



Improving estimates of the total cost of publication by recognising 'APCs paid in the wild'

THEO ANDREW

This paper describes a case study of how the University of Edinburgh gathered information about open access Article Processing Charges (APCs) that were paid by its affiliated academic authors, with particular emphasis on identifying those APCs which were not paid through centrally-managed funds which we colloquially call 'APCs paid in the wild'. We estimate that these costs could account for up to 20% extra in the total cost of publication that is not currently being accounted for. This additional cost is important to take into account when institutions are negotiating fair offsetting agreements for open access publishing.

READ REVIEWS

WRITE A REVIEW

CORRESPONDENCE:
theo.andrew@ed.ac.uk

DATE RECEIVED:
April 20, 2016

KEYWORDS:
open access, article processing charges, offsetting agreements, publishing

© Andrew This article is distributed under the terms of the [Creative Commons Attribution 4.0 International License](#), which permits unrestricted use, distribution, and redistribution in any medium, provided that the original author and source are credited.



1. INTRODUCTION

In recent years there have been calls for increased scrutiny on the amount organisations spend on APCs for Gold Open Access (OA) in academic publishing to make them more transparent. In the UK, the drive towards gold open access is being underwritten by a dozen or so major research funders – Research Councils UK (RCUK) and Charity Open Access Fund (COAF) - who provide central block grants to institutions to help allow their researchers and academics comply with their open access policies. These funders require annual report of costs, so via a standard report template (Jisc Collections, 2015), it is fairly trivial for institutions to gather and aggregate this data to give a picture indicative of APC expenditure (Lawson, 2015). However, this data is not comprehensive as it does not include:

1. Approximately 50 or so Institutions in the UK that do not receive a RCUK grant. In the 2014/15 financial year RCUK funded 107 institutions (RCUK, n.d.) ; However, there are at least 163 universities & higher education institutions listed in the UK (Wikipedia, 2016).
2. APCs paid by authors from other sources. This includes APCs paid by industrial sponsors or local departmental budgets.

This study focuses on the second grouping we are calling 'APCs paid in the wild' – a term we believe was initially coined by Neil Jacobs at Jisc. In our context these APCs paid in the wild are important because the identification of them will lead to a more accurate picture of total APC expenditure for our institution. It will also help identify local grants that can be reimbursed, and potentially identify uncompliant journals who are ineligible for RCUK/COAF funding.

2. FINANCE

The first approach we took to identify APCs paid in the wild was to use data sourced directly from the University's finance systems.

2.1 ACCOUNT CODES

Our initial method was to query and return all transactions with the official finance account code allocated for open access. In theory these transactions can then be traced back to corresponding invoices through which it may be possible to match to an individual article. If this is not possible because matching article information is not available on the invoice, then the associated cost centre and job codes can identify the specific department and research grant the APC was paid from. With the date/publisher information on the invoice and the grant information it should be possible to identify the corresponding article via information hosted on our Current Research Information System (CRIS). The preliminary results suggested that, apart from Information Services who are operating the open access block grants from RCUK and COAF, there is no unified coding rule for APC payments and code selections depend on particular grant, centre or school budgets. Thus, it was not possible for us to identify APCs paid in the wild in this way.

2.2 FOCUS ON SUPPLIER

The second methodology we tried was to collate reports from the Finance department for payments to specific publishers during 2015, and identify those that looked like APC payments by the transaction amount. The corresponding invoices could then be cross-referenced against our APC database and we could remove any duplicates. The remaining APCs paid in the wild could then be linked to papers by correlating the date and finance job code with publication and grant award information held on our CRIS. Unfortunately for us the timing of our work clashed with an extremely busy period in the Finance department and understandably we weren't able to secure the data within the timeframe we required. We hope to revisit this work in the future to see if we can extract any useful information.

3. GETTING THE INFORMATION DIRECT FROM PUBLISHERS

Due to the problems described in the previous section we then decided to test if it is feasible and desirable to gather information about an institutions annual open access publishing profile directly from an online database. To do this we wanted to select a platform that allowed search terms to return a list of articles filtered by time period, affiliation, open access status and publisher (for potential cost analysis).

The largest databases (e.g. Web of Science, Scopus, and Google Scholar) do not offer filtering by open access status or publisher, so for the purposes of this study we selected Elsevier's Science Direct platform. This decision was influenced by the fact that the biggest expenditure from the University of Edinburgh's combined OA block grants was with Elsevier so any search should return the largest pool of results. In 2015 there were 493 APCs paid by the University of Edinburgh's RCUK and COAF block grants, of which 101 APCs were attributed to Elsevier journals which accounts for 20% of the APC which are administered by the Library. This figure is similar to that reported by other institutions and organisations - 21% from the Wellcome Trust open access expenditure in 2014/15 (Kiley, 2016).

