 doi:10.20944/preprints201910.0057.v1

Scientific Publishing: Education as the Key Enabler for the Transition to Open Science

Mario Pagliaro[a]

Abstract: Published since the late 1600s, scholarly journals today are the products of a large industry comprised of for-profit and not-for-profit organizations, mostly based in western Europe and North America, whose annual income exceeds $25 billion ($10 billion for English-language scientific, technical and medical journals). Originally created for facilitating scientific communication, the Web in principle makes scientific journals no longer necessary. Yet, in an almost opposite fashion to what happened in retail publishing, the academic publishing industry has further flourished following the advent of the internet. Education of today's students and young researchers, we argue in this study, is the key enabler for the transition to open science.

Keywords: scientific publishing; science journals; academic publishing; open science

1. Introduction

Published since the late 1600s (the world's oldest scientific journal appeared in France, on January 5, 1665 as a twelve page pamphlet called the *Journal des sçavans* [1] followed by the *Philosophical Transactions of the Royal Society of London* published without interruption since March 1665 [2]), today's scholarly journals are the products of a large industry, mostly based in western Europe and North America, comprised of for-profit and not-for-profit organizations (not-for-profit publishers do make profits like any for-profit organizations, the only difference lies in the way these profits are used).

Selected figures show the relevance of this industry whose global annual turnover exceeds $25 billion ($25.2 billion in 2015 [3]. The annual revenues generated from English-language scientific, technical and medical (STM) journal publishing were about $10 billion in 2017 [4], when the number of peer reviewed English-language journals was about 10,000 (out of 33,100 academic journals) with 110,000 people employed, about 40% of which based in Great Britain, Germany, and Holland [4].

The aforementioned journals published over 3 million articles in 2017, but the annual growth rate has increased to 4% per year for articles due to the rising number of researchers [4], mostly originating from China and India.

In one of the first studies to include a discussion on the economics of scholarly publishing, Larivière and co-workers in 2015 found out that the industry is actually an oligopoly in which, the five major publishers in natural and medical sciences in 2013 accounted for 53% of all papers published [5].

In 1973, the share was slightly more than 20% [5].

Nicely explaining the unique nature of the scientific publishing market in which consumers (scholars) are isolated from the purchase because purchase and use are not directly linked (and thus price fluctuations do not influence demand) the latter study was published in a so called "open-access" (OA) journal published online by a nonprofit science, technology and medicine publisher charging authors a fee (Article processing charge, APC) per article, billed upon acceptance [6].

The aforementioned $10 billion revenues of STM journal publishing, indeed, chiefly originate from subscriptions paid by universities and research institutions to enable access to otherwise "pay-walled" research articles.

Developed since the mid 1990s, shortly after the introduction of the World Wide Web in 1991, the main alternative economic model for scholarly publishing is based on OA journals in which authors of accepted articles are charged with an APC which can vary from a few hundred dollars (for example $750 for publishing in *ACS Omega* [7]) through $3500 for publishing in *PLOS Biology* [6] or $5,000 in *Advanced Science* [8].

Publishing scientific papers on the Web eliminates the need for printing journals and disseminating them via post to subscribers across the world. Perhaps not surprisingly, most scientific journals are no longer printed. Rather, journal issues and issue journal covers (generally sold by publishers to the authors of selected paper), are currently published on the Web in different formats including hypertexts in HTML (hypertext markup language), PDF (portable document format) and ePub (open e-book standard format).

Customers wishing to receive a printed copy of a journal's issue are billed with its cost, and the selected journal's issue printed "on demand" is sent via post.

"Everything points to the fact" wrote in 2014 Bartling and Friesike introducing one of the most complete (open-access) books on open science, "that we are on the brink of a new scientific revolution" [9].

Seven years before, a group of major science journal publishers had hired a "public relations" (PR) agent "to combat the open access movement" [10] willing to make scientific articles freely accessible on the internet.

