for APCs are also highly variable depending on the per capital GNP of the authors' country, as well as the size of the APC (Solomon & Björk, 2012). According to MacKie-Mason (2016), one potential outcome of authors seeing the price of APCs and securing funding for them is that authors may begin to take the price of APCs into account (in addition to other factors such as prestige and topic) when selecting a journal for their research output, which may drive market competition and could as a consequence lower the price of APCs. However, a potential negative consequence of an increasingly APC-driven model of OA is that some researchers may struggle to procure funds in order to publish and conform to mandates at different levels. This might impact early-career researchers and those working in fields were research grants and publishing fees are more difficult to obtain.

Subscription-based publishers still frequently produce print versions of their journals, which increases production costs, potentially to justify charging for readership or to satisfy a small demographic who prefers this mode of reading. After all, subscriptions to print journals make sense and, if large-scale printing is still in place, simply transferring this idea to the digital versions creates continuity. Print versions are accompanied by logistical costs to print and ship each issue, but these are partially offset with reprint orders, additional charges for colour figures, and print-based advertising. For some of the largest subscription-oriented publishers the annual net profit on investment reaches up to 40 percent, which makes academic journal publishing highly lucrative for investors (Satyanarayana, 2013), further increases investment to sustain this type of publishing model, and allows maintenance of an oligopoly (Larivière *et al.*, 2015).

OA publishers only publish digitally and have opened up avenues for innovation. For example, *PeerJ* has introduced a wholly different OA business model, where readers pay nothing to access articles, but authors pay a membership fee once to publish for a lifetime. The Open Library of Humanities (OLH) is another innovative business model in which libraries pay a small fee to support OLH and scholars are able to publish for free (subscription for publishing rather than subscription for access); this support also enables the OLH to help journals transition from a subscription model to OA (for example, the recent case of *Lingua*; timeshighereducation.com/research-intelligence/open-library-humanities-aims-flip-journals-open-access). Library publishing has also developed in response to the OA movement; in this model, academic libraries begin publishing operations in the interest of providing added value to their patrons and contributing to the growth of knowledge (librarypublishing.org). In terms of innovating in the publishing platform itself, *eLife* have introduced the Lens as a novel way of viewing research articles online (lens.elifesciences.org/about/), and *F1000Research* has introduced so-called 'living figures' to enable researchers to interact with data underlying research findings (e.g., Colomb & Brembs, 2015).

With this innovation comes massive scope for reducing the costs associated with publishing through implementing more efficient procedures. In this case, costs are reduced by eliminating the need for type-setting and copy-editing, with web-hosting costing only $15/year, and a total operating cost of between $6.50–$10.50 per article. Other platforms such as ARPHA offer an end-to-end XML-based publishing service, utilised by publishers including Pensoft, with a more efficient and integrated publishing workflow, which should highlight and reduce the real costs of publishing. In addition, OA has the potential to increase the speed of publication, as seen in journals like *eLife* and *PeerJ* (blog.dhimmel.com/plos-and-publishing-delays/), which combined with 'pre-print' servers like *biorXiv* and platforms that offer post-publication peer review like *Research Ideas and Outcomes* (riojournal.com/), *F1000 Research*, and *ScienceOpen* (www.scienceopen.com/), can exceptionally accelerate the speed of research communication. Such innovations add value to the research communication process (contrary to services such as paying to print colour figures) and represent just several cases of innovation across the publishing ecosystem. One can imagine that publishing costs in OA journals become dependent on the value added on a per-article basis, which can help reshape and improve scholarly communication. As such, making publication costs dependent on the value added aligns the interests of publishers with those of scholars, where improving the quality of the process of scholarly communication is the end goal. The motivation behind this could come from the currently available data, which suggest that hybrid publishing options offered

by traditional publishers, while being of higher cost due to supposed prestige, provide a much lower overall quality publishing process (blog.wellcome.ac.uk/2016/03/23/wellcome-trust-and-coaf-open-access-spend-2014-15/). It is noteworthy that in spite of the higher costs of hybrid publishing compared to 'pure' or 'born' OA publishing, some reports, such as the highly influential and somewhat controversial Finch Report in the UK (www.researchinfonet.org/publish/finch/), favoured the former model and high-priced Gold OA over a Green model.

The effect on non-publishers

The implementation of OA models has implications beyond the publishing industry in terms of economics. Research funding comes from multiple sources, including national funding agencies and industries, as well as private funders. Much primary research actually takes place outside of academia, inside R&D departments; if R&D in the private sector can access more research findings, this will ultimately benefit the public interest as well. A report from 2004 by Arzberger and colleagues into the scientific, social and economic development of access to research results concluded that access should be promoted to the largest extent possible. According to this report, access to research results can only be responsibly restricted in the case of national security, privacy, or those involving IP rights of the authors (Arzberger *et al.*, 2004). A major principle underlying this is the ownership of research results: publicly funded research and data are public goods and because they have been produced in the public interest they should be considered and maintained as such. Indeed, such a principle has become one of the focal rallying points of the global OA movement. Appropriate licensing and accessibility can influence re-use through commercialization, and can empower citizens and industry to recognize great economic benefits. This apparently resonates with many organisations, as indicated by the increased numbers of OA policies on a global basis (see Figure 3).

With access to scholarly articles, entrepreneurs and small businesses can accelerate innovation and discovery, which is advantageous for advancing the 'entrepreneurial state' (Mazzucato, 2011). Access to research results has clear advantages for a range of industries and can help stimulate regional and global economies. Increased access to research results has been associated with considerable increases of return on financial investment (Beagrie & Houghton, 2014). Furthermore, OA facilitates collaborations between publishers and industrial partners to leverage the potential of structured information networks for advanced data mining projects, such as that recently announced between IBM Watson and PLOS (Denker, 2016). One of the major driving forces behind the development of OA in the UK on a national level, the 'Finch Report', also concluded that OA was an essential source for information and innovation to the civil service, commercial sectors, small- and medium-sized enterprises (SMEs), and the general public (www.researchinfonet.org/publish/finch/).

Taking UK cancer research as one high impact case study, there is substantial evidence for the economic benefit of OA. In 2011–12 prices, the total expenditure on research relating to cancer in the period of 1970–2009 was £15 billion (Glover *et al.*, 2014). 5.9 million quality adjusted life years were gained from the prioritized interventions in 1991–2010, of which the net-monetary

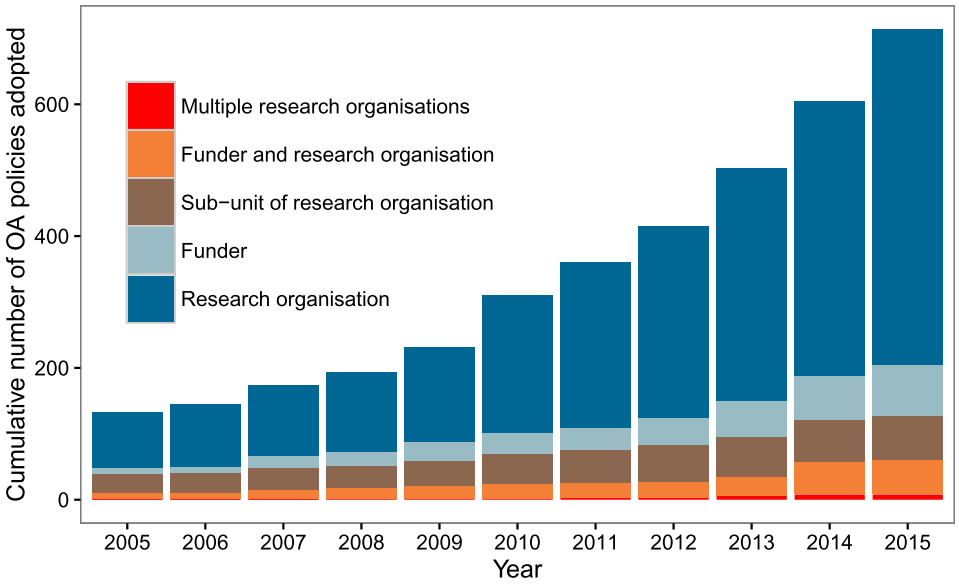

Figure 3. Cumulative number of Open Access policies adopted by multiple research organisations, funder and research organisation, sub-unit of research organisation, funder, and research organisation. Figures are given at the beginning of each year. Source: ROARMAP, accessed March 2016.

benefit was an estimated £124 billion (i.e., eight-fold return on investment). However, only 17% of the annual net-monetary benefit was estimated to be attributable to research performed in the UK (Glover *et al.*, 2014), suggesting that 83% of the economic return on cancer research is drawn from research from non-UK sources. Another example is from the area of environmental impact assessments, where Vickery (2011) has shown that OA to R&D results could result in recurring gains of around €6 billion per year. As such, opening up research for global access rather than localized and restricted use has the potential to increase the economic return, as demonstrated with the cases on cancer research and environmental impact assessments.

The price of Open Access

The question of the current publication cost is difficult and confounded by estimates of the total global publishing costs and revenue. Data provided by Outsell, a consultant in Burlingame, California, suggest that the science publishing industry generated $9.4 billion in revenue in 2011 and published around 1.8 million English-language articles. This equates to an approximate average revenue per article of $5,000. A white paper produced by the Max Planck Society estimated costs at €3,800–€5,000 per paper through subscription spending, based on a total global spending of €7.6 billion across 1.5–2 million articles per year in total (Schimmer *et al.*, 2015). Other estimates suggest that the total spending on publishing, distribution and access to research is around £25 billion per year, with an additional £34 billion spent on reading those outputs, a sum which equates to around one third of the total annual global spending on research (£175 billion; Research Information Network, 2008).

Such high costs are at odds with alternative estimates of the cost of OA publishing. For example, the Scientific Electronic Library Online (*SciELO*) is a pan-Latin American bibliographic database, digital library, and co-operative electronic publishing model of OA journals. It is estimated that their costs are between $70 and $600 per OA article depending on the services provided (Brembs, 2015). OA now dominates the Latin American publishing landscape, with an estimated 72–85% of articles now with full text OA articles publicly available (www.sparc.arl.org/news/open-access-latin-america-embraced-key-visibility-research-outputs). Furthermore, in countries such as Brazil, higher quality journals are more likely to be published OA (Neto *et al.*, 2016), implying that low-cost, high quality, and OA can all co-exist. Even more extreme estimates of the cost of OA come from Standard Analytics, who suggested the absolute minimum per-article costs of publishing could fall to between $1.36 and $1.61 with sufficient cloud-based infrastructure (Bogich *et al.*, 2016). However, it is likely that this estimate under-emphasizes marginal costs that are beyond a per-article cost basis. However, what is clear from these analyses is that OA has the opportunity to become a cost-reducing mechanism for scholarly publishing. Open Journals System (OJS), an open source software available for anyone to use and download without charge, is another example of this. Additionally, researcher-led initiatives such as the recently launched *Discrete Analysis* have costs that average around $30 per article, with no cost to authors or readers, and utilise the infrastructure offered by the arXiv to keep costs low (discreteanalysisjournal.com).

In her article, Sutton (2011) argued that current scholarly journals are digital products and that as such they are driven by very different economic principles and social forces than their print ancestors. Based on Anderson (2013b), the author made the case that changes in both the delivery of scientific content and in publishers' business models was inevitable when journals moved online. Sutton (2011) considered that scientific literature is no different from other digital products with respect to distribution costs and as such it is no exception to the 'zero is inevitable' rule of pricing.

The societal impact of Open Access

OA to the scholarly literature does not just benefit academics, but also has wider impacts on other domains in society. It makes research available to anyone with an Internet connection who has the ability to search and read the material. Therefore, it transcends academic affiliation and supports sustainable lifelong learning. Examples of groups who might benefit most from OA include citizen scientists, medical patients and their supporting networks, health advocates, NGOs, and those who benefit from translation and transformation (e.g., sight-impaired people). In theory, OA affects anyone who uses information, and opens up possibilities for knowledge to be used in unexpected, creative and innovative ways, far beyond the mainstream professional research.

Access to knowledge has been called a human rights issue, considering it is included in Article 27 of the United Nations Declaration of Human Rights. Willinsky (2006) has argued that "*Access to knowledge is a human right that is closely associated with the ability to defend, as well as to advocate for, other rights.*". This is not only true for access to knowledge from research that could save human lives, but also, as argued by Jacques Derrida, to the right of access to philosophy and the humanities disciplines that stem from it. Derrida writes about the field of Philosophy, "*No one can forbid access to it. The moment one has the desire or will for it, one has the right to it. The right is inscribed in philosophy itself* " (Derrida, 2002).

Society's ability to make research publicly accessible supports the long-term interest and investment in research. Citizens support research through taxes and therefore one could argue that efforts to support public access should be a fundamental part of the research process. While OA is not a solution to all aspects of research accessibility (e.g., filtering and language barriers, connectivity barriers and disability access remain continuing issues to be addressed; cyber.law.harvard.edu/hoap/Open_Access_(the_book)), it most certainly increases accessibility greatly and at the same time allows innovations to remove other barriers (e.g., OA articles can be freely translated to address language barriers and can be changed to different formats to accommodate screen readers). Anecdotal evidence suggests that public access to research is required from a range of public spheres (whoneedsaccess.org/). Nonetheless, the fact that access to knowledge continues to be prohibited in fields like public health should be of major concern to all stakeholders engaged in academic publishing.

In addition to professional research by, for example, academics, there is the dimension of citizen science. In citizen science, the

broader public participates in the research process itself and will have an increased interest in accessing previous research. Numerous projects such as Galaxy Zoo, Zooniverse, Old Weather, Fold It, Whale FM, Bat Detective, and Project Discovery, are all different initiatives in which citizens publicly and openly engage with research. These initiatives introduce new ways of knowledge creation and these groups also require thorough access to actually be able to do non-redundant research. Citizen science forms part of the societal case for OA, because it clearly indicates that anyone can be actively engaged with research, and not only professional scientists.

Some traditional publishers and some academics have argued that public access to research is not required because research papers cannot be understood by non-specialists (cyber.law.harvard.edu/hoap/Open_Access_(the_book) - see Section 5.5.1). However, citizen science initiatives already indicate the general public *is* interested in and understands the research. Whereas this understanding and engagement is highly variable, and strongly dependent on a range of extrinsic and intrinsic factors, the fact that a high level of public interest in science already exists is of relevance. These publishers and academics argue that specialization is a sufficient reason for confining access to professional research bodies through subscriptions. Such statements conflate a lack of desire or need for access with the denial of opportunity to access research, and makes false presumptions about the demand in access to the literature (i.e., unmet and unknown demand). Importantly, OA provides access to everyone who potentially needs or wants it, without making explicit and patronising statements or guesswork about who needs or deserves it. As Peter Suber says in his 2012 book: "The idea [of OA] is to stop thinking of knowledge as a commodity to meter out to deserving customers, and to start thinking of it as a public good, especially when it is given away by its authors, funded with public money, or both" (page 116). Isolated incidents such as the crashing of servers of Physical Reviews Letters upon the 'Gravitational Waves' announcement and OA publication (Feb, 2016; Abbott *et al.*, 2016) indicate that there are cases of extreme public interest in science that closed access would only impede. Moreover, one out of four people seeking medical information have hit a paywall at least once (pewinternet.org/2013/01/15/information-triage/). Claims that only experts can and should read research articles does little to break down the 'ivory tower' perception that still pervades academia, and undermines the enormous amounts of resources invested in science communication and public engagement activities. Such perceptions run counter to the idea of access to knowledge as a right, retaining it as a privilege based on financial or academic status.

Open Access in developing countries

The arguments outlined above form the basis for democratic and equal access to research, which come to light even stronger in the developing world. For low- and middle-income countries (LMIC), OA publishing breaks traditional financial barriers and allows unrestricted, equal access to scholarly information to people all over the globe. Due to the high prices of journal subscriptions, developing countries struggle with access just as in developed countries, but to a greater extent and consequently with greater negative repercussions. For example, a research paper from 1982 that indicated why Liberia should be included in the Ebola endemic zone was unknown to Liberian officials in the 2014 Ebola outbreak (Knobloch *et al.*, 1982); the paper was published behind a paywall, drastically reducing its discoverability. Even though the result is available in the abstract of the paywalled article, assessing the truth of the result certainly requires access to the full research article. In general, lack of access can have major deleterious consequences for students and researchers, in that they do not have sufficient material to conduct their own primary research or education.

OA provides a mechanism to level the playing field between developed and developing countries. This increases fair competition and the scientific potential of the developing world (Chan *et al.*, 2005). This aspect is linked to the wider issue of open licensing, which is essential for effective marketing of medicines and medical research in developing countries (Flynn *et al.*, 2009), and justifies the necessity of OA in the wider context of social welfare. Developing countries clearly acknowledge the need for access and as such have launched many repositories to increase access with self-archiving of research articles. In 2014, over 100 institutions in Africa launched a network of over 25 fully-operational OA repositories in Kenya, Tanzania and Uganda (www.ubuntunet.net/april2014#researchrelevant). Such developments suggest that African nations are leaning more towards a Green model of OA adoption.

