Fair, Affordable and Open Access to Knowledge: The Caul Collection and Reporting of APC Information Project

Stephen Cramond
University of Melbourne, Australia, oa.cramond@gmail.com

Cameron Barnes
University of New England, Australia, cbarnes@une.edu.au

Susan Lafferty
Australian Catholic University, Australia, susan.lafferty@acu.edu.au

Virginia Barbour
Australasian Open Access Strategy Group, aoasg.eo@gmail.com

Debbie Booth
University of Newcastle, Australia, debbie.booth@newcastle.edu.au

See next page for additional authors

Presenter Information

Stephen Cramond, Cameron Barnes, Susan Lafferty, Virginia Barbour, Debbie Booth, Karen Brown, Dianne Costello, Kate Croker, Robert O'Connor, Harry Rolf, Tom Ruthven, and Sally Scholfield

FAIR, AFFORDABLE AND OPEN ACCESS TO KNOWLEDGE: THE CAUL COLLECTION AND REPORTING OF APC INFORMATION PROJECT

Stephen Cramond
Formerly University of Melbourne, Australia, oa.cramond@gmail.com

Cameron Barnes
University of New England, Australia, cbarnes@une.edu.au

Susan Lafferty
Australian Catholic University, Australia, susan.lafferty@acu.edu.au

Virginia Barbour
Australasian Open Access Strategy Group, aoasg.eo@gmail.com

Debbie Booth
University of Newcastle, Australia, debbie.booth@newcastle.edu.au

Karen Brown
Monash University, karen.brown@monash.edu

Dianne Costello
Formerly Council of Australian University Librarians, caul@caul.edu.au

Kate Croker
University of Western Australia, kate.croker@uwa.edu.au

Robert O'Connor
Council of Australian University Librarians, caul@caul.edu.au

Harry Rolf
Council of Australian University Librarians, cpo@caul.edu.au

Tom Ruthven
University of Queensland, Australia, t.ruthven@library.uq.edu.au

Sally Scholfield
University of Technology Sydney, sally.scholfield@uts.edu.au

Abstract

Article processing charges (APCs) are fundamental to the business models of many Hybrid and Gold open access (OA) journals. The need to quantify the volume of APC payments paid on behalf of institutional researchers has therefore never been greater. New publishing models will have profound implications for future institutional budgets, and libraries urgently require better information about potential costs and savings.

In 2018, the Council of Australian University Librarians (CAUL) commissioned a project to examine the financial impact of APC payments on universities in Australia and New Zealand. The project aims to develop a methodology for the estimation of APC payments based on data from sources such as Scopus, Web of Science and Unpaywall. In order to test this methodology, the Working Group began a pilot project in February 2019. As part of this pilot, data on publications produced by researchers at six local universities in 2017 were collated and analysed.

This paper will explain the rationale behind the project methodology. It will present the preliminary findings of the pilot, and flag some of the lessons learnt to date. In addition, the

paper will identify future changes. It will be of interest to any librarian concerned with the potential impact of changing publishing models on institutional budgets.

Keywords

Academic publishing, Article processing charges, APCs, Gold open access, Hybrid journals, Open access

Introduction

There is an increasing move towards open access (OA) publishing of scholarly research, driven especially by funder and institutional policies, both nationally and internationally. Article processing charges (APCs) are fundamental to the business models of many OA publishers. Where a journal's sole business model is OA, i.e. all articles in a journal are available as OA, payment of an APC is often a pre-condition for publication. Subscription-based journals that offer an OA option (so-called Hybrid journals) always require the payment of an APC if a paper is to be made openly available. APCs are somewhat controversial. Hybrid journal publishers potentially benefit from two revenue streams: annual subscriptions and APCs [Prosser, 2015]. In such cases, a researcher or institution may end up paying twice for the same article [Björk & Solomon, 2014]. This paper explores the question of how to quantify the sums involved.

