

Open Access Book Publishing and Selection Bias

Marcel Knöchelmann | Marcel.knoechelmann.15@ucl.ac.uk

Department of Information Studies
University College London
119 Foster Court
Gower Street, London, WC1E 6BT

Published on *LePublikateur.de* in October 2018:

<https://www.lepublikateur.de/2018/10/22/open-access-book-publishing-and-selection-bias/>

Gold open access book publishing is prone to selection and competition. The reason for this is the financial demand of the high-priced book processing charges (BPC) on which gold open access book publishing regrettably thrives. The resulting selection bias is harmful for progress towards more equal scholarship. Selection bias weakens the moral imperative of immediate and unrestricted open access. Moreover, that the implications of this selection bias are being promoted as a result of open access is both erroneous and a revealing statement about some stakeholders' value system.

Open Access as a Moral Imperative

It is often argued that the assumptions that drive open access are based on a “moral economy in the context of digital capitalism” (Bacevic and Muellerleile, 2017). There seems to be a moral imperative to making knowledge freely accessible, immediately and unrestricted, and it is academia’s obligation to assure all have access to this knowledge, both within and beyond academia.

Yet, in the context of humanities and social sciences (HSS) book publishing, the notion of a digital capitalism and access to knowledge do not generally apply. Publishing in this realm thrives on a long tail of smaller and often not-for-profit presses which do not normally run on the foundations of capitalism. Not unusually, these entities are part of the academic structures and, though at times being driven by philistine managerialism, aim to support scholarly communication regardless of economic returns.

In addition, it is commonly not knowledge that is published in the HSS as opposed to STM or life sciences. HSS books are gateways to understanding, and, thus, deeply connected to individual authorship and qualitative judgement (Collini, 2012). If the moral imperative is still applied in this realm—in other words, open access being *the right way to publish* for HSS books—open access needs to get rid of a structural selection bias.

Selection bias is the propensity for favouring superior work in gold open access publishing. Especially in book publishing—and, thus, particularly in the monograph-centric HSS disciplines—selection bias is likely to be far-reaching. The extensive costs associated with BPC-based gold open access—revolving around £10,000 with some outliers above and below (Pinter, 2018)—are responsible for this: There is a substantial economic barrier to *access* gold open access and it can often only be overcome by proving some kind of superiority.

Open Access Selection Bias by Means of Research Funding

BPC-based gold open access is commonly chosen when funding for the charges is assured. This assurance often comes from the same entity that requires open access, universities and research funders. Those entities may allow self-archiving (green open access) instead of the cost-intensive BPC-based gold open access. But if the respective university or research funder would pay the charges, the researcher is likely to avoid the inferior self-archiving and choose gold. However, to get this funding in the first place, researchers must get through a competitive process, either specifically for the publication or for the whole research project.

Library and university funds for open access are small compared to other budgets. To get a gold open access book funded by the researcher's institution is, thus, a competitive process. For instance, the University of Reading states that an “an assessment will be made by a small panel”, based on the proposed publication. But: “Priority will be given to monographs which will be published before and submitted as part of the next REF” (University of Reading, n.d.).

The Ideals of Excellence and Superiority

External research funders are often even more competitive, pronouncing their ideals of excellence and superiority. For instance, the Austrian Science Fund states in their monograph funding scheme: “Only excellent scientific publications that meet the current scientific standards of the respective subject area in terms of content and form and which can be expected to broaden scientific/scholarly knowledge significantly or to advance research in the field in question can be funded” (FWF, 2006).

Similarly, the NWO “funds top researchers” (NWO, n.d.) and the Leverhulme Trust supports “research projects of high quality and potential” (The Leverhulme Trust, n.d.). The geographically wider-ranging European Research Council funds “the highest quality research in Europe through competitive funding” (European Research Council, n.d.).

Open Access Selection Bias in Consortial Funding

The largest platform for consortial funding is Knowledge Unlatched. All their about 950 titles (as of Summer 2018) were selected by a Title Selection Committee. Publishers can annually submit titles to Knowledge Unlatched to be included in the KU Select funding scheme. But of those titles, only a few are chosen for the programme.

It can be assumed that publishers would not submit their potential next bestsellers or crossover titles, because publishers would hope to rather sell a larger quantity of those titles (the cannibalisation myth holds on, see for the opposite: Snijder, 2014). Thus, the submitted titles might not be the publishers' top titles of a respective programme. But of those titles submitted,

only 30% make it into the consortial funding programme. This means that a high rejection rate makes Knowledge Unlatched a considerable gatekeeper, and consortial funding a highly competitive procedure.