The shift from a 'reader pays' to a pre-publication fee model (often conflated with 'author pays'; see subsection 'The effect on publishers') with OA potentially limits its adoption in developing countries. The pay-to-publish system is a potentially greater burden for authors in developing countries, considering that they are not used to paying publication costs, and funding systems for OA are not as well-established as those in the Western world. Publication fees present an even greater relative burden (Matheka *et al.*, 2014) given that they can often exceed a monthly salary for researchers. This has been at least partially mitigated with waiver fees for authors from developing countries and additional provisions in research grants, and around 70% of peer reviewed OA journals are fee-free. In November 2015, Research4Life (research4life.org) and DOAJ announced a working partnership that will help to ensure that the Research4Life users will have access to the largest possible array of OA journals from publishers with a certain quality standard. While Research4Life does not directly cover OA publication costs, a lot of publishers propose full or partial waivers if they are based in countries eligible by Research4Life. However, determining which countries qualify for access to scientific journals through these programs, and which journals they are provided access to, is a fairly closed process. They are also not entirely stable, as publishers can opt out of the initiative, or be selective about which countries they choose to serve. In 2011, publishers withdrew free access to 2500 health and biomedical journals for Bangladesh (Kmietowicz, 2011) through the HINARI programme. While access was subsequently reinstated, this demonstrates that such initiatives are not an adequate replacement for full OA (Chatterjee *et al.*, 2013). Despite these programs purporting to provide essential articles to researchers in poor nations, they exclude some developing countries (e.g., India) and limit access to researchers who work in registered institutions.

Initiatives such as the Journals Online Project developed by INASP (International Network for the Availability of Scientific Publications; inasp.info/en/) has helped to develop a number of online OA platforms in the Global South. These were launched in 1998 with the African Journals Online (AJOL) platform, a project currently managed in South Africa. More recently, INASP have set up Latin American Journals Online (LAMJOL) which hosts journals in El Salvador, Honduras, and Nicaragua. In Asia, Bangladesh Journals Online (BanglaJOL), Nepal Journals Online (NepJOL), and Sri Lankan Journals Online (SLJOL), all facilitated through INASP, continue to develop and now around 95% of their articles are full-text Open Access. As mentioned previously, improved access should not be limited to professional researchers only, considering that there is also global interest from the broader public, including health professionals.

Deceptive publishing practices

One negative effect of OA comes from entities that attempt to profit by exploiting the pay-to-publish system used by many OA publishers. These publishers operate a sub-category of OA journals known as vanity presses, predatory publishers (Beall, 2012) or pseudo-journals (McGlynn, 2013). These journals, referred to in this work as 'deceptive publishers', seem to be in the scholarly publishing business primarily to collect publication fees (i.e., APCs) in exchange for rapid publication without formal peer-review. Beall (2015) has defined a list of criteria for identifying deceptive publishers and an index of publishers and individual journals that meet these criteria is continuously updated (scholarlyoa.com).

While not all scholars and advocates agree with the criteria proposed by Jeffrey Beall (who controversially describes the OA movement as "an anti-corporatist movement that wants to deny the freedom of the press to companies it disagrees with" (Beall, 2013)), there are several factors that many agree on to identify a deceptive publisher, but these factors are not clear-cut indicators of deceptive publishing. One such indicator is that deceptive publishers tend to charge low publication fees (Xia, 2015), most below $100 and few charge more than $200. However, while this is a trait of almost all deceptive publishers, the reverse is not necessarily the case. For example, a single-authored paper with *PeerJ* would cost $99, but this is not a deceptive publisher. On the contrary, the average publication fee of journals indexed in the Directory of Open Access Journals (DOAJ) is around $900–$1,000 (Solomon & Björk, 2012) and leading universities in the UK and Germany pay on average $1,200–$1,300 per article (Schimmer *et al.*, 2015). The editorial and peer-review aspects of deceptive publishers are either non-existent or suspect; they also falsely claim to have ratings such as a Journal Impact Factor and to be indexed in major databases such as Scopus (Djuric, 2015). Editors from these journals solicit articles that have no relation to the topic of their journal and do not send the manuscripts out to be properly peer-reviewed (Bowman, 2014).

The problem of deceptive publishers in OA seems to highly affect countries where the academic evaluation strongly favors international publication without further quality checks (Shen & Björk, 2015). Xia *et al.* (2015) collected and analyzed the publication record, citation count, and geographic location of authors from the various groups of journals. Statistical analyses verified that deceptive and non-deceptive journals have distinct author populations: authors who publish in deceptive journals tend to be early-career researchers from developing countries with still little publishing experience. The spatial distribution of both the deceptive publishers and those authors who submit in pseudo-journals is highly skewed: Asia and Africa contributes three quarters of authors (Xia *et al.*, 2015) and Indian journals form the overwhelming proportion of deceptive publishers (Xia, 2015). An interesting finding is the very low involvement of South America, both among deceptive publishers (0.5%) and corresponding authors in deceptive journals (2.2%). The OA infrastructure in Latin America is different compared to other developing countries, which reveals a possible reason for this asymmetric situation. Latin American journals and universities are engaged in OA publication models at a higher degree than other regions (Alperin *et al.*, 2011). As a result, scholars from this region are not only more aware of OA issues, but they have more options for publishing OA than those from other regions (Alperin *et al.*, 2011). Moreover, SciELO (Packer, 2009) and the creation of Latin American databases (Alonso-Gamboa & Russell, 2012) have played a tremendous part in this process by bringing recognition and a good reputation to publishing outlets in Latin America.

Considerable attention is given to the subject of deceptive publishers, who have become conflated with the OA movement in general to the detriment of genuine OA publishers. For example, a 'sting' operation that outed bad peer-review instead got misinterpreted as bad peer-review in OA journals (Bohannon, 2013), but was probably more indicative of issues to do with the traditional closed and over-burdened system of peer review (scilogs.com/communication_breakdown/jon-tennant-oa/). Overall, the deceptive publisher phenomenon is one major negative aspect that spawns many misconceptions and misgivings about publishing OA. Recently launched industry-led initiatives such as "Think, Check, Submit" (thinkchecksubmit.org) provide a checklist to help researchers identify trustworthy journals, and will likely be a pivotal tool in combating deceptive publishers.

Open Access and Open Science

OA exists in a constantly evolving scholarly research ecosystem and the proliferation of "open" as a description of scientific activities has caused some confusion about what the term "open" means (for a more comprehensive discussion, see Pomerantz & Peek (2016)). As such, it is important to note how it is interconnected to other facets of the scholarly communication system. Here, we discuss the implications that the transition to OA has on developments in the broader context of Open Science (or Open Research).

Open Access and Open Data

The overall movement of OA has become conjoined with the drive for Open Data. Data sharing is fundamental to scientific progress, because data lead to the knowledge generated in research articles. Furthermore, data sharing has recently become a common requirement, together with OA, for both research funding and publication. The data sharing policy from PLOS illustrates the high degree of overlap between OA and Open Data; authors of articles published

in PLOS are required to share the data except if they have valid reasons not to (i.e., an opt-out system; journals.plos.org/plosone/s/data-availability). Many publishers, NGOs, and research funders have recently come together to commit to free research sharing in times of public health emergency, catalysed by the current Zika health threat (http://www.wellcome.ac.uk/About-us/Policy/Spotlight-issues/Data-sharing/Public-health-emergencies/index.htm). It is noteworthy that some of the largest publishers, including Wiley, Taylor and Francis, and Elsevier (with the exception of the journal *The Lancet*) did not commit to research sharing during ongoing or future public health crises (as of May, 2016).

The benefits of Open Data are diverse, including a citation advantage. Combined with the citation advantage for OA articles, providing data alongside publications can increase citations on average by 30% (Piwowar & Vision, 2013) and up to 69% (Piwowar *et al.*, 2007), but this evidence is entirely field-dependent (e.g., Dorch *et al.*, 2015)). Below we cover seven additional benefits of Open Data.

First, data sharing enhances reproducibility, a crucial aspect in a time where some scientific domains appear to have problems with reproducibility (e.g., Open Science Collaboration, 2015). Several factors could form the basis for this 'crisis', such as an overemphasis on novelty instead of rigour, selective reporting of results, an overemphasis on statistical significance, and insufficient documentation of the research methods. Publicly sharing data, code, and materials can certainly alleviate issues with reproducibility. This is especially pertinent in the modern sciences, where a substantial proportion of published results draw on quantitative experiments and computer simulations. As such, it is largely impossible to reproduce these experiments as they become more complex and the associated datasets increase in complexity. When full access to the data, metadata, and the code used to produce the ultimate results are provided alongside publication, this greatly improves reproducibility.

Second, publicly available data can be used to stimulate innovations, such as new analytical methods. An excellent example of this is provided by the neuroimaging OpenfMRI project, where shared data have been used to examine the effects of different processing pipelines on analysis outcomes (Carp, 2012) and test new methods to characterize different cognitive tasks (Turner & Laird, 2012). Another good example is the Protein Data Bank (PDB) (Berman *et al.*, 2000), a project which has enabled the re-use of the primary structural data and opened up new avenues of research, despite the latter not being expected.

Third, data sharing enables new research questions that can only be answered by combining datasets which now remain separated. Analyzing vast volumes of data can yield novel and perhaps surprising findings. This allows for integrated research hypotheses on the underlying processes behind the original data and observations. Exploratory approaches to large datasets can be seen as hypothesis generating tools, which later drives experimental testing to confirm or disprove these hypotheses (Wagenmakers *et al.*, 2012).

Fourth, the realization that data will ultimately be shared and visible to the community provides a strong incentive for researchers to ensure they engage in better data documentation and, therefore, research methods. For example, the willingness to publicly share data has been associated with fewer statistical errors in the final research article (Wicherts *et al.*, 2011).

Fifth, public data sharing provides a digital backup for datasets, protecting valuable scientific resources. Moreover, a considerable amount of data produced every day does not ultimately lead to publication and often remain hidden. Such data might remain in a hidden file-drawer despite being valid, creating a systematic bias in the information available. Public data sharing opens this file-drawer and, consequently, allows independent assessments of whether the data are valid or not.

Sixth, sharing data can certainly reduce the cost of performing research. A file-drawer has been indicated to greatly reduce the efficiency of research in detecting effects (van Assen *et al.*, 2014). Open Data, as such, discourages redundant data collection (i.e., data that have been already collected but never made publicly accessible) and simultaneously allows researchers to better approximate what is happening in their fields. This will have a large effect on research costs, resulting in savings that can be then be used for more productive research goals.

Finally, and tightly connected with the sixth point, Open Data potentially has a great economic value. For example, Open Data creates jobs for analysis and re-use of these data Capgemini (2015), and contributes to additional value of products and services in major sectors (Manyika *et al.*, 2013), ad well as benefits users of these data rich services (Stott, 2014).

Open Access and Open Science

Beyond OA and Open Data lies a more integrated approach to research, referred to more broadly as Open Science (i.e., Science 2.0, Open Scholarship). According to the European Commission's Horizon 2020 programme, Open Science is defined as "*The transformation, opening up and democratisation of science and research through ICT, with the objectives of making science more efficient, transparent and interdisciplinary, of changing the interaction between science and society, and of enabling broader societal impact and innovation*". Consequently, we see OA as only one of the multiple challenges currently facing the 'open transformation' of the scholarly publishing system (Watson, 2015), and OA should therefore be considered in the wider contexts and complimentary domains of research transparency and open source.

As Kriegeskorte *et al.* (2012) pointed out, OA is now widely accepted as desirable and becoming a reality in many academic spheres. However, the second essential complementary element to research, evaluation, has received less attention despite the large amount of research that has been done to document its current limitations (Benos *et al.*, 2007; Birukou *et al.*, 2011; Ioannidis, 2005, Ioannidis 2012a; Ioannidis 2012b; John *et al.*, 2012; Nosek & Bar-Anan, 2012; Simmons *et al.*, 2011).

Open evaluation, an ongoing post-publication process of transparent peer review and rating of papers, promises to address the problems of the current assessment systems Kriegeskorte *et al.* (2012),

as well as increasing the overall quality of the peer review process. As such, ongoing assessments of the development of OA must also consider the broader impact and concurrent changes to the peer review system (Leek *et al.*, 2011; van Rooyen *et al.*, 1999; Wicherts, 2016). Some assessment methods, such as the Research Excellence Framework (REF) in England and administered by HEFCE, have already made OA a core feature of evaluation in that all research papers submitted to the REF must be archived in an institutional or subject repository (www.hefce.ac.uk/pubs/year/2014/201407/). While it is too early to evaluate the impact of this policy, by tying OA compliance with research evaluation we might expect to see a national shift towards large-scale OA adoption. At the very least, such a combination is generating increasing interest and awareness about OA among researchers, increasing usage of institutional repositories, and increasing demand for funding for APCs (Tate, 2015).

Future research regarding better ways to improve scholarly communication will be instrumental in providing evidence to support the transformation of the publishing system and design new alternatives (Buttliere, 2014; Ghosh *et al.*, 2012; Kriegeskorte *et al.*, 2012; Pöschl, 2012), which will draw heavily upon on open publishing framework driven by developments and newly emerging models in OA. Finally, consideration of Open Science and OA will be important inclusions in evolving research standards such as the Transparency and Openness Promotion (TOP) guidelines (https://cos.io/top/).

Conclusions

This article provides an evidence-based review of the impact of OA on academy, economy and society. Overall, the evidence points to a favorable impact of OA on the scholarly literature through increased dissemination and re-use. OA has the potential to be a sustainable business venture for new and established publishers, and can provide substantial benefits to research- and development-intensive businesses, including health organisations, volunteer sectors, and technology. OA is a global issue, highlighted by inequalities between developing and developed nations, and largely fueled by financial disparity. Current levels of access in the developing world are insufficient and unstable, and OA has the potential to foster the development of stable research ecosystems. While deceptive publishing remains an ongoing issue, particularly in the developing world, increasing public engagement, development of OA policies, and discussion of sustainable and ethical publishing practices can remove this potential threat.

For libraries, universities, governments, and research institutions, one major benefit of lowering the cost of knowledge is the availability of extra budget that can be reallocated for other purposes. For researchers themselves, OA can increase their audience and impact by delivering wider and easier access for readers. For publishers, promoting OA is an answer to the desires and the needs of their research communities. Furthermore, subscription-based publishers have (partly) answered the call of the increasing global demand for OA, by giving their green light to author self-archiving (Harnad *et al.*, 2008), as well as by establishing numerous 'hybrid' OA options. In an author survey, Swan & Brown (2004) reported that the vast majority of their sample indicated that they would self-archive willingly if their employer (or funding body) required them to do so. Similarly, in a study by Swan & Brown (2005) the vast majority of researchers (81%) indicated that they would comply with mandates that made OA a condition of funding or employment. There is evidence that many funders and research organisations are moving in this direction: since 2005, the number of policies supporting OA publishing increased steadily, and there is similar growth in the number of institutional rights-retention policies. Consequently, it is now the responsibility of researchers to ensure OA to their publications either by choosing the Green or the Gold road, and for public research funders to employ policies that are in the best interests of the wider public while considering the financial sustainability of the scholarly publishing ecosystem.

The fact that OA impacts upon such a diverse range of stakeholders, often with highly polarised and emotional viewpoints, highlights the ongoing need for evidence-informed discussion and engagement at all levels. This is especially the case for research communities, who have exceptionally diverse perspectives about OA and in particular how it interacts with 'quality' and 'prestige' in publishing (Schroter & Tite, 2006; Schroter *et al.*, 2005). As Peter Suber, a leading voice in the OA movement, stated (dash.harvard.edu/handle/1/4391169).

> *"TA [toll-access] publishers are not the enemy. They are only unpersuaded. Even when they are opposed, and not merely unpersuaded, they are only enemies if they have the power to stop OA. No publisher has this power, or at least not by virtue of publishing under a TA business model. If we have enemies, they are those who can obstruct progress to OA. The only people who fit this description are friends of OA who are distracted from providing OA by other work or other priorities."*

Therefore, OA supporters should focus their efforts on working for new models and systems rather than trying to undermine or punish the existing ones. OA remains only one of the multiple challenges that the scholarly publishing system is currently facing. As highlighted in this review, the empirical evidence for OA is overwhelmingly positive, but further research is certainly required to move from investigating the effects of OA to researching the broader effects of Open Science. In particular, OA must be considered in the future to more broadly regarding the adverse effects of a system of journal-based research assessment (Brembs *et al.*, 2013), and the development of scholarly communication systems that are sustainable for, and in the best interests of, the commons.

Author contributions

All authors contributed equally to the writing of this manuscript using the Overleaf collaborative writing platform.

Competing interests

JPT currently blogs for the PLOS Paleo Community, and works for ScienceOpen. CHJH is a Center for Open Science ambassador. DCJ and FW are members of the Open Access Working Group of EURODOC. PM is a Research Data Alliance member. LBC

works for the University of Pittsburgh, which has an Open Access library publishing department. All views presented here are strictly personal.

Grant information

This research was partly funded by the Belgian National Fund for Scientific Research through a FRIA grant. PM acknowledges support from the European Commission Horizon 2020 Programme under Grant Agreement 634107 (PHC32-2014) 'MULTIMOT'.