[a] Dr M. Pagliaro
Istituto per lo Studio dei Materiali Nanostrutturati, CNR
via U. La Malfa 153
90146 Palermo (Italy)
E-mail: mario.pagliaro@cnr.it
Web: www.qualitas1998.net

"We're like any firm under siege" [10] commented the manager of a publishers association organization. "It's common to hire a PR firm when you're under siege" [10].

Said "siege" apparently had little effects if a scholar based in Canada, commenting in late 2019 his refusal to write for free another book chapter, emphasized on a social network how:

> «The science publishing industry which charges us to publish papers, and then charges us again for content access (though our universities or personal licenses), and not compensating us as associate editors, or being a reviewer, or to write book chapters as an expert, has to change [11]»

In an almost opposite fashion to newspapers and magazines (retail publishing), the scientific publishing industry not only has not been financially hurt by digitization process followed by the widespread adoption of the internet since the mid 1990s but actually, due the dramatic production costs reduction seen above, greatly benefited from the advent of the internet.

Will this be the case also in the course of the next decade? Expanding the education of today's students and young researchers to include the topics of scholarly communication through scientific publishing in the digital era, we argue in this study, is the key enabler for the transition to open science.

2. Citations, Tenure and Open Access

Noting how progress to open access recently stalled, with only 20% of new papers being published as OA articles, Green has lately called for a true digital transformation of scientific publishing [12]. Rather than using information technology to simply digitize the scientific publication processes, said transformation "is about changing the way you work and designing processes using internet-era principles to deliver value" [12].

Guiding OECD Publishing, namely the publishing agency of the Organisation for Economic Cooperation and Development (OECD), Green has pioneered OECD iLibrary, a platform that disseminates OECD work to academic and research institutions across the world.

The main reason for which OA journals did not replace pay-walled journals is likely due to the fact that only few of them reached high impact factor (IF) values. Some actually did but, almost invariably, they are those with the highest article publication charges (one of the exceptions being *Chemical Science*, a journal "free to read and free to publish with no APCs" with an IF of 9.556 in 2018 [13]).

The impact factor, however, is a worthless criterion to forecast the impact of a specific article. For example, up to 75% of the articles in any given journal has lower citation counts than a journal's IF [14].

Still, a recent study on the use of the journal impact factor in academic review, promotion, and tenure evaluations at North America universities found that 40% of universities granting the PhD degree explicitly mentioned the impact fact, or closely related citation-metrics, in their review, promotion, and tenure documents [15].

The picture was confirmed the editor of a prestigious European chemistry journal in a plea to authors calling for a better and fairer use of citations:

> «Impact Factor, and similar citation-based metrics, are likely to prevail and will continue to be used for evaluating not only journals but also researchers who publish in them [16]»

Until the academic reward system across the world will continue to evaluate scholars based on the IF of journals in which they publish, young researchers will continue to prefer publishing their work in high IF journals, even though open-access journals have significantly more citations overall compared to non-OA journals, whose IF does not correlate with citations [17].

In closer detail, studying a sample of 100 OA articles and 100 non-OA articles randomly selected among the 3,742 randomized controlled trials published in the international literature in January 2011, a team of scholars in south east Asia could lately confirm that, whereas the IF shows moderate correlation with citations for articles published in non-OA journals, the IF does not correlate with citations for OA journals [17]. The conclusions were clear: it is better to publish in an OA journal for more citations, and it is not worth paying high APCs for higher IF journals, because gain in terms of increased number of citations will be minimal.

Yet, as noted by Green only 20% of new papers are currently published as OA articles [12]. Professors guiding the work of doctoral students and post-doctoral scholars continue to prefer to submit their team's papers to high IF journals because both their own promotion and that of their tenure-seeking post-docs is driven by the journal's impact factor.

Feeding papers to scientific journals having high IF values, this academic closed-loop system explains why most scientific journals (and academic publishers) were not impacted by the advent of the internet.

All would remain unvaried for the next decades if it not were for the unexpected recent expansion to other scientific disciplines of an alternative until then used by physicists, mathematicians, and computer scientists only: the preprint.

3. A disruptive innovation?

Cutting time to publication, establishing priority, and eliminating subjective assessments of significance or scope, preprints enable scholars to publish to outcomes of scholarly work immediately after its completion.