The criteria behind this process are only loosely described: “The shortlist is made up of the most suitable books and journals for the programme – which are defined as those with the widest global appeal and relevance” (Knowledge Unlatched, 2018). This makes it quite hard to judge how KU steers its role as a key gatekeeper to open access for books. Purely systematically, a rejection rate of 70% means that three out of ten titles need to have more “global appeal and relevance” from the perspective of a few librarians, which will quite likely translate into those three titles also being more appealing and more relevant to be read by the wider audience. If this were not the goal, then the selection could also be a process of random choice.

The Implications of Selection Bias

Based on this, it cannot be consistently articulated what the superiority is that is sought by those entities who bear or distribute the costs of BPC-based gold open access. Especially for HSS publications, this kind of judgement will be hard to generalise. But regardless of the specific criteria, those entities do aim to separate the mediocre from the superior through competitive processes. This has implications for both how the fruits of open access are communicated and the moral argument that underpins open scholarship.

In a way, selection bias opposes the idea to diminish all economic barriers. In an ideal of open scholarship there would be no economic barriers—regardless of whether for authors or readers and without further restrictions like embargo periods. With the current processes of selection, the reality is far from this ideal. There are indeed quite high economic barriers to partaking in many gold open access programmes. To overcome those, researchers have to prove an intellectual superiority—however this is individually specified—to gain additional funding. Thus, if a researcher’s content is deemed worthwhile, she can bypass the economic barrier. Yet, if she works in a niche area, for a not so globally appealing audience, or simply not in the top ranges of scholarship, those economic barriers are fundamental.

Moreover, *excellent* content will be published the *morally right way to be published* instead of the inferior works which will be paywalled in addition to their intellectual inferiority. Surely, those titles can still be self-archived. But only after embargo periods, with extra labour by the author (formatting, layouting), and with diminished discoverability. To top-up this downward cycle, it is argued by some stakeholders that the gold-opened content will also gain many times more downloads and citations. But this reasoning is flawed.

The Opaque Foundations of Open Access Usage Advantages

Open access is said to drive usage and impact. Usage is commonly understood as the combination of downloads, citations, and online mentions. Impact is a diffuse statement in HSS and can seldomly be backed by such hard facts as usage figures. While statements on impact are often plain wrong (Knöchelmann, 2018), statements on usage are often wrongly attributed—or at least, the driving forces of usage can hardly always be attributed to specific aspects (Gatti, 2017).

The mis-attribution means that it *may* not be gold open access that is responsible for more usage, but the fact that the published material is superior to comparable paywalled publications. It is easy to state that free material gets higher download figures than paywalled material. But these figures are often inflated and never result in similar boosts in citations. For instance, Springer Nature titles were downloaded seven times more often than comparable paywalled titles but only one and a half times more often cited (Emery et al., 2017).

The citation figure has supposedly more meaning in a comparison, and this is where a statement of attribution to open access is even harder to justify. For the BPC-based gold open access publication, the content (or the researcher and her project) went through additional competitive funding. Thus, the content is supposedly superior—on whichever criteria—to a comparable and paywalled title. In other words, before publication, the content that is to be published gold open access needs to be deemed more valuable than other content. In return this means that the content alone will be more valuable for readers, regardless of it being open access.

Some of this can also be expressed more formally, as Snijder does in a statistical analysis with rather sobering results (2016): There is more usage, but this does not necessarily translate into a citation advantage. Regarding online discussions especially on Twitter, language and subject seem to be statistically more significant than the publication mode. Moreover, the numbers of citations and online mentions are only weakly associated. Overall, the figures of open access titles are more positive than those of paywalled titles, but even those small advantages are only correlational. A causation to open access can only be assumed—and a selection bias cannot be refuted.

Promoting Openness on the Basis of Selection Bias

Even if there were some degree of causation behind the download or citation – gold open access correlation, using this as reasoning towards more openness impedes the moral logic. It drives competition, puts more pressure on the researcher, and is detrimental to the foundational principles of openness: All knowledge should be open for everyone, not just the chosen few who are deemed worthwhile to get more usage or impact.

More equality through openness must be built on trust in the collaborative effort of sharing scholarship. When Springer Nature or Knowledge Unlatched argue: “We have established that OA books have an increased performance, based on downloads (7 times more), citations (50% higher) and mentions (10 times more)” (Emery et al., 2017), and: “We are delighted to see the KU titles performing so strongly in terms of usage – a clear indicator that the global availability of top quality content from our publishing partners is helping to advance research in the humanities and social sciences” (Knowledge Unlatched, 2017), they are rather driving their own agenda. Making scholarship even more commercially competitive by means of openness cannot be supported with the moral notion of openness.

In addition, real download stats should include figures from pirate sites. This would likely boost the statistics of paywalled content. The best way to avoid that such a boost in download numbers of paywalled content would provide open access opponents with a prolific argument against open access is to avoid promoting open access on the basis of the impact opportunities for the individual in the first place.