The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Acknowledgments

We would like to collectively acknowledge the OpenCon community for inspiring this paper, and for providing continuous discussion about the various aspects of Open Access. In particular, we are grateful to Brett Buttliere, Audrey Risser, Sarah Barkla, and April Clyburne-Sherin for contributing resources to the development of this paper, and Tracey Depellegrin Connelly, Matt Menzenski, and Joseph McArthur for helpful comments on an earlier draft. We also thank Neil Saunders who provided the base code to extract data from PubMed Central. We would also like to thank Andy Nobes for drawing our attention to the work of INASP. PM would like to thank Lennart Martens for insightful discussions on Open Science. For comments on the first published version of this manuscript, we would like to thank Philip Young, Ross Mounce, Anna Sharman, and David Wojick for their helpful comments. Reviews from Gwilym Lockwood, Peter Suber, Paige Brown Jarreau, Anne Tierney and Chris Chambers were supportive, constructive, and greatly improved the content and balance of this article.

References

Abbott BP, Abbott R, Abbott TD, *et al.*: **Observation of Gravitational Waves from a Binary Black Hole Merger.** *Phys Rev Lett.* 2016; **116**(6): 061102.
PubMed Abstract | Publisher Full Text

Adcock J, Fottrell E: **The North-South information highway: case studies of publication access among health researchers in resource-poor countries.** *Glob Health Action.* 2008; **1**.
PubMed Abstract | Publisher Full Text | Free Full Text

Adie E: **Attention! a study of open access vs non-open access articles.** *Figshare.* 2014.
Publisher Full Text

Allen HG, Stanton TR, Di Pietro F, *et al.*: **Social media release increases dissemination of original articles in the clinical pain sciences.** *PLoS One.* 2013; **8**(7): e68914.
PubMed Abstract | Publisher Full Text | Free Full Text

Alonso-Gamboa JO, Russell JM: **Latin American scholarly journal databases: a look back to the way forward.** In *Aslib Proceedings.* Emerald Group Publishing Limited, 2012; **64**(1): 32–45.
Publisher Full Text

Alperin JP, Fischman GE, Willinsky J: **Scholarly communication strategies in Latin America's research-intensive universities.** *Educación superior y sociedad.* 2011; **16**(2).
Reference Source

Aman V: **Is there any measurable benefit in publishing preprints in the arxiv section quantitative biology?** *CoRR.* 2014; abs/1411.1955.
Reference Source

Anderson B: **Funding sources of impactful and transformative research.** Master's thesis, San Jose State University, 2013a.
Reference Source

Anderson C: **Free: How today's smartest businesses profit by giving something for nothing.** Random House, 2013b.
Reference Source

Antelman K: **Do open-access articles have a greater research impact?** *Coll Res Libr.* 2004; **65**(5): 372–382.
Publisher Full Text

Arzberger P, Schroeder P, Beaulieu A, *et al.*: **Promoting access to public research data for scientific, economic, and social development.** *Data Sci J.* 2004; **3**: 135–152.
Publisher Full Text

Atchison A, Bull J: **Will open access get me cited? an analysis of the efficacy of open access publishing in political science.** *PS Polit Sci Polit.* 2015; **48**(01): 129–137.
Publisher Full Text

Beagrie N, Houghton JW: **The value and impact of data sharing and curation: A synthesis of three recent studies of UK research data centres**. 2014.
Reference Source

Beall J: **Criteria for determining predatory open-access publishers.** *Scholarly Open Access.* 2015. (accessed 2015-02-14).
Reference Source

Beall J: **Predatory publishers are corrupting open access.** *Nature.* 2012; **489**(7415): 179.
PubMed Abstract | Publisher Full Text

Beall J: **The open-access movement is not really about open access.** tripleC: Communication, Capitalism and Critique. *Open Access Journal for a Global Sustainable Information Society.* 2013; **11**(2): 589–597.
Reference Source

Benos DJ, Bashari E, Chaves JM, *et al.*: **The ups and downs of peer review.** *Adv Physiol Educ.* 2007; **31**(2): 145–152.
PubMed Abstract | Publisher Full Text

Bergstrom TC, Courant PN, McAfee RP, *et al.*: **Evaluating big deal journal bundles.** *Proc Natl Acad Sci U S A.* 2014; **111**(26): 9425–9430.
PubMed Abstract | Publisher Full Text | Free Full Text

Berman HM, Westbrook J, Feng Z, *et al.*: **The Protein Data Bank.** *Nucleic Acids Res.* 2000; **28**(1): 235–242.
PubMed Abstract | Publisher Full Text | Free Full Text

Berners-Lee T, De Roure D, Harnad S, *et al.*: **Journal publishing and author self-archiving: Peaceful co-existence and fruitful collaboration.** Commentary On: **http://www.alpsp.org/news/rcuk/default.htm**. 2005.
Reference Source

Bernius S, Hanauske M, Dugall B, *et al.*: **Exploring the effects of a transition to open access: Insights from a simulation study.** *J Am Soc Inf Sci Technol.* 2013; **64**(4): 701–726.
Publisher Full Text

Birukou A, Wakeling JR, Bartolini C, *et al.*: **Alternatives to peer review: novel approaches for research evaluation.** *Front Comput Neurosci.* 2011; **5**: 56.
PubMed Abstract | Publisher Full Text | Free Full Text

Björk BC, Laakso M, Welling P, *et al.*: **Anatomy of green open access.** *J Assoc Inf Sci Technol.* 2014; **65**(2): 237–250.
Publisher Full Text

Bjork BC, Roos A, Lauri M: **Scientific journal publishing: yearly volume and open access availability.** *Inform Res.* 2009; **14**(1).
Reference Source

Bloudoff-Indelicato M: **Text-mining block prompts online response.** *Nature News.* 2015; **527**(7579): 413.
Publisher Full Text

Bogich T, Ballesteros S, Berjon R, *et al.*: **On the marginal cost of scholarly communication**. 2016. Accessed: 2016-3-24.
Reference Source

Bohannon J: **Who's afraid of peer review?** *Science.* 2013; **342**(6154): 60–5.
PubMed Abstract | Publisher Full Text

Bohannon J: **Who's downloading pirated papers? everyone**. 2016; Accessed: 2016-5-4.
Reference Source

Bowman JD: **Predatory publishing, questionable peer review, and fraudulent conferences.** *Am J Pharm Educ.* 2014; **78**(10): 176.
PubMed Abstract | Publisher Full Text | Free Full Text

Brembs B, Button K, Munafò M: **Deep impact: unintended consequences of journal rank.** *Front Hum Neurosci.* 2013; **7**: 291.
PubMed Abstract | Publisher Full Text | Free Full Text

Brembs B: **What goes into making a scientific manuscript public**? 2015; Accessed: 2016-3-24.
Reference Source

Buttliere BT: **Using science and psychology to improve the dissemination and evaluation of scientific work.** *Front Comput Neurosci.* 2014; **8**: 82.
PubMed Abstract | Publisher Full Text | Free Full Text

Capgemini: **Creating value through open data: Study on the impact of re-use of public data resources.** *Digital Agenda for Europe.* 2015.
Reference Source

Carp J: **On the plurality of (methodological) worlds: estimating the analytic flexibility of fMRI experiments.** *Front Neurosci.* 2012; **6**: 149.
PubMed Abstract | Publisher Full Text | Free Full Text

Carr L, Harnad S, Swan A: **A longitudinal study of the practice of self-archiving.** *University of Southampton Working Paper.* 2007.
Reference Source

Chamberlain S, Boettiger C, Ram K: **rplos: Interface to the Search 'API' for 'PLoS' Journals**. 2016.
Reference Source

Chan L, Kirsop B, Arunachalam S: **Open access archiving: the fast track to building research capacity in developing countries**. 2005.
Reference Source

Chatterjee P, Biswas T, Mishra V: **Open access: the changing face of scientific publishing.** *J Family Med Prim Care.* 2013; **2**(2): 128–30.
PubMed Abstract | Publisher Full Text | Free Full Text

Cheng W, Ren S: **Evolution of open access publishing in Chinese scientific journals.** *Learn Publ.* 2008; **21**(2): 140–152.
Publisher Full Text

Colomb J, Brembs B: **Sub-strains of *Drosophila* Canton-S differ markedly in their locomotor behavior [version 2; referees: 3 approved].** *F1000Res.* 2015; **3**: 176.
Publisher Full Text

Costas R, Zahedi Z, Wouters P: **Do "altmetrics" correlate with citations? extensive comparison of altmetric indicators with citations from a multidisciplinary perspective.** *J Assoc Inf Sci Technol.* 2015; **66**: 2003–2019.
Publisher Full Text

Crawford W: **Open Access Journals 2014, DOAJ subset**. 2015.
Publisher Full Text

Dallmeier-Tiessen S, Darby R, Goerner B, *et al.*: **Highlights from the soap project survey. What scientists think about open access publishing**. arXiv preprint arXiv:1101.5260. 2011.
Reference Source

Davis P, Fromerth M: **Does the arXiv lead to higher citations and reduced publisher downloads for mathematics articles?** *Scientometrics.* 2007; **71**(2): 203–215.
Publisher Full Text

Davis PM, Lewenstein BV, Simon DH, *et al.*: **Open access publishing, article downloads, and citations: randomised controlled trial.** *BMJ.* 2008; **337**: a568.
PubMed Abstract | Publisher Full Text | Free Full Text

Davis PM: **Open access, readership, citations: a randomized controlled trial of scientific journal publishing.** *FASEB J.* 2011; **25**(7): 2129–2134.
PubMed Abstract | Publisher Full Text

Denicola R: **Copyright and open access: reconsidering university ownership of faculty research.** *Nebraska Law Review.* 2006; **85**(2).
Reference Source

Denker SP: **Collaboration with IBM watson supports the value add of open access**. 2016.
Reference Source

Derrida J: **Who's afraid of philosophy?: Right to philosophy 1.** Stanford University Press, 2002; **1**.
Reference Source

Djuric D: **Penetrating the omerta of predatory publishing: the Romanian connection.** *Sci Eng Ethics.* 2015; **21**(1): 183–202.
PubMed Abstract | Publisher Full Text

Dorch SBF, Drachen TM, Ellegaard O: **The data sharing advantage in Astrophysics**. arXiv, 2015.
Reference Source

EFF: **Big win for fair use in google books lawsuit**. 2015; Accessed: 2016-4-26.
Reference Source

Eger T, Scheufen M: **The past and the future of copyright law: Technological change and beyond**. *Liber Amicorum Boudewijn Bouckaert*, forthcoming. 2012; 37–64.
Reference Source

Elbakyan A, Bohannon J: **Data from: Who's downloading pirated papers? everyone.** *Dryad Digital Repository.* 2016.
Publisher Full Text

European Commission: **Communication from the commission to the european parliament, the council, the european economic and social committee, and the committee of the regions**. Towards better access to scientific information: Boosting the benefits of public investments in research. 2012.
Reference Source

Evans JA, Reimer J: **Open access and global participation in science.** *Science.* 2009; **323**(5917): 1025.
PubMed Abstract | Publisher Full Text

Eysenbach G: **Can tweets predict citations? Metrics of social impact based on Twitter and correlation with traditional metrics of scientific impact.** *J Med Internet Res.* 2011; **13**(4): e123.
PubMed Abstract | Publisher Full Text | Free Full Text

Eysenbach G: **Citation advantage of open access articles.** *PLoS Biol.* 2006; **4**(5): e157.
PubMed Abstract | Publisher Full Text | Free Full Text

Flynn S, Hollis A, Palmedo M: **An economic justification for open access to essential medicine patents in developing countries.** *J Law Med Ethics.* 2009; **37**(2): 184–208.
PubMed Abstract | Publisher Full Text

Frandsen TF: **The integration of open access journals in the scholarly communication system: Three science fields.** *Inf Process Manag.* 2009; **45**(1): 131–141.
Publisher Full Text

Gargouri Y, Hajjem C, Larivière V, *et al.*: **Self-selected or mandated, open access increases citation impact for higher quality research.** *PLoS One.* 2010; **5**(10): e13636.
PubMed Abstract | Publisher Full Text | Free Full Text

Gargouri Y, Larivière V, Gingras Y, *et al.*: **Green and gold open access percentages and growth, by discipline.** *arXiv preprint arXiv:1206.3664.* 2012.
Reference Source

Gaule P, Maystre N: **Getting cited: does open access help?** *Res Policy.* 2011; **40**(10): 1332–1338.
Publisher Full Text

Gentil-Beccot A, Mele S, Brooks T: **Citing and reading behaviours in high-energy physics.** *Scientometrics.* 2010; **84**(2): 345–355.
Publisher Full Text

Ghosh SS, Klein A, Avants B, *et al.*: **Learning from open source software projects to improve scientific review.** *Front Comput Neurosci.* 2012; **6**: 18.
PubMed Abstract | Publisher Full Text | Free Full Text

Glenisson P, Glänzel W, Janssens F, *et al.*: **Combining full text and bibliometric information in mapping scientific disciplines.** *Inf Process Manag.* 2005; **41**(6): 1548–1572.
Publisher Full Text

Glover M, Buxton M, Guthrie S, *et al.*: **Estimating the returns to UK publicly funded cancer-related research in terms of the net value of improved health outcomes.** *BMC Med.* 2014; **12**(1): 99.
PubMed Abstract | Publisher Full Text | Free Full Text

Gonzalez GH, Tahsin T, Goodale BC, *et al.*: **Recent Advances and Emerging Applications in Text and Data Mining for Biomedical Discovery.** *Brief Bioinform.* 2016; **17**(1): 33–42.
PubMed Abstract | Publisher Full Text | Free Full Text

Hajjem C, Harnad S, Gingras Y: **Ten-year cross-disciplinary comparison of the growth of open access and how it increases research citation impact**. 2006; arXiv preprint cs/0606079.
Reference Source

Handke C, Guibault L, Vallbé JJ: **Is Europe falling behind in data mining? Copyright's impact on data mining in academic research**. 2015.
Publisher Full Text

Harmston N, Filsell W, Stumpf MP: **What the papers say: text mining for genomics and systems biology.** *Hum Genomics.* 2010; **5**(1): 17–29.
PubMed Abstract | Publisher Full Text | Free Full Text

Harnad S, Brody T, Vallieres F, *et al.*: **The access/impact problem and the green and gold roads to open access: An update.** *Serials Rev.* 2008; **34**(1): 36–40.
Publisher Full Text

Harnad S, Brody T: **Comparing the Impact of Open Access (OA) vs. Non-OA Articles in the Same Journals.** *D-lib Magazine.* 2004; **10**(6).
Publisher Full Text

Harnad S: **Opening access by overcoming zeno's paralysis**. 2006.
Reference Source

Henneken EA, Kurtz MJ, Eichhorn G, *et al.*: **Effect of E-Printing on Citation Rates in Astronomy and Physics.** *arXiv.* 2006.
Reference Source

Hitchcock S: **The effect of open access and downloads ('hits') on citation impact: a bibliography of studies**. 2013.
Reference Source

Hoorn E, van der Graaf M: **Copyright issues in open access research journals: The authors perspective.** *D-Lib Magazine.* 2006; **12**(2): 6.
Publisher Full Text

Houghton JW, Oppenheim C: **The economic implications of alternative publishing models.** *Prometheus.* 2010; **28**(1): 41–54.
Publisher Full Text

Ioannidis JP: **Scientific communication is down at the moment, please check again later.** *Psychol Inq.* 2012b; **23**(3): 267–270.
Publisher Full Text

Ioannidis JP: **Why most published research findings are false.** *PLoS Med.* 2005; **2**(8): e124.
PubMed Abstract | Publisher Full Text | Free Full Text

Ioannidis JP: **Why Science Is Not Necessarily Self-Correcting.** *Perspect Psychol Sci.* 2012a: **7**(6): 645–654.
PubMed Abstract | Publisher Full Text

John LK, Loewenstein G, Prelec D: **Measuring the prevalence of questionable research practices with incentives for truth telling.** *Psychol Sci.* 2012; **23**(5): 524–32.
PubMed Abstract | Publisher Full Text

Khabsa M, Giles CL: **The number of scholarly documents on the public web.** *PLoS One.* 2014; **9**(5): e93949.
PubMed Abstract | Publisher Full Text | Free Full Text

Kmietowicz Z: **Publishers withdraw 2500 journals from free access scheme in Bangladesh.** *BMJ.* 2011; **342**: d196.
Reference Source

Knobloch J, Albiez EJ, Schmitz H: **A serological survey on viral haemorrhagic fevers in Liberia.** In *Annales de l'Institut Pasteur/Virologie.* Elsevier. 1982; **133**: 125–128.
Publisher Full Text

Kousha K, Abdoli M: **The citation impact of Open Access agricultural research: A comparison between OA and non-OA publications.** *Online Inform Rev.* 2010; **34**(5): 772–785.
Publisher Full Text

Kriegeskorte N, Walther A, Deca D: **An emerging consensus for open evaluation: 18 visions for the future of scientific publishing.** *Front Comput Neurosci.* 2012; **6**: 94.
PubMed Abstract | Publisher Full Text | Free Full Text

Kurtz MJ, Eichhorn G, Accomazzi A, *et al.*: **The effect of use and access on citations.** *Inf Process Manag.* 2005; **41**(6): 1395–1402.
Publisher Full Text

Kurtz MJ, Henneken EA: **Open Access does not increase citations for research articles from The Astrophysical Journal.** *arXiv.* 2007.
Reference Source