Routinely used and cited by physicists, astronomers, computer scientists and mathematicians since the launch of arXiv in 1991 [18], preprints are now commonly used also by biologists and life scientists [19]. Eventually, their slow uptake by chemists, forecasted to inevitably accelerate in a 2017 study [20], has lately recorded its first inflection point [21].

As mentioned above, driven by the existing citation-based metrics career advancement system, both tenure-seeking young scholars and tenured professors ("principal investigators" in the research jargon) are chiefly interested in getting citations to their research articles.

Preprints – research articles permanently published online with a digital object identifier (DOI) – are rapidly becoming highly cited scientific documents. For example, a recent regression analysis of preprints posted in bioRxiv revealed that bioRxiv preprints are directly cited in journal articles, regardless of whether the preprint has been subsequently published in a journal [22]. Furthermore, bioRxiv preprints are also shared online widely, particularly on Twitter and in blogs.

Now, the main service offered by scientific journals to author submitting their work is peer review. Manuscript editing and formatting services offered by publishers today are far less important. Accustomed with digital technology, most today's scholars are able to self produce elegant versions of their studies using for example article templates and word processing software freely available online.

Peer review is provided for free by scholars on a voluntary basis following a request from a journal's editor to review a manuscript. The aim is to provide authors with a critical and constructive review helping them in improving their work prior to publication.

The process, however, lacks transparency because reviews are usually not published and reviewers remain anonymous, leading numerous scholars to propose open peer review in which reviews are published online and even rated by peers [23].

Several journals today publish online the names and affiliations of the reviewers next to the published article (for example, the *Frontiers* OA journals) or even the reports of the anonymous reviewers (for example when authors opt for open peer review, *Proceedings of the Royal Society A* publishes reviewer reports, the substantive part of decision letter after review, and the associated author responses).

Similarly, authors of preprints are allowed to post online at any time a revised version of their original preprint incorporating changes suggested by other scholars usually received via e-mail or directly using the Comments form at the bottom of the web page presenting the preprint (a truly useful tool, for example, of Preprints.org).

In principle, it is already possible for a young tenure-seeking candidate to present a list of publications including highly cited preprints in place of peer reviewed articles.

If scientific quality is associated with citations from peers, why should a selection committee or a funding agency make a difference between a citation to a preprint and another to a publication in a peer reviewed journal?

Indeed, the authors of the 2018 report of a scientific publishers association recently emphasized in the conclusions how [4]:

> «There is some concern that preprints (which can be brought up to date) may become a go-to place for the version of record, undermining publisher business models.
>
> «Concerns have also been raised over the loss of citations from journals to preprints servers, with well over 8,000 citations to bioRxiv reported on Web of Science».

4. Educating young researchers

To enable the transition to open science, today's science, engineering, and medicine students and young scholars need to receive updated and proper education on scientific communication and scientific publishing in the digital era.

First, young researchers need to understand the new and central relevance of preprints both for their own career, and for society, as preprints dramatically speed up the time to publication, and allow scholars to regain control of their own research work [18,19,20,22].

Studies first published as preprints are received and evaluated by the broad scientific community based on their own intrinsic quality, independent of the hosting publication platform (*i.e.*, scientific journal), and without the need to pay any charge, neither to publish the preprint nor to access it.

Second, students and junior researchers need to learn to edit their own work, namely acting as journal editors starting from answering the main question lately suggested by the editor of a prestigious catalysis journal: is this research meaningful? [24]. After that, prior to publish their work, they will carefully check the title, abstract, introduction, conclusion, graphics and references [24].

Acquiring the latter skills requires the direct involvement of the senior scholars supervising their students and post-docs.

Numerous scientists across the world offer such education to their students. Suffice it to mention here Rothenberg's and Lowe's 'Write it Right' workshop on writing research articles, funding proposals and technical reports in English offered since 2002 [25]. Attended so far by more than 1600 people, the course has become part of the educational program of the Netherlands Organization for Scientific Research (NWO).