The future impact of APCs on library budgets

APC transfers are expected to have a growing impact on academic library budgets [Earney, 2017; Scott, 2018]. Several university libraries in Australia and New Zealand have already established APC funds [Barbour & Anderson, 2017]. Following overseas trends, local libraries have also signed up to offset deals, prepayment schemes, voucher programs and membership discounts with publishers [CAUL APC Working Group, 2018]. As calls increase for publicly-funded research to be made openly available, these arrangements are expected to become increasingly common. Better information about the costs and benefits of different model is urgently required.

In response, the Council of Australian University Librarians (CAUL) added the investigation of APCs to the Fair, Affordable and Open Access to Knowledge Program for 2017–19. A Working Group was formed to investigate the collection and reporting of APC information by Australian universities. The Group's initial task was to review national and international approaches to the gathering of APC payment data. Following this step, the members developed a set of options for collecting information on APC payments by local tertiary institutions. The Group then sought feedback from CAUL member-libraries. Following this consultation, the Group opted to run a pilot project to test a proposed methodology for estimating institutional- and sector-wide APC expenditures.

As part of this pilot, data on 2017 publications by researchers at six local universities were examined in detail. Stephen Cramond, formerly at the University of Melbourne, was engaged to coordinate the pilot and conduct the associated analysis during early 2019. This paper summarises the results of the pilot.

Literature review

The Group carefully considered earlier approaches, both within Australia and New Zealand and overseas. Published and unpublished research on different data collection methodologies were reviewed. Information was also sought from universities in the United Kingdom. None of the published studies provided a complete answer to the problems faced by the Working Group, but their different approaches have informed our own study design.

In the United Kingdom, Jisc (a not-for-profit company in the higher education sector) has played a major role in sponsoring and publishing research into APCs payments. For a number of years, Jisc collected institutional self-reported APCs data as part of its Total Cost of Ownership (TCO) project. An early set of estimates for the APCs paid by 24 universities in the United Kingdom was prepared by Woodward and Henderson [2014]. Another Jisc study published figures for Hybrid OA journals on the basis of data from 23 UK academic libraries [Pinfield, Salter & Bath, 2016]. In 2016, Jisc published one of the largest and most detailed studies of the field, combining TCO figures with financial information from Research Councils UK (RCUK) and the Charity Open Access Fund (COAF) [Shamash, 2016]. Although a significant step forward, this study omits APCs paid directly by researchers. As a result, its totals are partial and incomplete.

The study by Jahn and Tullney [2016], which seeks to provide figures for national-level APC expenditures by German universities, suffers from similar limitations. This study combines self-reported institutional APC data available through the national OpenAPC initiative with article publication data from CrossRef. However, institutional reporting to OpenAPC is entirely voluntary, leading to the possibility of undercounting on national level [Jahn & Tullney, 2016].

A number of papers have attempted to estimate APC transfers indirectly, rather than rely on centrally collected figures. An early example was the study by West, Bergstrom and Bergstrom [2014]. Researchers have sometimes preferred to combine indirect methods with financial data from different sources. For example, Gerritsma [2014] arrived at totals for APCs payments by Dutch researchers on the basis of publishers' list prices and the number of Dutch Gold OA articles appearing in Web of Science and Scopus. Using a similar approach, Henning [2017] calculated APC expenditures at the Swedish University of Göteborg.

Pieper and Broschinski [2018] describe the extended OpenAPC initiative, which collates self-reported data on APCs from over 200 universities world-wide. The resulting dataset can be used to generate estimates of APC transfers at the publisher, institutional and national level. While an invaluable resource for tracking APC payment amounts "in the wild", the limitations of OpenAPC data need to be considered. For example, some Swedish institutions estimate that they only capture data for 50% of their researchers' APC payments.

A local knowledge gap

As the Group continued its investigations, it quickly became apparent just how little was known for certain regarding the amounts paid in Australia and New Zealand each year as APCs. These charges are met from a range of different sources: research grants, personal research accounts, endowment funds, departmental funds, transfers from local or overseas research partners, and the researcher's personal funds [Angelo & Lund, 2014; Barbour & Anderson, 2017]. The results of the Working Group's survey of CAUL and Council of New Zealand University Librarians (CONZUL) libraries in 2018 highlighted the information gap at the institutional level. Sixteen universities in Australia and New Zealand responded to the survey. They reported that:

- 74.4% did not collect data about APC payments;
- 80.85% had no central APC fund;
- 97.87% receive no APC information from funding bodies; and
- 57.45% had no agreements or offset arrangements in place [CAUL APC Working Group, 2018].