Laakso M, Björk BC: **Anatomy of open access publishing: a study of longitudinal development and internal structure.** *BMC Med.* 2012; **10**(1): 124. ISSN 1741-7015.
PubMed Abstract | Publisher Full Text | Free Full Text

Laakso M, Welling P, Bukvova H, *et al.*: **The development of open access journal publishing from 1993 to 2009.** *PLoS One.* 2011; **6**(6): e20961.
PubMed Abstract | Publisher Full Text | Free Full Text

Lansingh VC, Carter MJ: **Does open access in ophthalmology affect how articles are subsequently cited in research?** *Ophthalmology.* 2009; **116**(8): 1425–1431.
PubMed Abstract | Publisher Full Text

Larivière V, Haustein S, Mongeon P: **The Oligopoly of Academic Publishers in the Digital Era.** *PLoS One.* 2015; **10**(6): e0127502.
PubMed Abstract | Publisher Full Text | Free Full Text

Lawrence S: **Online or invisible?** *Nature.* 2001; **411**(6837): 521.
Reference Source

Lawson S, Meghreblian B: **Journal subscription expenditure of UK higher education institutions [version 3; referees: 4 approved].** *F1000Res.* 2015; **3**(274).
PubMed Abstract | Publisher Full Text | Free Full Text

Lawson S: **Fee waivers for open access journals.** *Publications.* 2015; **3**(3): 155–167.
Publisher Full Text

Leek JT, Taub MA, Pineda FJ: **Cooperation between referees and authors increases peer review accuracy.** *PLoS One.* 2011; **6**(11): e26895.
PubMed Abstract | Publisher Full Text | Free Full Text

Leitner F, Valencia A: **A text-mining perspective on the requirements for electronically annotated abstracts.** *FEBS Lett.* 2008; **582**(8): 1178–81.
PubMed Abstract | Publisher Full Text

Liang X, Su LY, Yeo SK, *et al.*: **Building buzz (scientists) communicating science in new media environments.** *Journal Mass Commun Q.* 2014; **91**(4): 772–791.
Publisher Full Text

MacKie-Mason J: **Economic thoughts about "gold" open access**. 2016; Accessed: 2016-4-27.
Reference Source

Malhotra A, Younesi E, Gurulingappa H, *et al.*: **'HypothesisFinder:' a strategy for the detection of speculative statements in scientific text.** *PLoS Comput Biol.* 2013; **9**(7): e1003117.
PubMed Abstract | Publisher Full Text | Free Full Text

Manyika J, Chui M, Groves P, *et al.*: **Open data: Unlocking innovation and performance with liquid information.** *McKinsey global institute. McKinsey Center for Government and McKinsey Business Technology Office.* 2013.
Reference Source

Martone M, Murray-Rust P, Molloy J, *et al.*: **Contentmine/hypothes.is proposal.** *Research Ideas and Outcomes.* 2016; **2**: e8424.
Publisher Full Text

Matheka DM, Nderitu J, Mutonga D, *et al.*: **Open access: academic publishing and its implications for knowledge equity in kenya.** *Global Health.* 2014; **10**(1): 26.
PubMed Abstract | Publisher Full Text | Free Full Text

Mazzucato M: **The entrepreneurial state.** *Soundings.* 2011; **49**(49): 131–142.
Publisher Full Text

McCabe M, Snyder CM: **Identifying the effect of open access on citations using a panel of science journals.** *Econ Inq.* 2014; **52**(4): 1284–1300.
Publisher Full Text

McGlynn T: **The evolution of pseudojournals**. *Small Pond Science.* Dominguez Hills, CA, 2013.
Reference Source

McGuigan GS, D. Russel: **The business of academic publishing: A strategic analysis of the academic journal publishing industry and its impact on the future of scholarly publishing.** *Electron Journal of Academic and Special Librarianship.* 2008; **9**(3).
Reference Source

McKiernan E, Bourne PE, Brown CT, *et al.*: **The open research value proposition: How sharing can help researchers succeed.** *Figshare.* 2016; **1**.
Publisher Full Text

McVeigh ME: **Open access journals in the ISI citation databases: analysis of impact factors and citation patterns: a citation study from Thomson Scientific**. Thomson Scientific. 2004.
Reference Source

Metcalfe TS: **The citation impact of digital preprint archives for solar physics papers.** *Sol Phys.* 2006; **239**(1–2): 549–553.
Publisher Full Text

Metcalfe TS: **The rise and citation impact of astroph in major journals**. *arXiv preprint astro-ph/0503519.* 2005.
Reference Source

Moed H: **The effect of "open access" upon citation impact: an analysis of arxiv's condensed matter section**. 2006.
Reference Source

Mohammadi E, Thelwall M, Haustein S, *et al.*: **Who reads research articles? an altmetrics analysis of Mendeley user categories.** *J Assoc Inf Sci Technol.* 2015; **66**(9): 1832–1846.
Publisher Full Text

Müller-Langer F, Watt R: **Copyright and open access for academic works**. *Review of Economic Research on Copyright Issues.* 2010; **7**(1): 45–65.
Reference Source

Natarajan J, Berrar D, Dubitzky W, *et al.*: **Text mining of full-text journal articles combined with gene expression analysis reveals a relationship between sphingosine-1-phosphate and invasiveness of a glioblastoma cell line.** *BMC Bioinformatics.* 2006; **7**(1): 373.
PubMed Abstract | Publisher Full Text | Free Full Text

Neto SC, Willinsky J, Alperin J: **Measuring, rating, supporting, and strengthening open access scholarly publishing in brazil.** *Educ Policy Anal Arch.* 2016; **24**(0): 54. ISSN 1068-2341.
Publisher Full Text

Norris M, Oppenheim C, Rowland F: **Open access citation rates and developing countries.** In *ELPUB.* 2008; 335–342.
Reference Source

Nosek BA, Bar-Anan Y: **Scientific utopia: I. opening scientific communication.** *Psychol Inq.* 2012; **23**(3): 217–243.
Publisher Full Text

Nuijten MB, Hartgerink CH, van Assen MA, *et al.*: **The prevalence of statistical reporting errors in psychology (1985–2013).** *Behav Res Methods.* 2015; 1–22. ISSN 1554-351X, 1554-3528.
PubMed Abstract | Publisher Full Text

Odlyzko A: **Economic costs of toll access.** *Open Access: Key Strategic, Technical and Economic Aspects.* 2006; **4**: 39–43.
Publisher Full Text

Odlyzko AM: **Open Access, library and publisher competition, and the evolution of general commerce.** *CoRR.* abs/1302.1105, 2013.
Reference Source

Okerson A: **With feathers: Effects of copyright and ownership on scholarly publishing.** *Coll Res Libr.* 1991; **52**(5): 425–38.
Publisher Full Text

Open Access NL: **Amendment to copyright act**. 2015; Accessed: 2016-4-27.
Reference Source

Open Science Collaboration: **PSYCHOLOGY. Estimating the reproducibility of psychological science.** *Science.* 2015; **349**(6251): aac4716.
PubMed Abstract | Publisher Full Text

Packer AL: **The SciELO Open Access: A Gold Way from the South.** *Can J High Educ.* 2009; **39**(3): 111–126.
Reference Source

Piwowar H: **Altmetrics: Value all research products.** *Nature.* 2013; **493**(7431): 159.
PubMed Abstract | Publisher Full Text

Piwowar HA, Day RS, Fridsma DB: **Sharing detailed research data is associated with increased citation rate.** *PLoS One.* 2007; **2**(3): e308.
PubMed Abstract | Publisher Full Text | Free Full Text

Piwowar HA, Vision TJ: **Data reuse and the open data citation advantage.** *PeerJ.* 2013; **1**: e175. ISSN 2167-8359.
PubMed Abstract | Publisher Full Text | Free Full Text

Pomerantz J, Peek R: **Fifty shades of open.** *First Monday.* 2016; **21**(5): ISSN 13960466.
Publisher Full Text

Porter AL, Kongthon A, Lu JC: **Research profiling: Improving the literature review.** *Scientometrics.* 2002; **53**(3): 351–370.
Publisher Full Text

Pöschl U: **Multi-stage open peer review: scientific evaluation integrating the strengths of traditional peer review with the virtues of transparency and self-regulation.** *Front Comput Neurosci.* 2012; **6**: 33.
PubMed Abstract | Publisher Full Text | Free Full Text

Research Information Network: **Activities, costs and funding flows in the scholarly communications system in the UK.** Technical report, 2008.
Reference Source

Rosenwald MS: **This student put 50 million stolen research articles online. And they're free**. 2016; Accessed: 2016-4-26.
Reference Source

Sahu DK, Gogtay NJ, Bavdekar SB: **Effect of open access on citation rates for a small biomedical journal**. *Fifth International Congress on Peer Review and Biomedical Publication, Chicago, September 16-1*8. 2005.
Reference Source

Satyanarayana K: **Journal publishing: the changing landscape.** *Indian J Med Res.* 2013; **138**(1): 4–7.
PubMed Abstract | Free Full Text

Scheufen M: **Copyright Versus Open Access: On the Organisation and International Political Economy of Access to Scientific Knowledge.** Springer, 2015.
Publisher Full Text

Schimmer R, Geschuhn KK, Vogler A: **Disrupting the subscription journals' business model for the necessary large-scale transformation to open access**. 2015.
Publisher Full Text

Schimmer R, Geschuhn KK, Vogler A: **Disrupting the subscription journals' business model for the necessary large-scale transformation to open access**. 2015.
Publisher Full Text

Schroter S, Tite L, Smith R: **Perceptions of open access publishing: interviews with journal authors.** *BMJ.* 2005; **330**(7494): 756.
PubMed Abstract | Publisher Full Text | Free Full Text

Schroter S, Tite L: **Open access publishing and author-pays business models: a survey of authors' knowledge and perceptions.** *J R Soc Med.* 2006; **99**(3): 141–148.
PubMed Abstract | Publisher Full Text | Free Full Text

Schwarz GJ, Kennicutt RC Jr: **Demographic and citation trends in astrophysical journal papers and preprints**. arXiv preprint astro-ph/0411275, 2004.
Reference Source

Shatkay H, Pan F, Rzhetsky A, *et al.*: **Multi-dimensional classification of biomedical text: toward automated, practical provision of high-utility text to diverse users.** *Bioinformatics.* 2008; **24**(18): 2086–2093.
PubMed Abstract | Publisher Full Text | Free Full Text

Shavell S: **Should copyright of academic works be abolished?** *J Legal Analysis.* 2010; **2**(1): 301–358.
Publisher Full Text

Shen C, Björk BC: **'predatory' open access: a longitudinal study of article volumes and market characteristics.** *BMC Med.* 2015; **13**(1): 230.
PubMed Abstract | Publisher Full Text | Free Full Text

Shieber S, Suber P: **Good practices for university open-access policies**. 2016; Accessed: 2016-9-12.
Reference Source

Simmons JP, Nelson LD, Simonsohn U: **False-positive psychology: undisclosed flexibility in data collection and analysis allows presenting anything as significant.** *Psychol Sci.* 2011; **22**(11): 1359–66.
PubMed Abstract | Publisher Full Text

Solomon DJ, Björk BC: **A study of open access journals using article processing charges.** *J Am Soc Inf Sci Technol.* 2012; **63**(8): 1485–1495.
Publisher Full Text

Solomon DJ, Björk BC: **Publication fees in open access publishing: Sources of funding and factors influencing choice of journal.** *J Am Soc Inf Sci Technol.* 2012; **63**(1): 98–107.
Publisher Full Text

Stodden V: **Open science: policy implications for the evolving phenomenon of user-led scientific innovation.** *Journal of Science Communication.* 2010; **9**(1): A05.
Reference Source

Stott A: **Open data for economic growth**. Washington DC: World Bank, 2014.
Reference Source

Suber P: **Open Access**. MIT Press, Cambridge, Mass, 2012.
Reference Source

Sutton C: **Is free inevitable in scholarly communication? the economics of open access.** *College & Research Libraries News.* 2011; **72**(11): 642–645.
Reference Source

Swan A, Brown S: **Authors and open access publishing.** *Learn Publ.* 2004; **17**(3): 219–224.
Publisher Full Text

Swan A, Brown S: **Open access self-archiving: An author study**. 2005.
Reference Source

Swan A: **The open access citation advantage: Studies and results to date**. 2010.
Reference Source

Swanson DR: **Two medical literatures that are logically but not bibliographically connected.** *J Am Soc Inf Sci.* 1987; **38**(4): 228–233.
Publisher Full Text

Tate D: **Open access and research assessment: Dealing with uk open access requirements in practice**. 2015; 58–62.
Publisher Full Text

Turner JA, Laird AR: **The cognitive paradigm ontology: design and application.** *Neuroinformatics.* 2012; **10**(1): 57–66.
PubMed Abstract | Publisher Full Text | Free Full Text

United Nations: **Universal declaration of human rights**. 1948.
Reference Source

van Assen MA, van Aert RC, Nuijten MB, *et al.*: **Why publishing everything is more effective than selective publishing of statistically significant results.** *PLoS One.* 2014; **9**(1): e84896.
PubMed Abstract | Publisher Full Text | Free Full Text

Van Noorden R: **Elsevier opens its papers to text-mining.** *Nature.* 2014; **506**(7486): 17.
PubMed Abstract | Publisher Full Text

Van Noorden R: **Open access: The true cost of science publishing.** *Nature.* 2013; **495**(7442): 426–429.
PubMed Abstract | Publisher Full Text

Van Noorden R: **Trouble at the text mine.** *Nature.* 2012; **483**(7388): 134–135.
PubMed Abstract | Publisher Full Text

van Rooyen S, Godlee F, Evans S, *et al.*: **Effect of open peer review on quality of reviews and on reviewers' recommendations: a randomised trial.** *BMJ.* 1999; **318**(7175): 23–27.
PubMed Abstract | Publisher Full Text | Free Full Text

Vanclay JK: **Factors affecting citation rates in environmental science.** *J Informetr.* 2013; **7**(2): 265–271.
Publisher Full Text

Veletsianos G, Kimmons R: **Assumptions and challenges of open scholarship.** *The International Review of Research in Open and Distributed Learning.* 2012; **13**(4): 166–189.
Reference Source

Vickery G: **Review of recent studies on psi re-use and related market developments.** *Information Economics.* 2011.
Reference Source

Vincent-Lamarre P, Boivin J, Gargouri Y, *et al.*: **Estimating open access mandate effectiveness: The melibea score.** *J Assoc Inf Sci Technol.* 2016.
Publisher Full Text

Wagenmakers EJ, Wetzels R, Borsboom D, *et al.*: **An agenda for purely confirmatory research.** *Perspect Psychol Sci.* 2012; **7**(6): 632–638.
PubMed Abstract | Publisher Full Text

Wagner B: **Open access citation advantage: An annotated bibliography.** *Issues Sci Technol Librarianship.* 2010; (60): 2.
Publisher Full Text

Walters WH, Linvill A: **Characteristics of open access journals in six subject areas.** *Coll Res Libr.* 2011; **72**(4): 372–392.
Publisher Full Text

Wang X, Liu C, Mao W, *et al.*: **The open access advantage considering citation, article usage and social media attention.** *Scientometrics.* 2015; **103**(2): 555–564.
Publisher Full Text

Watson M: **When will 'open science' become simply 'science'?** *Genome Biol.* 2015; **16**(1): 101, ISSN 1465-6906.
PubMed Abstract | Publisher Full Text | Free Full Text

Wicherts JM, Bakker M, Molenaar D: **Willingness to share research data is related to the strength of the evidence and the quality of reporting of statistical results.** *PLoS One.* 2011; **6**(11): e26828.
PubMed Abstract | Publisher Full Text | Free Full Text

Wicherts JM: **Peer Review Quality and Transparency of the Peer-Review Process in Open Access and Subscription Journals.** *PLoS One.* 2016; **11**(1): e0147913.
PubMed Abstract | Publisher Full Text | Free Full Text

Willinsky J: **Copyright contradictions in scholarly publishing**. *First Monday.* 2002; **7**(11).
Publisher Full Text

Willinsky J: **The Access Principle: The Case for Open Access to Research and Scholarship**. MIT Press, Cambridge, Mass, 2006.
Reference Source

Wohlrabe K, Birkmeier D: **Do open access articles in economics have a citation advantage?** *Munich Personal RePEc Archive.* 2014.
Reference Source

Xia J, Harmon JL, Connolly KG, *et al.*: **Who publishes in "predatory" journals?** *J Assoc Inf Sci Technol.* 2015; **66**(7): 1406–1417.
Publisher Full Text

Xia J: **Predatory journals and their article publishing charges.** *Learn Publ.* 2015; **28**(1): 69–74.
Publisher Full Text

Xu L, Liu J, Fang Q: **Analysis on open access citation advantage: an empirical study based on oxford open journals.** In *Proceedings of the 2011 iConference.* ACM, 2011; 426–432.
Publisher Full Text

Zhang Y: **The effect of open access on citation impact: a comparison study based on web citation analysis.** *Libri.* 2006; **56**(3): 145–156.
Publisher Full Text

Zhao D: **Characteristics and impact of grant-funded research: a case study of the library and information science field.** *Scientometrics.* 2010; **84**(2): 293–306.
Publisher Full Text

Zuccala A: **Open access and civic scientific information literacy.** *Information Research: An International Electronic Journal.* 2010; **15**(1).
Reference Source

Open Peer Review

Current Referee Status: ✓ ? ? ✓ ✓

Version 3

Referee Report 26 September 2016

doi:10.5256/f1000research.10376.r16583

Peter Suber
Berkman Center for Internet & Society, Harvard University, Cambridge, MA, USA

I'm commenting on this sentence from the first paragraph of Version 3 of the article:

The Green route is also enabled through author rights retention, in which authors pre-emptively grant non-exclusive rights to their institutions before publishing any works. The institution then has the ability to make articles by these authors OA without seeking permission from the publishers (e.g., this is the case of the Dutch Taverne amendment that has declared self-archival of research after 'a reasonable period of time' a legal right (Open Access NL, 2015)).

