Article

Changing the Academic Gender Narrative through Open Access

Katie Wilson 1, *, Chun-Kai (Karl) Huang 2, Lucy Montgomery 2,3, Cameron Neylon 2,3, Rebecca N. Handcock 3, Alkim Ozaygen 2 and Aniek Roelofs 3

1 School of Education, Te Herenga Waka Victoria University of Wellington, Wellington 6140, New Zealand
2 School of Media, Creative Arts and Social Inquiry, Faculty of Humanities, Curtin University, Perth, WA 6845, Australia; karl.huang@curtin.edu.au (C.-K.H.); lucy.montgomery@curtin.edu.au (L.M.); cameron.neylon@curtin.edu.au (C.N.); alkim.ozaygen@curtin.edu.au (A.O.)
3 Curtin Institute for Computation, Faculty of Science and Engineering, Curtin University, Perth, WA 6845, Australia; rebecca.handcock@curtin.edu.au (R.N.H.); aniek.roelofs@curtin.edu.au (A.R.)

* Correspondence: katie.wilson@vuw.ac.nz

Abstract: In this article, we ask whether dominant narratives of gender and performance within academic institutions are masking stories that may be both more complex and potentially more hopeful than those which are often told using publication-related data. Influenced by world university rankings, institutions emphasise so-called 'excellent' research practices: publish in 'high impact', elite subscription journals indexed by the commercial bibliographic databases that inform the various ranking systems. In particular, we ask whether data relating to institutional demographics and open access publications could support a different story about the roles that women are playing as pioneers and practitioners of open scholarship. We review gender bias in scholarly publications and discuss examples of open access research publications that highlight a positive advantage for women. Using analysis of workforce demographics and open research data from our Open Knowledge Initiative project, we explore relationships and correlations between academic gender and open access research output from universities in Australia and the United Kingdom. This opens a conversation about different possibilities and models for exploring research output by gender and changing the dominant narrative of deficit in academic publishing.

Keywords: open access; publishing; gender bias; gender disparities; open research; open science; research productivity; academia; women

1. Introduction

Open Access (OA), defined as openly accessible knowledge and research, has origins in experimental publishing in the humanities and social sciences in the early 1990s and preprint archiving in the sciences [1]. It is also an activist movement exposing critical biases in academic structures and promotional, tenurial systems favouring white, global northern populations dominated by men. In 1990, bell hooks wrote that knowledge could be shared in many ways, drawing attention to the political commitment to achieving such sharing [2]. In the late 1990s, in response to both rising serial and journal subscriptions and barriers to access, scholars and institutions began to challenge the monopoly of commercial publishers, developing alternative means to share research openly. The Open Archives Initiative (OAI) metadata standards and free ‘eprints’ archiving software enabled the development of interoperable and searchable institutional repositories [3]. The Budapest Open Access Initiative [4] produced the first formal open access statement, detailing two options for opening research output: self-archiving and open access journals.

These include the UK Research and Innovation (https://www.ukri.org/our-work/supporting-healthy-research-and-innovation-culture/open-research/) and the Belt and Road Initiative.

Authors have the choice of several paths to sharing research openly. Green open access offers the option to self-archive or deposit a preprint or postprint (author-accepted manuscript or AAM) of an article in an institutional repository or a disciplinary repository, such as the Humanities Commons (https://hcommons.org/ accessed on 2 February 2022), preprint servers such as SocArXiv (https://osf.io/preprints/socarxiv/ accessed on 2 February 2022), or ArXiv (https://arxiv.org/ accessed on 2 February 2022). Practices and understandings of the “state” and the “standing” of preprints within preprint servers or repositories vary by discipline [5]. Commercial publishers may apply embargoes of up to 12 months on depositing or self-archiving in repositories and websites. Three possible options exist for authors who wish to publish openly in a journal without an embargo. Gold open access journals which allow a license to reuse are listed in the Directory of Open Access Journals (DOAJ) and other journals that are exclusively open access. Hybrid open access refers to open publications in a subscription journal with both closed and open articles and a reuse license. Bronze open access refers to journal publications that are free to read, usually the publisher’s decision, but without a defined reuse licence [6]. Some Gold and most Hybrid levy Article Processing Charges (APCs) to publish openly, paid by authors, institutions or funding bodies. The DOAJ includes approximately 70 per cent of journals without APCs. However, APCs can be exorbitant and, although fee waivers may be available, beyond the means of researchers and institutions in many countries. The popular and prominent journal Nature and 32 other titles within the Nature portfolio published by SpringerNature charge €9500, US$11,390 or £8290 per article [7].