As the survey indicated, at least some local universities have centralised APC funds or allocate specific fund codes in their accounting systems for APC payments [CAUL APC Working Group, 2018]. However, no institutions — even those with centralised funds for paying author APCs — have a complete picture of the total amounts paid by their researchers, or the exact sources and destinations of the funds involved. This may seem surprising, but the reality is that existing institutional arrangements record only a portion of transfers.

The best evidence of this comes from the United Kingdom. Even here, central APC funds are rarely the exclusive source of funding by authors, and centrally reported figures are often highly inaccurate. The University of Edinburgh refers to the unfettered payment methods it uncovered as "APCs in the wild" [Andrew, 2016]. The University of St Andrews indicated that, while it has a specific financial code for APCs, few researchers or research administrators use it, while sometimes the code is inadvertently used for other payments such as subscriptions [Proven, 2018]. At some universities, APC funds operate on a "first-come, first-served" basis, and hence not all eligible papers were funded. In the real world, central APC funds operate in a highly uncertain manner.

In Australia, APC funds typically impose a range of criteria for financial support. For example, rules often limit funding to cases where the first named or corresponding author is affiliated with the institution. At a stroke, such requirements exclude many papers co-authored with outside researchers. Other rules limit funding to permanent staff, explicitly excluding casual staff, adjunct and conjoint lecturers, along with visiting fellows, although these researchers frequently generate much of the research published by institutions. Higher degree students are ineligible for financial assistance at different Australian institutions. In addition, some funds do not cover articles submitted to Hybrid journals (for fear of "double-dipping") and many restrict eligibility to authors publishing in top-tier journals.

Another issue is that APCs are increasingly paid by grant providers. Grant bodies frequently include provision for the payment of APCs in research funding (Björk & Solomon, 2014; Monson, Highby & Rathe, 2014]. APCs may also be covered in whole or in part by research partners from outside the institution, particularly those from universities in the United Kingdom and Europe where financial support for the payment of APCs is better established. In such cases, there may be no reason for a researcher to make an application to any central fund.

These problems were highlighted by a recent unpublished study at an Australian university. Library staff sought an accurate picture of institutional APC expenditures by collecting data from a range of sources: CAUL surveys, the university's Finance system, Web of Science, Scopus, and publisher sites. After cross-checking the available data, the study authors concluded that its centrally collected figures were incomplete. Web of Science or Scopus indicated many more OA articles authored by institutional researchers than their central records indicated. A range of account codes and descriptions were used for APC payments reported to the university's finance department. Moreover, it was not possible to identify individual articles or authors based on the descriptions in the university's finance system.

Finally, overseas studies show that many eligible researchers may not even be aware of the existence of a central APC fund [Monson, Highby & Rathe, 2016; Pinfield & Middleton, 2016]. Figures for disbursements from central APC funds therefore have no direct relation to actual institution-wide expenditures. For this reason, the Working Group realised it would be impossible to scale up the figures disbursed by central APC funds at specific universities in Australia and New Zealand to the whole sector. Such a simple approach would be misleading.

The need for a pilot

Faced with the limitations of centrally-reported figures, the Working Group determined that indirect methods of estimation were likely to be the most useful. The Working Group decided to build on work already undertaken at several Australian and New Zealand universities, where staff could see the potential of marrying corresponding author data from traditional bibliographic indexes (e.g. Web of Science) with newer OA finding tools such as Unpaywall, a publicly-accessible database of over 22 million articles. Studies such as ours would not be possible without the rapid growth in coverage, reliability and accuracy of Unpaywall for locating OA status, version and licence type.