The authors added the bulk of this at my request in an earlier comment. I'm glad they did, but it's still only partially true. It's true that universities can adopt rights-retention OA policies that make it unnecessary to seek permission from publishers. But it's not true that Dutch Taverne amendment is an example. It's not a university policy, but legislation. It's legislation that gives authors permission for green OA regardless of the contracts they signed, and regardless of the rights they might have retained through a university policy. It's a very good idea and I recommend it everywhere. (There is already a similar law in Germany, and one is now emerging in France.) But what the authors need here is an example of a university rights-retention OA policy, or a thorough explanation of this kind of OA policy.

I'm not seeking a citation to my own university's activity, though it adopted the first OA policies of this kind. Nor am I seeking a citation to the good-practices guide that I maintain with Stuart Shieber, though it's the standard reference on this kind of policy. In fact, I was reluctant to follow up the authors' response to my prior suggestion because I didn't want to appear to seek additional citations.

But since I've gone this far, I'll mention these two sources anyway:

The Harvard open-access policies
https://osc.hul.harvard.edu/policies/

Good practices for university open-access policies
http://bit.ly/goodoa

Of the two, I'd recommend the second in this situation. But even if the authors include no citation on this point, at least they should stop citing the Dutch law as an example, and treat it separately as another path to the same goal.

I have read this submission. I believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.

Competing Interests: No competing interests were disclosed.

Version 2

Referee Report 25 August 2016

doi:10.5256/f1000research.9609.r14274

Gwilym Lockwood
Neurobiology of Language Department, Max Planck Institute for Psycholinguistics, Nijmegen, Netherlands

I have no further comments; my points or concerns have been addressed, and other issues are highlighted in further depth by other reviewers.

I have read this submission. I believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.

Competing Interests: Since I submitted my first review, I have been commissioned by Altmetric to write blogs using their data. However, all my reviewing was done before knowing about this opportunity, and so my original comments on the alternative metrics section were therefore unaffected.

Referee Report 27 June 2016

doi:10.5256/f1000research.9609.r14275

Peter Suber
Berkman Center for Internet & Society, Harvard University, Cambridge, MA, USA

In the first version I reviewed, you said, "This [green] route is dependent on journal or publisher policies on self-archiving (sherpa.ac.uk/romeo)."

That was untrue or incomplete and I offered this comment: It overlooks rights retention. Some individual authors retain enough rights to authorize green OA on their own. While this may be fairly rare, rights-retention OA policies at universities are increasingly common. More than 80 institutions in North America, Europe, Africa, and Asia have now adopted rights-retention OA policies. Under these policies, the faculty grant non-exclusive rights to their institution before they sign future publishing contracts. The institution then has permission to make those future articles OA without having to seek permission from publishers. (The institutions also grant the same non-exclusive rights back to authors.) For more detail on rights-retention OA policies, see Stuart Shieber and Peter Suber, "Good Practices for University OA Policies."
http://bit.ly/goodoa

You revised the text in a way that missed my point and misstated my position: "While academics themselves may have little power in debates regarding copyright, institutes could claim ownership of the work they likely already own by invoking their rights under the work made-for-hire doctrine (Denicola, 2006). However, it is difficult to imagine researchers favoring university-held rather than journal-held copyright, and a system of non-exclusive rights is preferred, as is reflected in OA policies and OA journals (Suber, 2012)."

Here's the main point: More than 80 universities around the world have adopted rights-retention OA policies. These policies are adopted by faculty votes, not administrative edicts. At these institutions the rights needed to authorize OA are not seized from faculty by the institution, or claimed by the institution through work-for-hire. These policies presuppose that the rights initially belong to faculty, not the institution. If the institution is to exercise them, faculty must voluntarily grant them. (There are details we needn't go into here, for example, that we're only talking about non-exclusive rights, and that these policies generally include waiver options when faculty don't want the institution to have rights to a given work.) At Harvard, which pioneered this type of policy, four of the school-level votes were unanimous. In short, it's not at all "difficult to imagine researchers favoring university-held...copyright." On the contrary, it's easy to imagine and widely attested. Green OA does not always depend on permission from publishers. Increasingly it depends on rights retention by authors, through carefully drafted and widely supported university policies. That's a fact. My opinion is that that's a good thing.

You needn't share my opinion, and needn't mention the fact. But please don't misrepresent the fact or reverse my opinion.

I have read this submission. I believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.

Competing Interests: No competing interests were disclosed.

Referee Report 10 June 2016

doi:10.5256/f1000research.9609.r14273

Chris Chambers
School of Psychology, Cardiff University, Cardiff, UK

All my suggestions have been addressed. I'm very happy approve this interesting and useful addition to the literature.

I have read this submission. I believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.

Competing Interests: No competing interests were disclosed.

Version 1

Referee Report 28 April 2016

doi:10.5256/f1000research.9110.r13295

Chris Chambers
School of Psychology, Cardiff University, Cardiff, UK

Tennant *et al* offer a timely and insightful review of the various effects of open access publishing on science and society. The paper is well structured and enjoyable to read. Although I am not an expert on open access publishing, I also found the discussion of the literature quite balanced and evidence-based.

I have just three recommendations for revisions:

1. In the discussion of the OA citation advantage (which is excellent), the authors are very careful to avoid asserting a causal link between the OA status of a paper and the number of citations it generates. However, in my view, the conclusion of the Gargouri (2010) paper should be caveated. While the results of Gargouri are consistent with the absence of a self-selection bias in producing the OA advantage, they cannot rule it out. It might be the case that funders or institutions that mandate OA are also those that are more likely to support/host higher impact research. Furthermore, given that OA mandates are never 100% effective, perhaps authors are more likely to comply with a mandate for work they perceive to be of high impact. A useful addition to this section would be specify the conditions under which we *could* determine whether or not OA causally influences citations. This would require a randomised controlled trial in which articles are randomly are assigned to OA and non-OA routes. To my knowledge, no such a trial has yet been undertaken, although the authors will be in a more informed position to know whether this is the case.

2. There is not much discussion in the paper of the distinction between full OA and hybrid OA. Given that the APCs for hybrid OA articles tend to be substantially higher than those for full OA articles, this may warrant more prominent coverage in the economic case for OA. It is also relevant to the brief mention of the Finch Report, which (controversially) favours gold OA (including the hybrid route) over green OA.

3. I was glad to see the link formed between OA and open science more generally, as many researchers and advocates draw a distinct (and somewhat arbitrary) line between these initiatives. In this context, it would perhaps be relevant to mention the TOP guidelines (https://cos.io/top/). TOP is very much an evolving entity so it would be interesting to consider the inclusion of an OA standard in future revisons of TOP.

Minor

p13 This sentence is difficult to parse: "Whereas this is hyper-variable, and strongly dependent on a range of factors, it is the fact that any public interest in science that is of importance."

I have read this submission. I believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.

Competing Interests: No competing interests were disclosed.

Reader Comment 30 Apr 2016

David Wojick, David E. Wojick Associates, USA

Chris, given my comments below about how one-sided this article is, in favor of OA, I am curious as to how you think it is balanced? Balanced in what respect?

Competing Interests: No competing interests were disclosed.

Reader Comment 01 Aug 2016

Charles Oppenheim, Retired, UK

David, in view of your well-known antipathy to OA, don't you think you should have declared your competing interest?

Competing Interests: I am in favour of OA

Referee Report 25 April 2016

doi:10.5256/f1000research.9110.r13292

Anne Tierney
Department of Learning and Teaching Enhancement, Edinburgh Napier University, Edinburgh, UK

This paper is a comprehensive review of the complexities of OA. I have come late to the discussion on this paper, and I find that the previous reviewers have been meticulous in their critique of the paper, to the point I have very little to add. However, there are a couple of points for consideration. What is the effect (if any) of the UK Research Excellence Framework on Open Access? To what extent is disciplinarity a factor in Open Access? I ask this question because of the high impact of the sciences and biomedical research, but wonder about Arts and Humanities (and other areas) as a comparision. As one of the other reviewers said, "this doesn't consider the fact that the prestige of some journals is advertising in and of itself." While this is true of scientific journals, the same can't be said, for example, of education journals, so there is a lack of parity between disciplines. There was also an assumption of the willingness of reviewers to continue to offer their services freely. This aspect of OA (and subscription-based) publishing is hardly ever critiqued, but it assumed to be part of the process.

All in all, this paper gives a lot of food for thought. I don't expect a rewrite of the paper, based on my comments, but I would welcome further discussion on where the authors (and readers) see OA going in the future.

I have read this submission. I believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.

Competing Interests: No competing interests were disclosed.

Referee Report 22 April 2016

doi:10.5256/f1000research.9110.r13291

Paige Brown Jarreau
Manship School of Mass Communication, Louisiana State University, Baton Rouge, LA, USA

This is an interesting and timely review of the issue of open access to scientific literature.

The two other reviewers have highlighted specific issues that should be addressed in the revision of this article, and I agree with these issues. I've added other notes below. I think this review article would benefit from a re-write to correct potentially biased language in support of open access and to round out the review with further evidence of open access impacts on citation rates, altmetrics, scientific literacy / public engagement and research quality.

Notes:

The authors cite *"fostering a culture of greater scientific literacy"* as a benefit of open access. While this is theoretically a benefit, has more or less open access in particular scientific fields been tied to greater or lesser scientific literacy in those areas? Is this potential benefit supported by research literature? The authors should be clear on what the evidence-based benefits of open access are, and also what the potential drawbacks are. References to related research should be provided on this topic.

The structure the authors use for laying out their evidence and the language they use (e.g. *"[the] case for Open Access"*) appear to lean more toward the positive impacts / benefits of open access from the outset. The authors should be very careful to review the evidence first before making value-based statements or arguments about open access, even if the evidence-based benefits outweigh any potential drawbacks, or lack of significant benefits, in the end.

Correct the typo in the following sentence: *"[In] A longitudinal study Eysenbach (2006) compared…"*

The authors write: *"One alternative explanation for the citation advantage could be that researchers choose to publish OA when a finding is more impactful, but empirical evidence contradicts this selection effect. Gargouri et al. (2010) compared citation counts within a cohort of OA articles that had either been self-selected as OA or mandated as OA (e.g., by funders). The study concluded that both were cited significantly more than non-OA articles. As such, these findings rule out a selection bias from authors as the cause for the citation advantage (Gargouri et al., 2010)."* However, couldn't funded research also have a tendency to be considered "more impactful," because it was chosen in the first place to be funded and mandated as OA? The authors should discuss this, and whether there is any research that experimentally investigates whether open access provides a citation advantage. This could perhaps be suggested as future research. The authors should also discuss how/why some studies have found no citation advantage for OA papers.

Related to social media mentions of research papers and citation counts, the authors might also consider citing Liang, X., Su, L. Y. F., Yeo, S. K., Scheufele, D. A., Brossard, D., Xenos, M., ... & Corley, E. A. (2014). Building Buzz (Scientists) Communicating Science in New Media Environments. *Journalism & Mass Communication Quarterly*, 1077699014550092. [PDF]

In discussing the economics of OA, the authors should also discuss any evidence of potential drawbacks for various stakeholders, such as where funds for pay-to-publish fees will come from and how these fees may affect individual researchers. Pay-to-publish models of OA may also burden early career researchers and researchers working in fields where research grants are more difficult to obtain.

The authors do not discuss the potential impact of OA on research quality or reproducibility (reproducibility is only mentioned in the context of open data). As this has been a controversial issue in the past (e.g. the mentioned 'sting' operations) the authors should discuss any research that has

investigated the impact of open access on the rigor of peer reviews, research quality, presence of replication studies / reproducibility, etc. There has also been some discussion of whether open peer review (like that used by F1000Research) affects the quality of reviews, e.g. (Rooyen *et al.* 1999). The authors should mention this and/or subsequent literature when addressing open peer review.

There have also been studies on scientists' / journal article authors' perceptions and attitudes toward open access, e.g. Schroter and Tite (2005; 2006). The authors might considering summarizing some of this research, as it gives context to some of the existing barriers to open access and perceived drawbacks among researchers.

Summary:

In summary, the topic of this review is important and timely. However, this paper falls short of what I would expect from a systematic review in terms of systematically summarizing previous research findings related to the impact of open access on scientific publishing, public engagement with science, science literacy and altmetrics. The authors should be careful to hold back value judgements / arguments related to the case for open access until having systemically reviewed the evidence-based benefits, drawbacks, and/or lack of significant benefits. The authors should also provide some discussion of how we might objectively weigh any evidence-based benefits with potential drawbacks for various stakeholders including researchers and especially early career researchers. The authors should avoid summarizing mostly the findings of previous studies that find positive impacts of open access on the various domains of potential impact they consider in their paper. The evidence already presented in the paper is rigorous and detailed. However, I would recommend a revision that rounds this review out with more systematic evidence.

References

1. Liang X, Su L, Yeo S, Scheufele D, Brossard D, Xenos M, Nealey P, Corley E: Building Buzz: (Scientists) Communicating Science in New Media Environments. *Journalism & Mass Communication Quarterly*. 2014; **91** (4): 772-791 Publisher Full Text
2. van Rooyen S, Godlee F, Evans S, Black N, Smith R: Effect of open peer review on quality of reviews and on reviewers' recommendations: a randomised trial. *BMJ*. 1999; **318** (7175): 23-27 Publisher Full Text
3. Schroter S, Tite L, Smith R: Perceptions of open access publishing: interviews with journal authors. *BMJ*. 2005; **330** (7494): 756 PubMed Abstract | Publisher Full Text
4. Schroter S, Tite L: Open access publishing and author-pays business models: a survey of authors' knowledge and perceptions.*J R Soc Med*. 2006; **99** (3): 141-8 PubMed Abstract | Publisher Full Text

I have read this submission. I believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.

Competing Interests: No competing interests were disclosed.

Referee Report 18 April 2016

doi:10.5256/f1000research.9110.r13430

Peter Suber
Berkman Center for Internet & Society, Harvard University, Cambridge, MA, USA

The article is very well-done, unusually thorough and detailed. Here are a few ways to improve it.

When I refer to page numbers, I mean the page numbers in the PDF of v1, April 11, 2016.
http://f1000research.com/articles/5-632/v1

"You" refers to the authors.

Apologies in advance if I sometimes cite my own work in these comments.

p. 1. In the abstract you say, "The economic case for Open Access is less well-understood, although it is clear that access to the research literature is key for innovative enterprises, and a range of governmental and non-governmental services."

This understates the economic case. For example, some subscription journals convert to OA precisely for economic benefits.

See the preliminary version of David Solomon, Bo-Christer Björk, and Mikael Laakso, "Converting Scholarly Journals to Open Access: A Review of Approaches and Experiences" now open for public comment. (The final version will be published this summer.)
https://osc.hul.harvard.edu/programs/journal-flipping/public-consultation/

See especially section 4.6, "Increased revenue and financial viability."
https://osc.hul.harvard.edu/programs/journal-flipping/public-consultation/4/6/

p. 3. You say, "The Green route refers to author self-archiving, in which a version of the peer-reviewed article is posted online to a repository or website."

Green OA also applies to preprints, which are not peer-reviewed.

p. 3. You say, "This [green] route is dependent on journal or publisher policies on self-archiving (sherpa.ac.uk/romeo)."

This is importantly incomplete. It overlooks rights retention. Some individual authors retain enough rights to authorize green OA on their own. While this may be fairly rare, rights-retention OA policies at universities are increasingly common. More than 80 institutions in North America, Europe, Africa, and Asia have now adopted rights-retention OA policies. Under these policies, the faculty grant non-exclusive rights to their institution before they sign future publishing contracts. The institution then has permission to make those future articles OA without having to seek permission from publishers. (The institutions also grant the same non-exclusive rights back to authors.)

For more detail on rights-retention OA policies, see Stuart Shieber and Peter Suber, "Good Practices for University OA Policies."
http://bit.ly/goodoa

p. 3. You say, "A subscription to all peer-reviewed journals is not affordable for any single individual, research institute or university (Odlyzko, 2006)."

This is true and important, but it's a pity you don't cite more recent evidence than 2006.

An important kind of evidence for this proposition is that not even Harvard University can afford all the journals needed by its faculty and students, and must cancel journals every year for budgetary reasons alone. I've collected seven public statements from Harvard to this effect (2008-2012) in the supplements to p. 30 of my 2012 book (Open Access, MIT Press, 2012).
http://bit.ly/oa-book#p30.2

p. 3. You say, "Much of the driving force behind this global change has been through a combination of direct, grassroots advocacy initiatives in conjunction with policy changes from funders and governments."