Third, young scholars need to know more closely science evaluation practices and citation-based metrics such as the *h*-index or the age-normalised *m* quotient [26].

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 6 October 2019 doi:10.20944/preprints201910.0057.v1

Fourth, students need to be trained on how to effectively use social media, such as for example Twitter and ResearchGate, to share the outcomes of their research and start interacting with researchers within and outside their research field [27].

Fifth, scholars usually unfamiliar with copyright legal aspects, need to regain control on their work using the Scholarly Publishing and Academic Rights Coalition Author Addendum [28]. The latter is a legal instrument that modifies the publisher's agreement and allows authors to keep key rights to their research, "instead of blindly giving it away to publishers" [29].

5. Outlook and Conclusions

Trying to answer the question why the Web, created by Berners-Lee in 1991 "to disrupt scientific publishing" [30] actually did *not* radically change it in the course of the subsequent two decades, Clarke in late 2016 concluded that this was mostly due to the journal's role of "designation", namely a *cultural* function ("the hardest to replicate through other means") for which, based on a scientists' publication record in existing scientific journals of high reputation (*i.e.*, impact factor), academic institutions and funding agencies base career advancement and award decisions [30].

Thanks to the advent of preprints in all main scientific fields beyond physics, the number of citations and thus the impact of a scientist's work is now independent of the publishing platform.

For the very first time, in principle a scientist can become a highly cited scholar without having her/his work published in conventional scientific journals following submission and peer review.

Having received no formal training on scientific publication, most today's scholars are unaware that "the practice of sending manuscripts to experts outside of the journal's editorial offices for review was not routine until the last half of the 20^{th} century" [30].

The seminal article in which Mullis published the discovery of the polymerase chain reaction (PCR) was rejected by *Science* [31] and eventually was published in *Methods in Enzymology* [32]. Three years later *Science* proclaimed PCR Molecule of the Year, emphasized Mullis in his Nobel Lecture [31]. The impact factor of *Methods in Enzymology* in 2018 was 1.984, even though Mullis' article up to September 2019 had been cited 7876 times (Google Scholar).

"We now receive many more interesting papers than we can publish…" typically reads the e-mail (a letter when Mullis in 1986-1987 repeatedly submitted his revolutionary molecular biology work [32]) with which an editor communicates her/his decision to not send for review a manuscript.

"Hence, we regret that we are unable to process your manuscript further and suggest that you consider submission to a more specialized journal."

The time and money-wasting publication cycle of conventional scientific publication begins, with the manuscript going from a journal to another, and peer review starting each time afresh. Eventually, after months or even years, the manuscript is accepted for publication.

Whether or not manuscripts are selected on the basis of "methodological rigor, novelty, quality of writing, and general significance", until the advent of preprints, there was little or nothing scholars could do to object the subjective opinion of journal's editors and reviewers.

Now, posting online her/his preprint, an author or a team of authors is granted priority on the discoveries and the novel ideas reported therein. Novelty is established, and any media embargo until the date of online publication for which prior to embargo lifting there can be no public mention of the upcoming paper becomes unnecessary.

As mentioned in the introduction, global scientific output is increasing at fast pace, and is expected to double approximately every nine years [33]. This huge scientific output - most of which will shortly originate from China and India due to their huge population (amounting to 37% of world's current population) *and* excellent scientific schools - will appear first as preprints.

Academics, like other professionals, "work to the system in which they find themselves" [34]. Hence, argued Laurillard calling for a better academic system facilitating and rewarding excellence in teaching and not only in research, promotion of excellence in teaching requires to change the rewarding system.

Said reformed academic system will comprise the ability to evaluate scholars for hiring and career advancements based on true scientific merit, and thus also on citation-based metrics, independent of the publishing platform in which the candidate's papers have been published.

I agree with Clarke [30]: "designation" of a scientist by scientific journals is a *cultural* trait. The etimology of a word is most often revealing. The word "submission" originates from the Latin verb *submittere* and was first recorded in English around mid 15^{th} century with a clear meaning: "humble obedience" [35].