As part of the pilot, total APCs transfers for a single year (2017) were estimated on the basis of publication data in Web of Science and Scopus Information from these databases was cross-referenced with data from Unpaywall. This data allowed the Working Group to distinguish between articles made OA through authorial self-archiving of pre- and post-prints (Green OA) and those made available through the payment of APCs (Gold and Hybrid OA). Armed with this distinction, the Working Group then assigned APC costs on the basis of publisher price lists, and other sources which provide aggregated pricing information (such as DOAJ, OpenAPC and FlourishOA).

Initial assumptions

In designing the pilot project, the Group made a number of initial assumptions. Outputs would be assigned to institutions on the basis of the institutional affiliation of reprint authors (Web of Science) or corresponding authors (Scopus). Corresponding authors are typically the principal authors of papers [González-Alcaide, Park, Huamaní, & Ramos, 2017; Mattsson, Sundberg, & Laget, 2011; Weiss, 2012]. For this reason, such authors are most likely to take responsibility for the payment of APCS. It was also decided to exclude records with at least one international corresponding author (approximately 15–16% of papers in the pilot), based on the assumption that the APCs for most of these papers would be paid externally, as grant funding bodies in other countries now routinely make allocations for this purpose.

As Gumpenberger, Hölbling and Gorraiz [2018] point out, multiple corresponding authors for papers poses series methodological problems. Due to the increasing use of fixed term and casual contracts, it is not unusual for local researchers to have multiple institutional affiliations

[Ryan, Connell & Burgess, 2017]. In the case of the single-authored articles in our sample, only 52% of authors had a single institutional affiliation. The remainder (48%) claimed multiple affiliations, often across multiple countries. It was decided to count only the first listed institutional address, as it was assumed that this would be the author's primary affiliation.

The number of corresponding authors in the author lists has grown significantly in recent years (Hu, 2009; Weiss, 2012). This presented particular problems when the different corresponding authors were based at different institutions. There is good bibliometric warrant for the assumption that position in author lists is significant and that, where the first author is also a contributing author, he or she is probably the most important contributor [Mattsson, Sundberg, & Laget, 2011]. On this basis, articles were attributed to the affiliated institution of the first listed co-author.

Where there was doubt regarding the status of an output (as in the case of chapters in monographs-in-series), the presence or absence of an ISSN was used to identify articles. All possible article types were included, as APCs could be charged for any article type under different publishing models. To ensure accurate matching with OA status data from Unpaywall, articles without DOIs were excluded (their omission does not materially influence the outcome of the study). On the basis of these steps, institutional data sets were then de-duplicated and then submitted to the Unpaywall API (Applications Programming Interface). The API returned a list of the articles and their OA status. Articles not available in OA or available only in Green OA (such as those in institutional repositories) were then excluded. The remainder consisted of articles for which an APC may have been paid.

Results

All these assumptions must be regarded as approximations, and it was necessary to calculate a number of simulations based on different scenarios. Each simulation lead to broadly similar outcomes, reflecting the fact that 87% of articles in our sample had a single corresponding author. In each scenario, where there was a single corresponding author, his or her first-listed institution is assumed to have paid the APC. The different scenarios were:

- **Scenario 1**: In instances where there were multiple corresponding authors, the first-listed is assumed to be the corresponding author and the APC is attributed to the first institutional affiliation claimed by that author.

- **Scenario 2**: In instances where there were multiple corresponding authors, the last listed is assumed to be the corresponding author and the APC is attributed to the last institutional address.

- **Scenario 3**: Where there are multiple corresponding authors, but the first and last author both listed the same institution as their first affiliation, it is assumed that institution paid the APC.

- **Scenario 4**: Where there were multiple corresponding authors, from multiple institutions, the APC payment was based on the fraction of each pilot site address as a proportion of the total address count for each record.

The results of these different scenarios are reported in the table below. The names of the institutions have been anonymised.