Please add *university policies* to this list. They're on a par with funder policies in importance, and they're far more numerous. ROARMAP shows that 7+ times more universities have OA policies than funders.

p. 3. You say, "The Open Access movement is intrinsically tied to the development of the Internet and how it redefined communication and publishing (Laakso et al., 2011)."

For more documentation on how the OA movement arose as soon as the internet arose, see my Timeline of the Open Access Movement.
http://legacy.earlham.edu/~peters/fos/timeline.htm

In 2009, I moved the timeline to the Open Access Directory wiki, and you should probably cite that version:

Timeline of the open access movement...
http://oad.simmons.edu/oadwiki/Timeline

...especially subsection on developments before 2000.
http://oad.simmons.edu/oadwiki/Timeline_before_2000

(You might cite this timeline again on p. 5, when you introduce your own timeline.)

p. 3. You say, "One result of the growing OA movement is the rise of OA-only publishers...."

Somewhere in this paragraph, I'd mention that some OA publishers are for-profit (e.g. BMC) and some are non-profit (e.g. PLoS).

p. 4. Your section on the impact advantage is very well-done. Most treatments are much briefer, less careful, and less detailed than yours.

I have just these suggestions. You cite authors of individual studies, and Alma Swan's 2010 literature review. But you don't cite the mother lode of literature on this topic: Steve Hitchcock's annotated bibliography, "The effect of open access and downloads ('hits') on citation impact: a bibliography of studies."
http://eprints.soton.ac.uk/354006/1/oacitation-biblio-snapshot0613.html

Or more precisely, you cite it once, 10 paragraphs before the section on the impact advantage begins. You should cite it again within the section on the impact advantage. You should mention that it's

comprehensive and annotated.

Hitchcock stopped updating it in 2013. But you should mention that SPARC Europe has committed to update it through its Open Access Citation Advantage Service.
http://sparceurope.org/oaca/

Finally, in the same place where you cite Swan's literature review, you should cite Ben Wagner's literature review, "Open Access Citation Advantage: An Annotated Bibliography," Issues in Science and Technology Librarianship, Winter 2010.
http://www.istl.org/10-winter/article2.html

p. 5. In the timeline entry for 2002, the BOAI was released on February 14, not January 14.

p. 6. In the timeline entry for 2013, I'd say that the suicide of Aaron Swartz "increases" (not "gains") international attention for the OA movement, or "draws new attention" to the OA movement. The current language suggests that the OA movement didn't have international attention before that, which is very far from the truth.

p. 9. You say, "Shifting copyright to stay with the author allows for wider re-use, including TDM, and forms the basis for a robust and developing public domain."

You shouldn't use "public domain" here. In copyright law, the term has a specific meaning which you don't mean here.

p. 10. You say, "Only recently has any transparency into the detailed costs of subscriptions been gained by using Freedom of Information Requests to bypass non-disclosure agreements between libraries and publishers (Lawson & Meghreblian, 2015)."

Here you overlook the earlier Big Deal Contract Project in the US, from Ted Bergstrom, Paul Courant, and Preston McAfee. It too used public records laws and Freedom of Information requests. I'm not sure when it launched, but it was before 2009.
http://www.econ.ucsb.edu/~tedb/Journals/BundleContracts.html

p. 10. You say, "The average production cost for one paper is estimated to be around $3500–$4000 (Van Noorden, 2013)."

I've seen dozens of widely varying estimates of this cost, most of them much lower than Van Noorden's. Unfortunately I don't have time to hunt them down. I hope you can introduce at least a few more, if only to show that estimates differ widely here.

p. 10. You say, "Philip Campbell (Editor-in-Chief of Nature) stated that his journal's internal costs were at $20,000–$30,000 per paper...."

To clarify, I think he meant that this was the cost per **published** paper. If Nature rejects x articles for every one it publishes, then this includes the cost of peer reviewing x rejected articles. Since Nature is very selective, x is high. But this "cost per published paper" should not be compared to costs for peer-reviewing a single paper or the production costs of publishing an accepted paper.

p. 10. You say, "OA publishing is most prevalent in the form of 'pay-to-publish'...."

This is either false or misleading. About 70% of peer-reviewed OA journals charge no APCs at all. In that sense, the fee-based model is not the most prevalent. It's a minority model. On the other hand, about 50% of the articles published in peer-reviewed OA journals are published in the fee-based variety.

On my claim that most OA journals charge no APCs:

See my article, "Good facts, bad predictions," SPARC Open Access Newsletter, June 2006.
https://dash.harvard.edu/handle/1/4391309

And my article, "No-fee open-access journals," SPARC Open Access Newsletter, November 2, 2006.
https://dash.harvard.edu/handle/1/4552050

The DOAJ used to make it easy to see what percentage of listed journals were fee-based and what percentage were no-fee. But it has temporarily made that difficult by combining the categories of "no-fee journals" and "journals for which we don't have enough information to say."

On my claim that about half the articles published in peer-reviewed OA journals are published in the fee-based variety, see the updates to p. 170 of my 2012 book. There I cite three studies and quote the relevant excerpts.
http://bit.ly/oa-book#p170

p. 10. When you describe ways in which fee-based OA journals mitigate some problems arising from the model, you mention the firewall between the editorial and business side of the journal (good), and you mention fee waivers (good).

You should also mention fee discounts, which many journals give in lieu of fee waivers.

You should also mention that most fees charged by fee-based journals are paid by funders (59%), or the author's employer (24%), and only 12% are paid by authors themselves. These numbers are from Suenje Dallmeier-Tiessen et al., "Highlights from the SOAP project survey. What Scientists Think about Open Access Publishing," arXiv, January 29, 2011, Table 4.
http://arxiv.org/abs/1101.5260

p. 12. You say, "In his article Sutton (2011)..."

please change "his" to "her". The article is by Caroline Sutton.

p. 12. You say, "While OA is not a solution to all aspects of research accessibility (e.g., language barriers and disability access remain continuing issues to be addressed)...."

See my 2012 book (Open Access, MIT Press, 2012, http://bit.ly/oa-book), at pp. 26-27, where I make much the same point. "OA isn't universal access" and by itself doesn't overcome "filtering and censorship barriers", "language barriers", "handicap access barriers", or "connectivity barriers".

p. 13. You refer to "the fact that access to knowledge is actively prohibited in fields like public health...."

I don't know what you mean here by "actively prohibited".

p. 13. You say, "Some traditional publishers, and some academics, have argued that public access to research is not required because research papers cannot be understood by non-specialists...."

Here you might want to cite Section 5.5.1 ("OA for Lay Readers," pp. 115-119) of my 2012 book.

p. 13. You say, "The shift from a 'reader pays' to an 'author pays' mode...."

I recommend avoiding the term "author pays" for the reasons I gave in my fourth comment to p. 10 above. Most OA journals don't charge author-side fees, and among those who do, most fees are not paid by authors.

p. 13. You say, "This has been at least partially mitigated with waiver fees for authors from developing countries and additional provisions in research grants...."

Yes. But again, don't forget that the majority of peer-reviewed journals are no-fee journals. See my fourth comment to p. 10 above.

pp. 15-15. You say, "Fortunately, it seems that funders and research organisations are moving in that direction. Since 2005, the number policies supporting OA publishing increased steadily. Consequently, it is now the responsibility of researchers to ensure OA to their publications either by choosing the green or the gold road."

Since you're recapitulating some grounds for optimism here, I'd also reiterate the growth of rights-retention OA policies. See my second comment on p. 3 above.

p. 16. You say, "As Peter Suber, a leading voice in the OA movement, stated: 'As long as they do not have the power to stop Open Access, the toll-access publishers are not the enemy'."

Thanks for quoting me. I wanted to give you the source to cite. Unfortunately, I don't think you're using an exact quote. Here's the closest one I can find:

"TA [toll-access] publishers are not the enemy. They are only unpersuaded. Even when they are opposed, and not merely unpersuaded, they are only enemies if they have the power to stop OA. No publisher has this power, or at least not by virtue of publishing under a TA business model. If we have enemies, they are those who can obstruct progress to OA. The only people who fit this description are friends of OA who are distracted from providing OA by other work or other priorities."

It's from "Two distractions," SPARC Open Access Newsletter, May 3, 2004.
https://dash.harvard.edu/handle/1/4391169

I have read this submission. I believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.

Competing Interests: No competing interests were disclosed.

Reader Comment 29 Apr 2016

Alexander Grossmann, Faculty of Media, Leipzig University of Applied Sciences, Germany

This is a very useful reviewer's report which may assist the authors to modify and further improve their paper in it's next version.
A question to Peter about his claim that most OA journals charge no APC: You say "About 70% of peer-reviewed OA journals charge no APCs at all. In that sense, the fee-based model is not the most prevalent. It's a minority model."
Would you give a more recent reference for this statement, please? The cited publications which were mentioned in your report seem to be not up-to-date at least because they were published 10 years ago. Meanwhile there are much more OA journals on the market, both hybrid or OA-only, in comparison with 2006. Moreover we should consider the fact that new OA journals are mostly free of charge for the first two years in their market introduction phase. After this period they are likely to become author-payed journals with APCs. I am not sure if the latter observation has been reflected by those statistics, too.
Thanks for clarifying this point.

Competing Interests: No competing interests were disclosed.

Author Response 10 May 2016

Jon Tennant, Imperial College London, UK

Hi Alexander,

I've been looking into this, and the most up-to-date statistics for this based on the DOAJ come from this source: http://citesandinsights.info/civ16i4.pdf. Figures here seem to suggest that 71% of journals in the DOAJ do not levy an APC. This is likely to change slightly with the updated 'crackdown' from the DOAJ (http://www.nature.com/news/open-access-index-delists-thousands-of-journals-1.19871), but I'll add a reference to this in.

Jon

Competing Interests: I work with Alexander at ScienceOpen.

Referee Response 24 May 2016

Peter Suber, Harvard University, USA

Here are the latest stats from the DOAJ (May 24, 2016).
https://goo.gl/LejTAw

Competing Interests: No competing interests were disclosed.

Author Response 25 May 2016

Jon Tennant, Imperial College London, UK

Dear Alexander,

I'm just posting the link that Peter was kind enough to send us yesterday, with updated statistics on this matter: https://plus.google.com/+PeterSuber/posts/HjrRDcrZS8p

Important points:

Here are the numbers as of May 24, 2016:

Total number of journals listed in DOAJ = 8,858
Yes (fee-based) = 1,424 = 16%
No (no-fee) = 2,601 = 29%
No info = 4,833 = 55%

Further references and information are provided in that post, and will be integrated into the next version of this manuscript.

Best,

Jon

Competing Interests: No competing interests were disclosed.

Referee Report 15 April 2016

doi:10.5256/f1000research.9110.r13293

Gwilym Lockwood
Neurobiology of Language Department, Max Planck Institute for Psycholinguistics, Nijmegen, Netherlands

This is the first time I've written an open peer review, although I always sign private peer reviews. Normally I'd make comments directed to the authors and the authors alone, but since this is open, I've also included a section for other readers of this paper. This may sound a bit like an Amazon or Airbnb review or something.

Short summary for readers

This is an excellent paper about the academic, economic, and societal benefits and impacts of Open Access. It's a good introductory text for people who don't know much about OA and would like to know more. It's also a good persuasive text for stakeholders in policy, universities, publishing, funding, etc. positions who may be interested in including OA in their decision making.

In addition to its attention to detail, its main strengths are its focus, its brevity, and its relative impartiality.

One of the difficulties with writing about OA is that there are so many overlapping issues; this paper is very good at giving a brief overview or description of the other issues, pointing the reader in the direction of somewhere with more information, and then getting back onto the topic.

Another thing about OA is that its advocates are very passionate about it. As with any cause, that's a good thing for its supporters, but overwhelmingly pro-OA resources can seem potentially off-putting to neutrals. This paper does an excellent job of presenting an evidence-based pro-OA viewpoint in a measured tone

and without coming across as ideological.

One possible caveat is that the paper presents extensive evidence of what OA does, but it doesn't tackle the meatier issue of how to implement it successfully. However, I feel that's a separate issue which is beyond the scope and purpose of this paper.

Suggestions and comments for authors

First of all, great article! Well done and thank you for pulling together what is a disparate collection of links and literature into a one-stop shop which is both useful and coherent. I like this article a lot... but my role here is to criticise and make it better, so the rest of this review will focus on that.

This article is well-written and well-structured. That's made it much easier as a reviewer to simply go through the article and highlight my issues with it paragraph by paragraph, rather than having to make it coherent first and then sort out the smaller things.

The vast majority of the issues I have with this paper are minor ones, so it didn't make sense to have separate major/minor sections; rather, I'll just go through them in order in the text.

(I printed this out to underline/comment on, so for me, tables 1 and 2 came during the academic case for OA section. Online, they're supplementary materials, and I think it's best that way, but this is why I'm commenting on the tables during that section)

Abstract

"We recommend that OA supporters focus their efforts on working to establish viable new models and systems of scholarly communication, rather than trying to undermine the existing ones..."
In general, I agree with this sentiment. However, I feel that its inclusion in the abstract is a bit jarring as the text of the article doesn't really cover recommendations to OA supporters at all, other than in the very last paragraph. I think that's good, as I feel this paper is best suited as relatively neutral source of information rather than a preaching to the converted or ideology discussion kind of purpose. So, I think this part can be left out of the abstract; it doesn't refer to any particular "recommendations to OA supporters" bit in the text and potentially clouds the strength of the relative impartiality of the paper.

A brief history of OA

"BioMed Central and ... PLOS were founded in the early 2000s and remain successful businesses to date." (p3, col2)
Technically, PLOS is a non-profit. I suggest changing *successful businesses* to successful business models. This both highlights the financial sustainability of OA (increased APCs at PLOS notwithstanding) and also sets it apart from traditional publishers, which are definitely successful businesses.

The academic case for OA

figure 1 (p4, top)
I have difficulty interpreting the y-axis on figure 1. It's labelled as cumulative number of PubMed articles relative to 2000, but I'm not sure how to read it. Reading off 2014, non-OA is c.22 on the y-axis, and OA is c.33 on the y-axis. Based on the figure 1 caption about the ratio, I'm interpreting this as meaning that, in 2014, the ratio of cumulative PubMed articles was approx 33:22 OA to non-OA, or in other words, 60% of

PubMed articles in 2000-2014 were OA. However, I'm not sure if this is how it's meant to be interpreted. I think that it's well visualised, and really makes it clear how OA has taken off, but exactly what the numbers represent on the y-axis is unclear to me: number of articles? number of times more articles? It could use some relabelling.

"Napster moment" (p4, col1)
I like the comparison, but it could use a citation (even just the Napster wikipedia article) and/or a little more explanation to clarify what that means.

Table 1:
"1991 ... by the American physician Paul Ginsparg"
He's a physicist, not a physician.

Table 2:
Xu et al (2011)
I don't think this reference was very well cited. Firstly, *Oxford Open Journals* are listed as a discipline, when they're the source of papers across disciplines. The actual disciplines were Medicine, Social Sciences, Mathematics & Physical Sciences, Life Sciences, and Humanities. Secondly, you list the citation advantage as 138.87%. However, one of the main findings of this paper was the disparity in citation advantages; it ranged from 163.16% for OA articles in Mathematics & Physical Sciences to an actual citation disadvantage of -49.24% for OA articles in Humanities. Given the pro-OA nature of the paper, I feel like you have an extra responsibility to report the few anti-OA pieces of evidence.

Gargouri et al (2010) (page 8, col1)
This paragraph is about a possible confound for the OA citation advantage, where it could be that researchers choose to publish OA for extra cool findings, and you use the Gargouri et al. study to counter this... which is totally correct. You write:
"Gargouri et al. (2010) compared citation counts [for articles which were] self-selected as OA or mandated as OA. The study concluded that both were cited significantly more than non-OA articles. As such, these findings rule out a selection bias"
This is true that both OA types were cited more than non-OA. However, it's also missing the crucial point that there was no difference in citation between self-selected OA articles and mandatory OA articles. Including this would strengthen your point to show that it's OA itself which leads to the citation advantage.

The whole section about altmetrics (subhead societal impact of the academic case for OA, p8, col2) could use some attention. It's not clear until much later what the difference is between alternative metrics (i.e. altmetrics), i.e. the various types of metrics which are alternative to journal impact factors, and Altmetric, i.e. the company which is often confusingly referred to as Altmetrics (not in this paper, to be fair, but elsewhere). A quick disambiguating sentence or two would be really useful here.

In the following paragraph (page 8, col2), you write about OA altmetrics advantage, and say that there's a logical assumption that OA articles should have one. However, this doesn't consider the fact that the prestige of some journals is advertising in and of itself. You can, and do, get a lot of closed-access papers which generate high altmetrics (social media attention, Mendeley readhership) from academics who do have access. And sure enough, in the next paragraph, (page 8 and 9), the Wang et al. 2015 article finds that the OA altmetric advantage doesn't extend to the most impactful articles. I think this section can be made more nuanced and informative by quickly discussing the role of journal prestige. Nothing in depth, just as something that exists and needs to change (for example, you could point people to Brembs *et al.* and the Deep Impact paper in Frontiers).