As we enter the third decade of the 21^{st} century, the time has come for world's scholars to replace journal article "submissions" with a free and open system widely based on freely accessible and freely reproducible preprints in which the value (*i.e.*, quality) of scientific articles no longer needs peer review but is open to the evaluation and use of the whole scientific community which has all the tools and an obvious incentive to separate "wheat from chaff", identifying quality work for further valorization via subsequent utilization (and citation).

The time foreseen by Évariste Galois in 1831 during which "*on s'associera pour étudier, au lieu d'envoyer aux académies des plis cachetés, on s'empressera de publier ses moindres observations pour peu qu'elles soient* *nouvelles, et on ajoutera: 'Je ne sais pas le reste'* [36]" has finally come.

List of abbreviations

APC = Article Processing Charge
DOI = Digital Object Identifier
ePub = Electronic Publication
HTML = Hyper Text Markup Language
IF = Impact Factor
OA = Open-Access
OECD = Organisation for Economic Cooperation and Development
PCR = Polymerase Chain Reaction
PDF = Portable Document Format
PR = Public Relations
STM = Science, Technical and Medical

[1] H. Brown, History and the Learned Journal, *J. Hist. Ideas* **1972**, *33*, 365–378.
[2] By H. Oldenburg, who acted as editor and publisher. See: M. Hunter, Promoting the new science: Henry Oldenburg and the Royal Society, *Hist. Sci.* **1988**, *26*, 165-181.
[3] Enago Academy, The World of Academic Publishing, May 21, 2018. See at the URL: www.enago.com/academy/the-world-of-academic-publishing.
[4] International Association of Scientific, Technical and Medical Publishers, *The STM Report*, Fifth edition, October 2018. See at the URL: www.stm-assoc.org/2018_10_04_STM_Report_2018.pdf.
[5] V. Larivière, S. Haustein, P. Mongeon, The Oligopoly of Academic Publishers in the Digital Era, *PLoS ONE* **2015**, 10(6): e0127502.
[6] PLOS One, Publication Fees, **2019**. https://journals.plos.org/plosone/s/publication-fees
[7] ACS Publications, ACS Omega, Overview, **2019**. See at the URL: https://pubs.acs.org/page/acsodf/submission/authors.html
[8] Wiley, Advanced Science, Article Publication Charges, **2019**. See at the URL: https://onlinelibrary.wiley.com/page/journal/21983844/homepage/2749_charges.html
[9] S. Bartling, S. Friesike, Towards Another Scientific Revolution In *Opening Science*, S. Bartling and S. Friesike (Ed.s), Springer, Berlin: **2014**; pp.3-15. The open-access book is accessible at the URL: https://link.springer.com/book/10.1007/978-3-319-00026-8
[10] J. Giles, PR's 'pit bull' takes on open access, *Nature* **2007**, *445*, 347.
[11] T. Stellingwerff, Twitter, 15 August 2019. See at the URL: https://twitter.com/TStellingwerff/status/1161816793956012032
[12] T. Green, Is open access affordable? Why current models do not work and why we need internet-era transformation of scholarly communications, *Learn. Publ.* **2019**, *32*, 13-25.
[13] RSC Publishing, Chemical Science. See at the URL: www.rsc.org/journals-books-databases/about-journals/chemical-science
[14] V. Larivière, V. Kiermer, C. J. MacCallum, M. McNutt, M. Patterson, B. Pulverer, S. Swaminathan, S. Taylor, S. Curry, A simple proposal for the publication of journal citation distributions, *bioRxiv* **2019**, 062109.
[15] E. C. McKiernan, L. A. Schimanski, C. Muñoz Nieves, L. Matthias, M. T. Niles, J. P. Alperin, Use of the Journal Impact Factor in academic review, promotion, and tenure evaluations, *PeerJ Prepr.* **2019**, 7:e27638v2.
[16] H. Ross, Citation Bribery, *Chem. Eur. J.* DOI: 10.1002/chem.201903777.