	Scenario 1	Scenario 2	Scenario 3	Scenario 4
University A	$ 158,588.21	$161.942.32	$158,588.21	$161,842.32
University B	$ 422.684.71	$ 414.947.77	$ 414.947.22	$ 420.856.29
University C	$1,276,903.56	$1,348,756.37	$1,246,232.27	$1,352,391.95
University D	$ 98,369.71	$ 100,369.73	$ 98,369.71	$ 99,369.72
University E	$ 405,662.12	$ 410,312.60	$ 403,974.50	$ 410.229.78
University F	$1,339,996.91	$1,327,270.21	$1,299,767.22	$1,366,536.07
Total	$3,702,205.22	$3,763,499.00	$3,621,879.67	$3,811,226.13

The relatively low variance between the different scenarios give grounds for some confidence in the final figures as minimum estimates for annual total APC expenditures at each institution. In addition, the pilot results were informative in terms of the spread of APC payments. A total 66 publishers were identified in the study. However, only six of these publishers (Springer Nature, Frontiers Media, Elsevier, PLoS, MDPI and Wiley-Blackwell), accounted for 75% of possible payments.

Study limits

This paper provides only part of the picture. Its estimates assume that an APC was paid every time an article was published in a Gold or Hybrid journal, whereas some publishers do in fact offer waivers. In addition, the pilot only reports on articles with DOIs in Web of Science or Scopus. Both Web of Science and Scopus are weak in terms of their coverage of the social sciences and humanities, and in their indexing of Australian and New Zealand academic journals (Harzing & Alakangas, 2016; Mongeon & Paul-Hus, 2016). They also omit many less prestigious OA academic journals patronised by early career researchers and other authors with weaker publication track-records. Although arguably peripheral to mainstream of world research, these less-prestigious titles attract a steady trickle of APCs from local researchers.

Conclusions

The pilot project provides robust estimates for minimum APC expenditures at six local universities. They confirm the suspicion that APC charges are likely to be a significant burden on academic budgets as new publishing models (such as offset deals) gain traction. University librarians would be well advised to alert other institutional decision-makers of this reality. It is crucial not to underestimate the potential impact of future publishing deals. Transparency in relation to total costs (especially APCs) will be critical.

Acknowledgements

Funding for this study was provided by the Council of Australian University Librarians (CAUL) under the Fair, Affordable and Open Access to Knowledge Program for 2017–19.

References

Andrew, T. (2016, April 20). Improving estimates of the total cost of publication by recognising 'APCs paid in the wild'. *The Winnower*. Retrieved from: https://thewinnower.com/papers/4241-improving-estimates-of-the-total-cost-of-publication-by-recognising-apcs-paid-in-the-wild

Angelo, A., & Lund, P. (2014). *An Evolving Business Model for Scholarly Publishing: Exploring the Payment of Article Processing Charges (APCs) to Achieve Open Access*. Paper presented at LIANZA Conference: Pou Whakairo: Connect and Thrive. Auckland, New Zealand. Retrieved from: http://hdl.handle.net/10092/9730

Barbour, V. & Anderson, G. (2017). *Principles for Setting up an APC Fund*. Australasian Open Access Strategy Group.

Björk, B., & Solomon, D. (2014). *Developing an Effective Market for Open Access Article Processing Charges*. Retrieved from https://fwf.ac.at/fileadmin/files/Dokumente/Downloads/Dev_Effective_Market_OA_Article_Processing_Charges.pdf

Council of Australian University Librarians [CAUL] APC Working Group. (2018). [Survey of CAUL and CONZUL members, September 2018]. Unpublished raw data.

Earney, L. (2017). Offsetting and its discontents: challenges and opportunities of open access offsetting agreements. *Insights*, *30*(1), 11-24. doi: 10.1629/uksg.345

Gerritsma, W. (2014) *The Cost of Going Gold in the Netherlands*. Retrieved from: https://www.slideshare.net/Wowter/the-costs-for-going-gold-in-the-netherlands

González-Alcaide, G., Park, J., Huamaní, C., & Ramos, J. M. (2017). Dominance and leadership in research activities: Collaboration between countries of differing human development is reflected through authorship order and designation as corresponding authors in scientific publications. *PloS One*, *12*(8), e0182513. doi: 10.1371/journal.pone.0182513