"Essentially, copyright is a tool wielded by traditional publishers for financial gain rather than fostering creativity..."
I don't disagree with this. However, I feel it comes on too strong. I think it's fair to say that most people's immediate opinion of copyright is "well, I'd like my stuff to be copyrighted, as that means people can't steal it and pass it off as their own". I think that you need a little more detail here, even just two or three sentences to explain how and why copyright is used for financial gain rather than author protection. Otherwise, it just sounds political/ideological, and counterintuitive for people who haven't read much about copyright.

Glenisson et al. (2005) citation (page 9, col2).
You write that TDM has "proven to be useful for a large variety of applications", and use the Glenisson citation to back this up. I have to say here that I know very little about TDM; however, following through to the Glenisson paper, I don't see how it supports that conclusion. I read it and it seems to show a proof-of-concept kind of study: that TDM can group a set of papers into themes in the same way that an expert can. This is really cool and everything, but I don't think that that substantiates your point that TDM is useful for a large variety of applications. Rather, I'd like to see a couple of specific examples, which you then describe more fully in the next paragraph. One good one is Swanson 1987 (I think - taken from here: http://people.ischool.berkeley.edu/~hearst/papers/acl99/acl99-tdm.html), who used TDM to make the link between migraines and magnesium deficiency.

"...simply because one can no longer keep up with the published literature".
Small point, but I think it's worth stressing that this is due to the amount of literature that there is.

The economic case for OA

the pay-to-publish part (p10, col2)
I feel this glosses over problems with pay-to-publish. You come back to predatory OA later, but this isn't quite the same: I think it could use a couple of extra sentences describing what the conflict of interest for researchers is, and also stress that pay-to-publish makes it potentially in a journal's interest to accept more papers than they necessarily should. One of the most common anti-OA arguments I see in non-scientific media is that OA is pay-to-publish, which is often misrepresented as "pay-to-publish is publication bribery". I think this section needs a little more substance to it to acknowledge/address this.

"making publication costs dependent on the value added..." (page 11, col1)
When talking about the value added by journals, this paragraph ignores the elephant in the room: journal prestige. Again, I know that this isn't the purpose of this article, but I think it could really be strengthened by mentioning it before moving on.

"Much primary research actually takes place outside of academia inside research and development departments" (page 11, col2)
The part following this sentence is muddy. First, you talk about R&D outside academia (i.e. presumably private research), and then you talk about access to research results because they're publicly financed public goods. So, what does that mean, that R&D from private businesses who've invested their own capital in it should be made available to all? (maybe I agree with that, in some cases, but a lot of people sure won't)
I think this paragraph could be honed a bit; otherwise, it's straying into the ideological territory of saying that all private research should be made public for the public good. That transcends OA in scholarly publishing, and makes OA in scholarly publishing too easy to dismiss.

The cancer research paragraph (page 12, col1) is also unclear. It took me a while to figure out it's talking about UK expenditure - my first assumption of "total expenditure" meant worldwide. It's also not totally clear what the point is - the geographical origin of research is unrelated to its open status. I think that it's quite a leap to write (apologies for paraphrasing) "83% of UK economic benefit from cancer research comes from research outside UK, therefore open access is good", because I think it conflates two different things.

Also, small point, "17% of the annual net-monetary was estimated" is missing the word benefit after net-monetary.

The societal case for OA

Small point: as somebody who wears a linguistics hat quite often, it rankles to read on page 11 "Examples of [non-academic] groups who might benefit include... those who work in linguistics and translation". Translation, for sure, but linguistics is an academic field - you even mention the Lingua to Glossa movement organised by academic linguists later in the manuscript! To me, this is like writing "...those who work in biology and vets", lumping the academic field and a practical use of that field together. Just referring to translation is fine.

Citizen engagement (page 13, col 1)
I agree that these are great examples of citizen engagement with science, but at the risk of sounding like an Elsevier representative, interest in projects like Galaxy Zoo does not entail desire to download and read papers. In fact, you could even make the (spurious) argument that those projects come into existence precisely because citizens aren't interested in downloading and reading papers. I don't actually agree with that, I agree with your general point... but I think that citizen science project interest and citizen science paper interest. Obviously I think it is in the public interest to have science journals OA, but this isn't the right argument (and I think the sentence "Such statements conflate a lack of desire or need for access with the denial of opportunity to access research" is perfect). I think a stronger argument would be to look at existing OA journals, such as PLOS and Frontiers, and see how many views and downloads come from people who aren't academics. If you can point to, say, some of the most viewed/downloaded PLOS papers and say "look, 30% (or whatever, that's a random number) of these readers aren't academics, they're real people who are interested in it", that would make for a stronger argument.

Quibble about the "yes, we were warned about Ebola" example: the finding from that paper (that Liberians have Ebola antibodies in their blood, suggesting the endemic presence of Ebola) is actually written on the first page preview of the paper (
http://www.sciencedirect.com/science/article/pii/S0769261782800282/part/first-page-pdf, accessed from my laptop outside my institution). It could be argued that anybody could see this finding anywhere in the world, meaning that it's not a problem of OA, it's a problem about searching and indexing. A good counterargument to *that* is obviously that this paper would have been unsearchable with TDM at the start of the outbreak when people were combing through all West African Ebola literature.

" 'green' model of OA adoption" (page 13, col2)
You generally refer to Green and Gold routes, with the colours capitalised. Just a small terminology thing to keep consistent.

A much more important thing is also on page 13, col2:
"The pay-to-publish system is a potentially greater burden for authors in **developed** countries,

considering that they are not used to paying publication costs, and funding systems for OA are not as well-established as those in the Western world."
--> developing countries, not developed countries!

Predatory publishers (page 14, col1)
I agree with Ross Mounce's comment on the paper: you give Beall too much importance. I think it can be a useful list and should be mentioned, but definitely include some caveats like the ones Ross writes, or the fact that he added Frontiers to the list because of a couple of editorial mistakes.

Conclusions

Peter Suber (page 16, col1)
You describe him as "a leading voice in the OA movement", but I think you should write what his positions are (see http://cyber.law.harvard.edu/~psuber/wiki/Peter_Suber) in order to justify his importance.

Other general things

There are no proposed solutions in this paper, which is totally fine, because it's beyond the scope of the paper. I feel it could benefit by putting in a couple of sentences here and there about who is needed for driving this change: academics, funders, governments, etc.

I was disappointed not to see anything about the Dutch government and university library organisations' collective drive towards OA. They've changed the national law on copyright, they've reached agreements with most major publishing groups, they may well introduce mandatory OA publishing in the Netherlands in 2016, and they've made it one of the main priorities of their EU presidency this year. It's like the best example of how a whole country can take the lead and sort it out. I think including a quick reference to the Netherlands as an example of excellent OA policy (in the same way that you mention sciELO in Latin America) would go a long way towards convincing the people who are reading this thinking, "ah, yes, I guess OA makes sense in the developing world, but we're doing fine here in the West and it would be too difficult to change things". A good summary of that is here:
http://openaccess.nl/en/in-the-netherlands/current-situation

Final remarks

That's the end of my 2800-odd word review. I really enjoyed reading this paper, going through it, and trying to find ways to improve it. Thanks to the authors for writing an excellent paper.

References

1. Brembs B, Button K, Munafò M: Deep impact: unintended consequences of journal rank.*Front Hum Neurosci*. 2013; **7**: 291 PubMed Abstract | Publisher Full Text
2. Swanson D: Two medical literatures that are logically but not bibliographically connected. *Journal of the American Society for Information Science*. 1987; **38** (4): 228-233 <228::AID-ASI2>3.0.CO;2-G">Publisher Full Text

I have read this submission. I believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.

Competing Interests: Non-financial: I'm an Open Access advocate, and so I have a vested ideological interest in seeing papers like this succeed and reach wide audiences. However, if anything, I believe that makes my peer review more critical, as I want this paper to be the best paper it can be.

Discuss this Article

Version 2

Author Response 11 Jul 2016

Jon Tennant, Imperial College London, UK

Dear Ruth,

Many thanks for pointing out that typo. It looks like just a keyboard malfunction of some sort! We'd be happy to fix it in future versions of our paper.

Competing Interests: No competing interests were disclosed.

Reader Comment 08 Jul 2016

Ruth Lewis, Biology Library, Washington University, USA

Very useful review! Pardon my ignorance but I don't understand the fggd in Open Accessfggd policies - Figure 3?

Competing Interests: No competing interests were disclosed.

Author Response 17 Jun 2016

Lauren Collister, University of Pittsburgh, USA

Elizabeth,

Thank you very much for your response and your focus on copyright. Unfortunately we did not have enough time to devote to the fine issues of copyright and OA and hope that others will take up the opportunity to do that. However, here are some explanations of our thinking on these two particular points:

1. We are aware that academic authors are uneasy about the commercial and derivative uses of their work. However, we don't make any arguments about the behavior of others or the moral rights held by the authors; there simply isn't time in this article to address this issue. Moral rights and CC licenses have been debated for a long time (e.g. Larry Lessig in 2005: http://www.lessig.org/2005/02/on-the-challenge-of-moral-righ/) and even the CC-BY license indicates that "other rights such as publicity, privacy, or moral rights may limit how you use the material" (https://creativecommons.org/licenses/by/4.0/). I've come across some interesting

thoughts on this matter, for example this post from Anne Barron (http://blogs.lse.ac.uk/impactofsocialsciences/2013/10/18/open-access-creative-commons-moral-rights/), that give food for thought in this area. I haven't seen any definitive legal decisions that deal with this particular area; if you know of any, we would be very interested to learn about them.

2. The act of applying a CC license indicates that an author is explicitly reserving copyright for themselves while granting certain usage rights to the public, which was the meaning behind our statement. (However I, a linguist, do appreciate the semantics of the verb RETAIN that you bring up here!) Your statement is also correct, in that the focus of CC licenses is on how the users may engage with the item, and I thank you for bringing up this perspective.

Thanks again,
Lauren (on behalf of the authors)

Competing Interests: I am a co-author of this paper.

Reader Comment 16 Jun 2016

Elizabeth Gadd, Loughborough University, UK

I really enjoyed this paper - a great overview of the key themes and issues around open access. Thank you.

I would query two points you make:

1. You write: "The only constraint on reproduction and distribution, and the only role for copyright in this domain, should be to give authors control over the integrity of their work and the right to be properly acknowledged and cited. This definition is broadly equivalent to the Creative Commons Attribution license (CC-BY), which is widely considered to be a standard for OA (creativecommons.org/licenses/)."

   Whilst CC-BY gives authors the right to be properly acknowledged and cited, it does not give authors control over the integrity of their work. A CC-BY licence allows anyone to do anything with the work, even for commercial purposes. This is something academic authors are very uneasy about.

   See: Gadd, E., Oppenheim, C. & Probets, S., 2003. RoMEO Studies 2: How academics want to protect their open-access research papers. Journal of information science. Available at: http://jis.sagepub.com/content/29/5/333.short [Accessed March 4, 2013].

   2) You write: "The majority of 'born OA' journals and publishers do not request or receive copyright from authors. Instead, publishers are granted non-exclusive rights to publish, and copyright is retained by authors through a Creative Commons license (typically CC-BY)."

   When an author signs a non-exclusive agreement with a publisher, copyright is retained by the author by virtue of copyright law, not by virtue of a Creative Commons licence. However, users of the research may engage with it by virtue of a Creative Commons licence. Is this what you mean?

Competing Interests: No competing interests were disclosed.

Version 1

Author Response 06 May 2016

Jon Tennant, Imperial College London, UK

Thank you for the additions to the discussion. We agree that econometric studies can also be subject to debate and are interested in hearing what you think might be problematic about them.

Additionally, in this article we only wanted to stress that the APC model does not necessarily destroy all business and that it does seem to work for some journals/publishers. Nonetheless, you raise a valid point and we will stress in the revised version of the article that not all OA models necessarily sustain the current ecosystem. The discussion you raise about the level of APCs and the necessity of journals as publishing systems are interesting, and we fully agree that they are relevant to the question whether journals survive. We will add this remark, but refrain from providing an answer to these issues, because they are normative and hence do not fit in the scope of this review.

Competing Interests: No competing interests were disclosed.

Reader Comment 30 Apr 2016

David Wojick, David E. Wojick Associates, USA

I am impressed by your authors' detailed response, Jon, but hardly convinced. In particular, I do not regard speculative econometric studies of possibilities as strong evidence. However, I cannot respond to all of this (unless a small grant comes from somewhere).

So just to pick a point, you say "The comment that APCs will be unable to sustain the current journal ecosystem is also not supported by current evidence."

That is not what I said.

My point is that there are several obvious reasons why APCs might not sustain the current journal ecosystem, not that they will not, as your authors claim. (We are talking about possibilities here, not making predictions. At least I am not and I hope you are not.)

The first reason is that switching the journal ecosystem to APCs requires massive systematic changes by millions of people and thousands of institutions. This is unlikely to occur, to say the least. Other scenarios are at least equally possible and these need to be considered. One cannot just say that since the money is there now nothing is going to change, which is essentially what these macroeconomic studies are saying. That is not how the world works.

But the second reason is far stronger. There is a large OA constituency that advocates the transition to a low APC model. Something on the order of a few hundred dollars an article, or less, compared to the thousands of dollars per article funding the present ecosystem.

Clearly it is impossible to cut revenue by an order of magnitude and maintain the same ecosystem. The study we need to see goes like this. Take a typical scholarly society, or a commercial publisher, and look at how it spends its journal revenue. Then cut that revenue by 90% and ask what goes away? Then estimate the academic, economic and social disbenefits of that loss of revenue. Do this for a variety of cases then project the results to the entire journal ecosystem.

Perhaps the authors do not consider the loss of jobs to be an adverse economic impact, although public policy analysis does. The deeper question is what value those labors provide, that is then lost? When I see such studies I will begin to believe that a balance has been struck in the OA discourse.

On the green side there have probably been few subscription losses to date simply because the repository system is a disjointed mess. But surely you do not deny that if everything were freely available then no one would subscribe. Immediate deposit per se is self defeating. It can only be used to force the journals into the APC model.

Embargoed access is far more difficult to assess. I have made a beginning here: https://scholarlykitchen.sspnet.org/2014/01/07/estimating-the-adverse-economic-impact-of-imposed-embargo .

My basic point is that if there are going to be journals then someone has to pay for them. If we pay less we will get less. But note that there is another OA constituency that says that we do not need journals. They advocate simple author posting and post-posting review via comments. I hope your authors do not consider this no journal scenario to be "sustaining the current journal ecosystem," which no longer exists.

As I have said repeatedly, your authors are glossing over the vast differences between the various OA models being proposed and the disbenefits each model entails.

Competing Interests: No competing interests were disclosed.

Author Response 27 Apr 2016

Jon Tennant, Imperial College London, UK

Dear David,

Many thanks for taking the time to provide an extended comment, we, the collective authors, appreciate it. Below, we provide our response to the points raised.

- The link to the Wellcome Trust blog post can be found here: https://blog.wellcome.ac.uk/2016/03/23/wellcome-trust-and-coaf-open-access-spend-2014-15/ - we will amend this in our future revision of this MS.
- Thank you for pointing out the phraseology of the sub-headings. We agree with you that this does appear to inflict bias on our MS, and is a hangover from a previous draft of this review. We will amend this in the next version to more accurately reflect the content of the MS, and also incorporate the comments from the referee reports, especially that of Paige Brown Jarreau, into this.
- You indicate that an “obvious adverse impact is the loss of subscription revenue” regarding green open access. We comment upon this in the first paragraph of the paper: “Some publishers require an embargo period before deposit in public repositories is allowed. These embargoes are applied in

order to avoid putative reductions in subscription income due to such self-archiving." However, empirical evidence fails to support this 'obvious' claim, and we refer you to Berners-Lee et al., 2005; Swan, 2005; Henneken et al., 2007, Houghton et al., 2010, Bernius et al., 2013, and the Committee for Economic Development, 2012), which all indicate that not only is there no loss of subscription income from green open access, but that both can indeed coexist as a sustainable business model for publishers. Furthermore, there is little evidence to support the commonly held inference that shorter or zero-length embargoes will have a detrimental effect on publisher incomes (BIS Select Committee report, 2013-14; but see also Jubb, 2011). According to the RoMEO journals database, only 13% of journals do not permit immediate self-archiving of either the accepted (60%) or submitted (an extra 27%) versions of articles (as of 2011: https://romeo.jiscinvolve.org/wp/2011/11/24/60-of-journals-allow-immediate-archiving-of-peer-reviewed), therefore it is unclear why such a widespread practice would exist for journals if it was detrimental to their finances. Indeed, the only case in which there would be a subscription loss would be if there was no value added throughout the publication process subsequent to acceptance of a peer reviewed manuscript, which is clearly not the case. Therefore, there is little empirical evidence that we are aware of to support the 'obvious adverse impact' suggested, including loss of functionality for scholarly societies and termination of journals. We will amend the text to clarify this point.