3.1 METHODOLOGY

The search [Affiliation = University of Edinburgh, refined by Open Access status = true, item type = article, time period = 2015] returned 348 results on 01/03/2016. The full citation list was exported as a text file and DOIs extracted. The articles were then individually analysed to remove remaining non UoE papers, to identify University of Edinburgh corresponding authors, to determine whether funding was facilitated via the University of Edinburgh (internal) or another institution (external), and to recognise funders in the cases where the article was authored by a University of Edinburgh corresponding author, and the funding was confirmed to be internal. In total, it took one hour to run, filter and de-duplicate the search results, then six hours to analyse 260 individual DOIs.

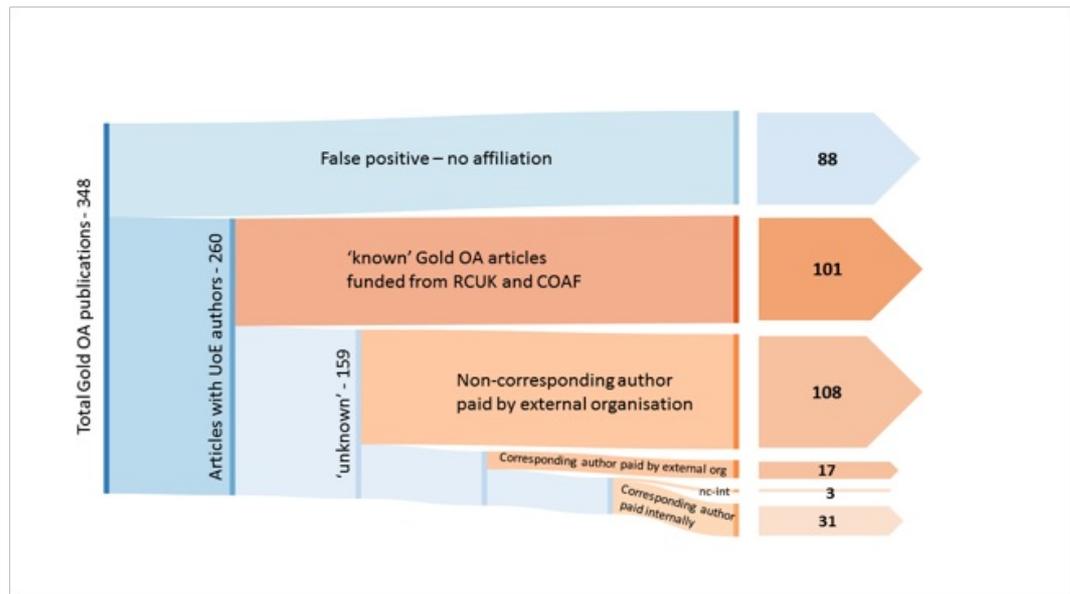


Figure 1: Sankey diagram showing the number of Elsevier open access articles published in 2015 with University of Edinburgh affiliation

3.2 DISCUSSION OF RESULTS

The results are summarised in Figure One and discussed here. From the criteria listed above a number of groupings were identified:

1. Non-University of Edinburgh articles [n=88]
2. 'Known' articles from Open Access block grants [n=101]
3. UoE non-corresponding author and externally funded [n=108]
4. UoE corresponding author and externally funded [n=17]
5. UoE internally funded, corresponding author [n=31], and non-corresponding author [n=3]

GROUP ONE: NON-UNIVERSITY OF EDINBURGH ARTICLES

Unfortunately the affiliation search produced a large number of false positive matches (n=88). The Science Direct platform searches the text in the authors address information to match the input search affiliation terms. Due to the occurrence of the name Edinburgh in the address field it has also picked up authors associated with a wide range of institutions based in Edinburgh, for example Heriot-Watt University, Queen Margaret University, Edinburgh Napier University and Scotland's Rural College (SRUC). These results were noted and set aside.

GROUP TWO: 'KNOWN' GOLD OA ARTICLES

Data was then extracted from four reports previously submitted to research funders - the 2014/15 and 2015/16 RCUK and COAF compliance reports - to give details of all Elsevier Gold OA articles published in the calendar year 2015 that were supported by open access block grants awarded to the institution. This list of 'known' Gold OA articles (n=101) was then removed from the results to give a list of 'unknown' Gold OA articles remaining.

GROUP THREE: APCS PAID BY ANOTHER INSTITUTION

Each of the articles on the remaining list were individually accessed and the corresponding author identified. The Science Direct platform names the major funders who enabled the Gold OA so it is

possible to positively identify where the APC was paid by another institution (n=108). The key identifying criteria being the lead corresponding author is based at another institution and the APC is not on our 'known' gold OA list.

GROUP FOUR: SPECIAL EDITION CONFERENCE PROCEEDINGS

We could further identify the situation where the corresponding author was based at the University of Edinburgh, but the gold open access was facilitated by another group or body (n=17). All of these identified articles were conference proceedings where the entire special issue has been paid to be open access by the conference organisers. These publications do not even consider single articles from individual authors, for example Energy Procedia. This method of funding open access for conference proceedings seems to be becoming extremely popular for subject disciplines like Engineering and Computer Science.