[17] S. K. Chua, A. M. Qureshi, V. Krishnan, D. R. Pai, L.a B Kamal, S. Gunasegaran, M. Z. Afzal, L. Ambawatta, J. Y. Gan, P. Y. Kew, T. Winn, S. Sood, The impact factor of an open access journal does not contribute to an article's citations, *F1000Res.* **2017**, 6:208.
[18] P. Ginsparg, Preprint Déjà Vu, *EMBO J.* **2016**, 35, 2620-2625.
[19] R. J. Abdill, R. Blekhman, Meta-Research: Tracking the popularity and outcomes of all bioRxiv preprints, *eLife* **2019**, 8:e45133.
[20] P. Demma Carà, R. Ciriminna, M. Pagliaro, Has the time come for preprints in chemistry?, *ACS Omega* **2017**, *2*, 7923-7928.
[21] For example, ChemRxiv hosted 1,067 preprints in its first 17 months by the end of January 2018. Its owners had planned for the end of 2020 to have a total of 5,600 preprints, a figure that at current posting rate will be reached by the end of November 2019. See: J. Winter, Magical and Malleable Research Stories: the ChemRxiv platform, alchemy.ie, September 13, 2019. See at the URL: www.alchem.ie/blog/magical-and-malleable-research-stories-the-chemrxiv-platform-a-conversation-marshall-brennan
[22] N. Fraser, F. Momeni, P. Mayr, I. Peters, The effect of bioRxiv preprints on citations and altmetrics, *bioRxiv* **2019**, 673665.
[23] J. M. Wicherts, R. A. Kievit, M. Bakker, D. Borsboom, Letting the daylight in: reviewing the reviewers and other ways to maximize transparency in science, *Front. Comput. Neurosci.* **2012**, 6:20
[24] M. A. Rowan, Are Editors Real People?, *20th IUPAC International Symposium on Organometallic Chemistry Directed Towards Organic Synthesis – OMCOS 20*, Heidelberg, 21-25 July 2019.
[25] The standard Write it Right workshop involves twelve hours of training spread over two days. Lectures occupy six hours, exercises related to the lecture content occupy five hours and one hour is for discussion. See at the URL: http://write-it-right.org.
[26] R. Ciriminna, M. Pagliaro, On the use of the h-index in evaluating chemical research, *Chem. Centr. J.* **2013**, 7:132.
[27] M. Carrigan, *Social media for academics*. SAGE Publications, London: **2016**.
[28] Scholarly Publishing and Academic Resources Coalition, Author Rights: Using the SPARC Author Addendum, **2019**. See at the URL: https://sparcopen.org/our-work/author-rights/brochure-html/
[29] J. Tennant, Scholarly publishing is broken. Here's how to fix it, aeon.co, July 17, 2018. See at the URL: https://aeon.co/ideas/scholarly-publishing-is-broken-heres-how-to-fix-it
[30] M. Clarke, Revisiting: Why Hasn't Scientific Publishing Been Disrupted Already?, *The Scholarly Kitchen*, October 26, 2016. See at the URL: https://scholarlykitchen.sspnet.org/2016/10/26/revisiting-why-hasnt-scientific-publishing-been-disrupted-already
[31] K. B. Mullis, The Polymerase Chain Reaction, *Nobel Lecture*, December 8, 1993. See at the URL: www.nobelprize.org/prizes/chemistry/1993/mullis/lecture
[32] K. B. Mullis, F. A. Faloona, Specific synthesis of DNA in vitro via a polymerase-catalyzed chain reaction, *Methods Enzymol.* **1987**, *155*, 335-350.
[33] L. Bornmann, R. Mutz, Growth rates of modern science: A bibliometric analysis based on the number of publications and cited references, *arXiv* **2014**, 1402.4578 [cs.DL].
[34] D. Laurillard, Rethinking Teaching for the Knowledge Society, *Educause Review* **2002**, *37*, 16-25.
[35] Online Etymology Dictionary, **2019**. See at the URL: www.etymonline.com/word/submission
[36] É. Galois, *Préface aux «Deux mémoires d'Analyse pure»*, St. Pélagie, December 1831. See at the URL: https://evariste.readthedocs.io/en/latest/

doi:10.20944/preprints201910.0057.v1