- The comment that APCs will be unable to sustain the current journal ecosystem is also not supported by current evidence. We refer to the White Paper issued by the Max Planck Society, which provides evidence for a sustainable global switch from subscription to APC-driven open access publishing (Schimmer et al., 2014). We recognise that these are different economic models, but present evidence suggests that the potential disbenefits outlined cannot be supported.
- The comment that APCs will be taken from research funding, thus reducing the amount of research, is also not strongly supported by current evidence. While we recognise that this is a possibility in some cases, research funders now typically, but not entirely, provide additional funds where OA is required in order to support APCs. The case of the RCUK at a national level in the UK is good evidence to support this (http://www.rcuk.ac.uk/media/news/121108/), as well as funding bodies such as the Wellcome Trust (http://www.wellcome.ac.uk/about-us/policy/spotlight-issues/Open-access/Guides/wtx036803.htm). One could could make the simple counter-argument that the amount spent on subscriptions (estimated at €7.6 billion per year) by university libraries and research organizations is much more detrimental in removing income from the research system than the relatively lower amounts currently spent on APCs.
- We agree with the comment that " those who cannot afford the APCs will not be able to publish". We comment on this in the context of the developing world: "The shift from a 'reader pays' to an 'author pays' model with OA potentially limits its adoption in developing countries. The pay-to-publish system is a potentially greater burden for authors in developed countries, considering that they are not used to pay publication costs, and funding systems for OA are not as well-established as those in the Western world." , and will revise this statement to indicate that APCs are also often currently unaffordable for researchers beyond the developing country to, to better address your concern. We will furthermore note that many journals also have fee waivers for APCs in place for researchers who cannot afford them, including PeerJ and PLOS ONE (not only researchers from developing countries). These issues were discussed in a recent presentation on the economics of OA by Jeffrey MacKie-Mason (summarized here: http://madlibbing.berkeley.edu/economic-thoughts-about-gold-open-access/), which we also plan to incorporate into the MS. This also ties in with the previous comment from Ross Mounce, who suggests that current levels of APCs are not sustainable and can be reduced in order to be more affordable for researchers. We will append the MS accordingly to accommodate this.

- Regarding the comment about APCs being too low and forcing peer review out, it should be noted that peer reviewers are not typically paid anyway, so it is not clear why would they be forced out if there isn't enough APC funding. Moreover, some OA models (e.g., Collabra) now include optional payment for the reviewers, so from a traditional economic perspective one might argue that reviews in these OA journals would improve. Nonetheless, it is more likely that paid copy editors, formatters, and people devoted to the marketing and management of the journal will be more likely to lose out, which is certainly problematic, but does not involve peer review, and we will stress this point in the next version of the MS. This also relates to the previous comment, including that by Ross Mounce, in which it has been demonstrated that the system is fully functional at much lower APC costs. It should be noted that it is presently unclear in which context “unsustainable” is used: this means very different things to publishers, funders, and researchers, for example.
- As for the last comment, we have addressed much of these disbenefits in the text already. Others noted are not empirically supported. However, in each instance we will provide additional context, as noted here, to clarify any additional potential disbenefits. This is most clearly the case in the comment that APCs at their current level are unaffordable to many, which we will be sure to comment more upon as a side effect of current OA policies, and also by incorporating information provided by the other commenters and referees.

We thank you for your thoughtful and constructive comment, and we will make sure to amend the text appropriately to accommodate your points.

References

Berners-Lee, Tim, De Roure, Dave, Harnad, Stevan and Shadbolt, Nigel (2005) Journal publishing and author self-archiving: Peaceful Co-Existence and Fruitful Collaboration. Available at: http://eprints.soton.ac.uk/261160/

Bernius, Steffen, et al. "Exploring the effects of a transition to open access: Insights from a simulation study." Journal of the American Society for Information Science and Technology 64.4 (2013): 701-726. http://onlinelibrary.wiley.com/doi/10.1002/asi.22772/epdf

Henneken, Edwin A. et al (2007) E-prints and journal articles in astronomy: a productive co-existence. Learned Publishing 20 (1), 16-22. Available at: http://arxiv.org/abs/cs/0609126

Houghton, John W., and Charles Oppenheim. "The economic implications of alternative publishing models." Prometheus 28.1 (2010): 41-54. http://www.tandfonline.com/doi/pdf/10.1080/08109021003676359

Jubb, Michael. "Heading for the open road: costs and benefits of transitions in scholarly communications." Liber Quarterly 21.1 (2011): 102-124

Schimmer R, Geschuhn KK, Vogler A: Disrupting the subscription journals’ business model for the necessary large-scale transformation to open access. 2014. Available at: http://pubman.mpdl.mpg.de/pubman/faces/viewItemOverviewPage.jsp?itemId=escidoc:2148961

Swan, Alma (2005) Open access self-archiving: an introduction. Truro: Key Perspectives Ltd. Available at: http://eprints.ecs.soton.ac.uk/11006/01/jiscsum.pdf

Competing Interests: I am the corresponding author for this paper.

Reader Comment 20 Apr 2016

David Wojick, David E. Wojick Associates, USA

Having been asked by Jon Tennant to provide additional comments regarding the potential disbenefits of OA, here are some. These will also appear as the latest issue of my subscription newsletter *Inside Public Access*.

Green OA

The most extreme form of green OA is immediate deposit of the article in a repository. Here the obvious adverse impact is the loss of subscription revenue. The more widespread and discoverable this deposit practice becomes the more subscriptions will be canceled. Basic economics says that one cannot sell what is available for free.

Many OA advocates want subscriptions to end so they may not see this as an adverse impact, but it certainly is from a public policy point of view. The loss of revenue is the very definition of adverse economic impact.

The loss of revenue readily translates into adverse academic and societal impacts, especially when this loss is sustained by a scholarly society. Societies use subscription revenue to conduct various activities that benefit their members directly and society indirectly thereby.

In the extreme case, which might be common, loss of revenue might lead to the termination of the very journals that OA seeks to provide access to. Here the academic and societal disbenefit is loss of the journal, making OA self defeating.

Transition to APC gold OA

OA advocates often argue that loss of subscription revenue will not lead to a loss of journals because the journals will simply transition to an author pays (APC) business model. But this transition is difficult and expensive, so some journals may not make it. Moreover, those that do may not attract the APC revenue needed to continue.

In fact there is no a priori reason to think that a system of all APC journals will duplicate and sustain the present system of subscription journals. The present system is based on libraries buying journals for large numbers of researchers. The APC system is based on individuals paying to be published. These are fundamentally different economic models.

The all APC system might well turn out to be much smaller than the subscription system. If so then the disbenefits will flow from the loss of a significant number of journals. Both recognition and communication might be greatly reduced in an all APC world.

Moreover, even if this does not happen, the APC funds are taken from research funding, thereby reducing the amount of research. Also, those who cannot afford the APCs will not be able to publish, further reducing both communication and recognition.

There is also the very real danger that funders will force APCs down to an unsustainable level. Peer review could be forced out, thereby reducing quality, or journals terminated altogether, if APC limits are set too low.

In fact the use of mandates to get OA has the potential adverse effect of turning scholarly publishing into a regulatory regime. In particular, de facto government control of who gets to publish, and where, is fraught

with danger. Most published research is funded by governments and lowering allowable APCs is politically attractive (as is shortening green embargo periods).

All of these potential adverse impacts of OA are well known. Some are obvious. If the Tennant et all article aims to be an balanced account of the impact of OA then it should include the recognition of these possible disbenefits, in at least as much detail as is given to the possible benefits. Either that or re-title their article to make clear that they are only considering the potential benefits of OA, not including the disbenefits.

Competing Interests: No competing interests were disclosed.

Author Response 18 Apr 2016

Jon Tennant, Imperial College London, UK

Dear David,

Many thanks for your comment. In the revised version of this article, we will pay extra attention to the language used, in particular with the headings to address potential issues with objectivity. While we were careful to provide a balanced argument, and provide numerous references to support these, it is possible of course that we have missed some key studies or counter points. Based on your comments, we would greatly appreciate any insight or links to the proposed counter-arguments so that we can strengthen the discussion throughout the paper.

Best,

Jon

Competing Interests: I am the corresponding author for the paper.

Reader Comment 15 Apr 2016

David Wojick, David E. Wojick Associates, USA

This single sentence on hybrid models seems dismissive and the link seems not to work:
"The motivation behind this could come from the currently available data that suggests that hybrid publishing options offered by traditional publishers, while being of higher cost, provide a much lower overall quality publishing process (blog.wellcome.ac.uk/ 2016/03/23/wellcome-trust-and-coaf-openaccess-spend-2014-15/)."

The hybrid model is an obvious transition route to widespread no-embargo OA, probably the most feasible and possibly the only feasible route.

Competing Interests: No competing interests were disclosed.

Reader Comment 15 Apr 2016

David Wojick, David E. Wojick Associates, USA

The repeated use of the phrase "the .... Case for OA" indicates that this is an OA advocacy document. As such it is an excellent review of the various arguments for OA. The title should be something like "A Review of the Case for OA."

But it is most certainly not a review of the academic, economic and societal impacts of OA, because there is no discussion of the various potential adverse impacts. Every argument presented has well known counter arguments based on adverse impacts. None of these counter arguments are mentioned, except the fact that significant APCs will make it harder for researchers in developing countries to publish. Thus the review is almost completely one-sided.

Another problem is that both the benefits and disbenefits of OA are highly sensitive to the model used to get it. This is not discussed, so the false impression may be given that all of these benefits will flow from every model. In some cases the benefits may be slight. Where they are potentially large the disbenefits may well be larger.

As an independent issue analyst who has studied this issue for years I have no competing interest. See https://scholarlykitchen.sspnet.org/author/dwojick/.

Competing Interests: No competing interests were disclosed.

Reader Comment 14 Apr 2016

Anna Sharman, Cofactor Ltd, UK

This is a really great review - balanced and informative. I learnt a lot from it despite being relatively well informed in this area. I will be recommending it as a resource on open access.

A few small improvements could be made:

1. 2nd paragraph: "A core issue remains: universal or even marginal access to ~75% of articles is not directly possible unless one is in a privileged position to work at an institute which has subscription access to a portion of these articles." - might it be worth mentioning the unlikely alternative, which is that someone has the budget to pay to buy access to all articles they require on a per-article basis? If so, include the fact that not all journals have a per-article price, and of course that these prices add up very quickly so are unrealistic for all but occasional uses.

2. Predatory publishers section: I agree with Ross Mounce that the sentence "Predatory publishers tend to charge low publication fees" could be misleadingly read as suggesting that this is a criterion on which publishers can be judged as predatory or not. Some clarification should be provided here. I also suggest expanding on "not all scholars and advocates agree with the criteria proposed by Jeffrey Beall" - can you provide a link or two to criticisms of his criteria and of his views on open access?

3. Figure 1: I found the "Ratio of the cumulative sum" hard to grasp - is there perhaps a more straightforward way of illustrating this point? Or could it be explained more clearly?

The following are small suggestions for improving wording:

- 1st paragraph: "OA pertains only to documents made available via two main pathways" - delete 'only' as it contradicts 'main'.
- Societal impact section: "Mohammadi et al. (2015) found that the number Mendeley readers" should be "Mohammadi et al. (2015) found that the number of Mendeley readers"

Competing Interests: None.

Author Response 14 Apr 2016

Jon Tennant, Imperial College London, UK

Dear Ross and Philip,

Many thanks for these thoughtful and constructive comments. They will be given full consideration, and integrated into the next version of this manuscript pending the organised peer review process.

Best,

Jon

Competing Interests: I am the corresponding author for this paper.

Reader Comment 13 Apr 2016

Philip Young, Virginia Polytechnic Institute and State University, USA

Thanks for this great overview of open access. I am already finding many of the arguments and references useful in my own work.

In the academic case for OA, I think the case for ordinary academic reuse is understated, particularly in comparison to the extensive description of TDM benefits (which is great). For CC-licensed OA, the everyday decisions about reuse of material in a class, or usage of a figure or passage in a thesis or dissertation are greatly simplified. Arguably this kind of permission-free reuse will benefit more academics than TDM.

In Table 1, for 2012 you mention PeerJ's membership model, but they also offer an APC if the authors choose. I realize you are focusing on innovative models, but it may be misleading to suggest this is the only way they receive payment.

In the section on text and data mining, I am not sure that your reference to "a robust and developing public domain" is completely correct, since "public domain" has a specific legal meaning- perhaps rephrase as "a robust and developing scholarly ecosystem" or something like that. Also in this sentence, isn't it open licensing rather than shifting copyright to stay with author that makes TDM possible?

I have serious reservations about the mention of work made for hire in this section. In the U.S. at least, proposing this would result in a tremendous amount of criticism from faculty groups such as the American Association of University Professors (AAUP) who claim that copyright retention is part of academic

freedom (see their statement in this regard here: http://www.aaup.org/report/statement-copyright). This would also be a problem for faculty who make creative works such as novels, films, artwork, etc. In addition, the argument here is contradicted in the paragraphs above, where you cite that 71% of the research community favors author-retained copyright. It's hard to imagine researchers favoring university-held rather than journal-held copyright. Perhaps faculty could voluntarily grant non-exclusive rights, as for green OA policies, but that is not the same as work for hire.

In the discussion of Freedom of Information requests, you could also cite Bergstrom et al. (2014), Evaluating big deal journal bundles, https://doi.org/10.1073/pnas.1403006111. In the next paragraph regarding pay-to-publish, it may be worth surfacing how authors pay APCs (Solomon & Bjork 2011, http://dx.doi.org/10.1002/asi.21660 see table 4) and the extent to which waivers are available (Lawson 2015, http://dx.doi.org/10.3390/publications3030155).

I also echo Ross Mounce's comment - someone has suggested “deceptive publishing” instead.

Minor issues:

- Suggest specifying "U.S." before National Institutes of Health
- question mark after “potential citation rates”
- Open Journal Systems (OJS) https://pkp.sfu.ca/ojs/ (not Journals System)
- in the conclusion, "number of policies” (not number policies)

Thanks again, great job

Competing Interests: None

Reader Comment 12 Apr 2016

Ross Mounce, University of Cambridge, UK

This is in no way a review of the entire manuscript, I don't have time for that at the moment.

This is merely an attempt at constructive criticism of a particular section, namely 'The case for publishers'.

Reading the manuscript as it currently is, one might come away from this section thinking that the average published research paper costs $3500–$4000 publish. I do not believe this to be true. For instance, at the Journal of Machine Learning Research (JMLR) it costs between $6.50 and $10 per paper - it is a notably efficient journal with technically competent submitting authors that submit manuscripts in formats that are more immediately publishable (e.g. LaTeX), thus Shieber (2012) absolutely must be cited in any discussion of the cost of publishing. Likewise, another estimate is given by Ubiquity Press: 300 is all they need to sustainable publish manuscripts submitted in less convenient formats (such as MS Word documents), so please do cite: http://www.ubiquitypress.com/site/publish/ too. Scale and 'selectivity' (the arbitrary rejection of signifcant percentage of papers that some journals like impose) are also very important factors in the cost of a particular journal's cost per paper. The high cost at Nature is largely down to the cost of rejecting all those papers, not the cost of the papers they actually end up accepting!

You may also want to mention the end-to-end XML based publishing system called ARPHA (Mietchen *et al* ., 2015) as used by the open access publisher Pensoft and a similar(-ish?) system licensed by River Valley. Adoption of these more efficient, integrated publishing workflows will inevitably bring down the real

costs of publishing - it's a madly inefficient workflow at most biology journals at the moment with completely different systems for authoring, peer-review and making the final version / typesetting / proofs.

Mike Taylor also has useful information on the real cost of typesetting here: https://svpow.com/2015/06/11/how-much-does-typesetting-cost/

I'm hugely surprised you don't seem to mention the effect of open access journals on decreasing publication delays. That's also a major value-add for most (all?) stakeholders. Daniel Himmelstein's (2015) excellent data & blog post showing that newer open access journals like eLife and PeerJ are decreasing the average time between submission and publication - faster publication. Furthermore, journals like F1000Research and Research Ideas and Outcomes (http://riojournal.com/) and preprint servers are eliminating the delay before submitted manuscripts are made public to zero days.

Finally, I'd advise extreme caution in adopting (willingly?) Beall's phrase "predatory publishers". You seem to indicate that if a journal charges a low APC it may be a predatory publisher? Would a solo author article at PeerJ then thus make PeerJ a "predatory publisher"?

If you still do insist on including a section on Beall's views (although why? whhyyyyyyy?), please do point out by citation/quotation that he is a controversial figure who thinks (his own words) that "The OA movement is an anti-corporatist movement that wants to deny the freedom of the press to companies it disagrees with".

Beall 2013. The Open-Access Movement is Not Really about Open Access. Triple C

Shieber 2012. An efficient journal https://blogs.harvard.edu/pamphlet/2012/03/06/an-efficient-journal/

Mietchen D, Mounce R, Penev L (2015) Publishing the research process. Research Ideas and Outcomes 1: e7547. doi: 10.3897/rio.1.e7547

Competing Interests: I am one of the founding editors of the open access journal Research Ideas and Outcomes (RIO). I am a Software Sustainability Fellow. I am a Panton Fellow for Open Data in Science I believe open access publishing should be the default I am one of nearly 16,000 academics who publicly boycott Elsevier (http://thecostofknowledge.com/)