GROUP FIVE: INSTITUTIONALLY PAID APCS

The final group (n=31) consisted of articles where the APC was paid by the University of Edinburgh, but the corresponding author paid for it from sources other than the central block grants that the Library administers. We also identified a sub-group (n=3) where a non-corresponding author paid to facilitate the gold open access. It was possible to spot these because a UK funding organisation visibly paid for the gold open access (e.g. as shown in Fig. 2), but all of the authors were based at foreign institutions apart from the one University of Edinburgh co-author. In this case it is reasonable to assume the Edinburgh author paid for the APC. Together this grouping of previously unknown institutionally paid gold OA articles is referred to as 'APCs in the Wild'.

Table 1: Funder identification

Source of funding	Count
Biotechnology and Biological Sciences Research Council (BBSRC)	3
British Heart Foundation (BHF)	1
Cancer Research UK (CRUK)	1
Chief Scientist's Office (CSO)	4
Department of Health (DoH)	2
Engineering and Physical Sciences Research Council (EPSRC)	6
European Research Council (ERC)	2
Medical Research Council (MRC)	6
Scottish Funding Council (SFC)	2

Wellcome Trust	3
Industry sponsor	1
Unknown	3

The largest portion (59%) of the Elsevier 'APCs in the Wild' group were RCUK (15) and COAF (5) funded. For one reason or another these costs were paid directly from research grants and not from the institutional block grant. With further investigation it may be possible to reimburse these research grants.

Nearly a quarter (24%) of the group was made up by research funders, CSO (4), ERC (2) and DoH (2), that have funds available to researchers - either post-grant awards that researchers can apply for, or budgeted as direct costs in the grant proposal. These costs are currently not tracked by the institution, but it may be possible to configure a report in our Current Research Information System (CRIS) to gather this data.

It is interesting to note that industrial sponsorship of open access seems to be relatively small - we detected only one case of this. In three of the articles it was not possible to determine funding sources without contacting the corresponding author directly.

4. CONCLUSIONS

If we can assume that Elsevier represent approximately 20% of the APC market – which as discussed before is a reasonable figure – then we can speculate on the bigger picture. As such we would expect that there could be as much as 170 APCs paid at the University of Edinburgh in 2015 that we are not currently aware of. This could potentially be as much as £325,000 not being accounted for in the various offsetting agreements for open access publishing being negotiated at the University of Edinburgh alone.

Offsetting agreements are important, and will increasingly become more so, as academic institutions begin to face a rising 'total cost of publication' – this is where you take into account subscription costs plus APCs and other additional administration costs like staffing. Jisc Collections have published a set of 'Principles for Offset Agreements' (Jisc Collections, 2015) which describes a series of recommendations for both publishers and academic institutions to improve and evaluate the offset systems.

To support a managed transition to fully gold open access then all parties need to be aware of all of the costs, which includes APCs 'paid in the wild', which we have illustrated here could account up to 20% more in the 'total cost of publication' than previously thought.

So what can we do about this? Being aware of the problem is a good first step from which we can then act and design better methods to identify and evaluate the true costs involved. We would recommend that when institutions are looking at negotiating open access offsetting deals with publishers that they don't just rely on known costs from the central open access funds that they manage, but that they also try to quantify and include the cost of APCs paid in the wild, which will give them a more accurate picture of the total cost of publication.

REFERENCES

Jisc Collections. (2015). APC data collection. Retrieved April 15, 2016, from Jisc Collections: <https://www.jisc-collections.ac.uk/Jisc-Monitor/APC-data-collection/>

Jisc Collections. (2015). *Principles for Offset Agreements*. Retrieved April 15, 2016, from [https://www.jisc-collections.ac.uk/Global/News files and docs/Principles-for-offset-agreements.pdf](https://www.jisc-collections.ac.uk/Global/News%20files%20and%20docs/Principles-for-offset-agreements.pdf)

Kiley, R. (2016). *Wellcome Trust and COAF Open Access Spend, 2014-15*. Retrieved April 15, 2016,

from Wellcome Trust Blog: <http://blog.wellcome.ac.uk/2016/03/23/wellcome-trust-and-coaf-open-access-spend-2014-1\n5/>

Lawson, S. (2015). 'Total cost of ownership' of scholarly communication: managing subscription and APC payments together. *Learned Publishing*, 28, 9-13. doi:10.1087/20150103

RCUK. (n.d.). *RCUK open access block grant distribution*. Retrieved April 15, 2016, from Esearch Councils UK: <http://www.rcuk.ac.uk/documents/documents/rcuk-apcfunddistribution-pdf/>

Wikipedia. (2016). *List of universities and higher education institutions by size*. Retrieved April 15, 2016, from Wikipedia: https://en.wikipedia.org/wiki/List_of_universities_in_the_United_Kingdom_by_enrollment