Equality Instead of Competition

All of this should not mean that books should not be open access. Only the way stakeholders attribute the success of open books may be erroneous because of an underlying selection bias and, likewise, the principles of open scholarship are damaged especially if such a bias increases.

To alleviate the problem of selection bias, progress in openness should move away from single output-focused financial compensation for openness (the BPC-based gold open access). Library and institutional publishing, compensation of publishers for programmes (being an open access publisher/imprint), and improved infrastructures and support for green open access (diminish the inferiority in green open access) can help reduce the damage of competition through selection bias. Most of all, promoting open access must emphasise collaboration and equality—not competition and erroneous usage and impact statements. There is enough funding in the system (Ferwerda et al., 2017), it just needs to be re-allocated more efficiently.

It is hard to assume a future of open scholarship that is possible based on selection and competition. Likewise, a future that is not based on selection and competition, but equality, would demand a level of collaboration and trust that is unseen today. If we wish to progress towards more equal openness, we must invest heavily in this trust. Otherwise, the movement is stuck in a dilemma: The moral argument on which reasoning for more openness is based currently stands on slippery grounds when this openness leads to more selection and competition. Hence, to make the *this is the morally right way to publish*-argument work, we need to base the concept on the more enduring ideals.

References

- Bacevic, J. and Muellerleile, C. (2017) ‘The moral economy of open access’, *European Journal of Social Theory*, vol. 21, no. 2, pp. 169–188.
- Collini, S. (2012) *What are universities for?*, London, Penguin.
- Emery, C., Lucraft, M., Morka, A. and Pyne, R. (2017) *The OA Effect: How Does Open Access Affect The Usage of Scholarly Books?* [Online]. Available at <https://media.springernature.com/full/springer-cms/rest/v1/content/15176744/data/v3> (Accessed 21 October 2018).
- European Research Council (n.d.) *Mission Statement* [Online]. Available at <https://erc.europa.eu/about-erc/mission> (Accessed 21 October 2018).
- Ferwerda, E., Pinter, F. and Stern, N. (2017) *A landscape study on open access and monographs* [Online]. Available at http://repository.jisc.ac.uk/6693/1/Landscape_study_on_OA_and_Monographs_Oct_2017_KE.pdf.

FWF (ed) (2006) *Application guidelines (including checklists) for the Stand-Alone Publications Funding Programme* [Online]. Available at https://www.fwf.ac.at/fileadmin/files/Dokumente/Antragstellung/Selbstaendige_Publikationen/pub_application-guidelines.pdf (Accessed 21 October 2018).

Gatti, R. (2017) *Handle with Care: pitfalls in analysing book usage data* [Online]. Available at <https://rupertgatti.wordpress.com/> (Accessed 21 October 2018).

Knöchelmann, M. (2018) *Downloads and Impact in Scholarly Book Publishing* [Online]. Available at <https://www.lepublikateur.de/2018/09/25/downloads-impact-scholarly-book-publishing/> (Accessed 21 October 2018).

Knowledge Unlatched (ed) (2017) *Usage of Knowledge Unlatched titles doubled within one quarter* [Online]. Available at <http://knowledgeunlatched.org/2017/10/usage-of-knowledge-unlatched-titles-doubled-within-one-quarter/> (Accessed 21 October 2018).

Knowledge Unlatched (ed) (2018) *How were titles selected?* [Online]. Available at <http://support.knowledgeunlatched.org/ku-select-faq/ku-select-2017/how-were-titles-selected> (Accessed 21 October 2018).

NWO (ed) (n.d.) *Funding* [Online]. Available at undefined (Accessed 21 October 2018).

Pinter, F. (2018) 'Why Book Processing Charges (BPCs) Vary So Much', *The Journal of Electronic Publishing*, vol. 21, no. 1.

Snijder, R. (2014) 'The Influence of Open Access on Monograph Sales', *Logos*, vol. 25, no. 3, pp. 13–23.

Snijder, R. (2016) 'Revisiting an open access monograph experiment: Measuring citations and tweets 5 years later', *Scientometrics*, vol. 109, no. 3, pp. 1855–1875.

The Leverhulme Trust (ed) (n.d.) *Research Project Grants* [Online]. Available at <https://www.leverhulme.ac.uk/funding/grant-schemes/research-project-grants> (Accessed 21 October 2018).

University of Reading (ed) (n.d.) *OPEN ACCESS MONOGRAPH FUNDING REQUEST FORM* [Online]. Available at <https://www.reading.ac.uk/closed/library/lib-oa-monograph-funding-form.aspx> (Accessed 22 October 2018).