Gumpenberger, C., Hölbling, L. and Gorraiz, J. I. (2018) On the Issues of a "corresponding author" field-based monitoring approach for Gold Open Access publications and derivative cost calculations. *Frontiers in Research Analytics and Metrics*, *3*(1). doi: 10.3389/frma.2018.00001

Harzing, A. W., & Alakangas, S. (2016). Google Scholar, Scopus and the Web of Science: A longitudinal and cross-disciplinary comparison, *Scientometrics*, *106*(2), 787-804. doi: 10.1007/s11192-015-1798-9

Henning, K. (2017). *Survey of Author Fees at the University of Gothenburg* [Blog post] [Translated using Google Translate]. Retrieved from: https://openaccess.blogg.kb.se/2017/11/15/gastbloggare-karin-henning-kartlaggning-av-forfattaravgifter-vid-goteborgs-universitet/

Hu, X. (2009). Loads of special authorship functions: Linear growth in the percentage of "equal first authors" and corresponding authors. *Journal of the American Society for Information Science and Technology, 60*(11), 2378-2381. doi: 10.1002/asi.21164

Jahn, N., & Tullney, M. (2016). A study of institutional spending on open access publication fees in Germany. *PeerJ, 4*, e2323. doi: 10.7717/peerj.2323

Mattsson, P., Sundberg, C. J., & Laget, P. (2011). Is correspondence reflected in the author position? A bibliometric study of the relation between corresponding author and byline position. *Scientometrics, 87*(1), 99-105. doi: 10.1007/s11192-010-0310-9

Mongeon, P., & Paul-Hus, A. (2016). The journal coverage of Web of Science and Scopus: a comparative analysis. *Scientometrics*, *106*(1), 213-228. doi: 10.1007/s11192-015-1765-5

Monson, J., Highby, W., & Rathe, B. (2014). Library involvement in faculty publication funds. *College & Undergraduate Libraries*, *21*(3-4), 308-329. doi: 10.1080/10691316.2014.933088

Pieper, D., & Broschinski, C. 2018. OpenAPC: A Contribution to a transparent and reproducible monitoring of fee-based Open Access publishing across institutions and nations. *Insights, 31*, *39*. doi: 10.1629/uksg.439

Pinfield, S., & Middleton. C. (2016). Researchers' adoption of an institutional central fund for open-access article-processing charges: A case study using innovation diffusion theory. *SAGE Open 6*(1), 1-18. doi: 10.1177/2158244015625447

Pinfield, S., Salter, J., & Bath, P. A. (2016). The "total cost of publication" in a hybrid open-access environment: Institutional approaches to funding journal article-processing charges in combination with subscriptions. *Journal of the Association for Information Science and Technology*, *67*(7), 1751–1766. doi: 10.1002/asi.23446

Prosser, D.C. (2015). The costs of double dipping. Retrieved from http://www.rluk.ac.uk/about-us/blog/the-costs-of-double-dipping

Ryan, S., Julia, J., & Burgess, J. (2017) Casual academics: A new public management paradox, *Labour & Industry, 27*(1), 56-72. doi: 10.1080/10301763.2017.1317707 S

Scott, A. M. (2018). Article processing charges threaten academic libraries: A librarian's opinion. *Journal of Scholarly Publishing*, *49*(2), 260–266. Doi: 10.3138/jsp.49.2.260

Shamash, K. (2016). *Article processing charges (APCs) and subscriptions: monitoring open access costs*. Retrieved from: https://www.jisc.ac.uk/sites/default/files/apc-and-subscriptions-report.pdf

Weiss, P. S. (2012). Who are corresponding authors? *ACS Nano, 6*(4), 2861. doi: 10.1021/nn301566x

West J., Bergstrom T., and Bergstrom C. (2014). Cost effectiveness of open access publications. *Economic Inquiry. 52*(4), 1315-1321. doi: 10.1111/ecin.12117

Woodward, H., & Henderson, H. (2014). *Report for Jisc Collections on Total Cost of Ownership Project: Data Capture and Process*. Retrieved from https://www.jisc-collections.ac.uk/News/Releasing-open-data-about-Total-Cost-of-Ownership/