The Open Access movement embraces critical thinking and diverse perspectives in academic publishing [2], promoting open knowledge institutions, openness and diversity in research through cultural change. Brabeck [8] emphasises the importance of ethics and the need for transparency and understanding diversity in participation and performance within open science, open data and open access, including developing policies and practices. Our project the Curtin Open Knowledge Initiative analyses research output data and undertakes a critical investigation of the characteristics and performance of global research and the output, openness, collaboration, publishers and funders of higher education research. This includes demographic analysis of institutional workforces to understand the gendered and diversified nature of research output: who is contributing to and producing research [9].

Analysis of research productivity and performance in academia highlights gender disparities and deficits for women compared to men in most disciplines [10]. However, there is limited investigation into the gender balance of open access research [11]. In this paper, we investigate open access research output through a binary gender perspective to understand how this may change the deficit gender narrative around research productivity. The binary approach is a result of the limitations of author name gender disambiguation methods used in the analysis of publications, identifying only names of women and men. We discuss this further in Section 4.

The first section of the paper outlines the methodologies for identifying the relevant literature, the data and the methods used in our analysis of gender-based research performance indicators in universities in Australia and the United Kingdom. Next, we review the biases that underlie and contribute to the deficit narrative. We investigate examples of women’s participation in open access publishing that counteract the predominant narrative.
Finally, we discuss alternative methodologies for analysis undertaken by our project for looking at OA through a gender perspective, including an analysis of academic workforce gender demographics and correlations with research performance in Australian and UK universities.

Research questions:

1. Is there evidence that open access publishing can counteract the predominant negative gender narratives surrounding research outputs?
2. To what extent does a researcher’s institutional context, including the balance of gender within that context, interact positively with open access publishing?

2. Literature Review

2.1. The Prevailing Deficit Narratives Linked to Gender


Gender inequality in higher education continues across the globe, with a slow rate of change over the last two decades [12]. Much attention focuses on STEM (Science, Technology, Engineering, Mathematics and Medicine), but disparities exist in other disciplines, including the Social Sciences, Humanities and Psychology disciplines. A large body of literature documents multiple reasons for the complexities of gender disparities in research productivity, but gender alone and gendered characteristics do not cause the deficit. Institutional procedures with embedded gender biases and behaviours persist [13]. Career promotion and appointment practices based on the institutional measurement and the counting of productivity outcomes, in turn, produce deficit narratives. Assumptions about gender in academia continue the “Matilda effect” of biases that deemed research, writing and discovery by women to be under-recognised and under-credited, especially if married or in collaborations with men. Margaret Rossiter [14] named this after Matilda Joslyn Gage (1826–1898), an American sociologist and feminist who recognised and was subject to such behaviour. Asplund and Welle [15] note the persistence of implicit bias and slow progress in STEM disciplines, despite efforts to change gender and other diversity biases [16]. Hierarchies and inequalities persist in the funding and resourcing of research, with gender often determining academic rank and subsequently, higher rank produces an advantage in funding, collaboration and publication [17] (p. 5). Australian programs to improve the gender imbalance in STEM focus on white women and overlook the bias and microaggressions towards “women of colour” [18]. Faniko, Ellemer and Derks reproduced research into the “Queen Bee phenomenon in academia” investigation), finding women in senior roles self-distancing from “negatively stereotyped” early-career women to enhance their success, a consequence of continued gender discrimination [19] (p. 395).

Structural, organisational and behavioural factors emerge from analyses as contributing to gender productivity differences within the deficit narrative. Bibliometric and scientometric author analysis of research output identifies differential and mixed outcomes for women and men in terms of impact and visibility through publications, citations and altmetrics [20]. Mayer and Rathman [21] analyse publications by tenured, full psychology professors (men and women) in Germany of articles in the top 10 per cent of Web of Science journals ranked by their own methods and book chapters in PSYNDEX to understand the effects of the organisation (size, number of academics, external funding) and individual factors (career age, rank, sub-discipline), collaboration and co-authorship. Controlling for these multivariate factors, they find gender on its own to be a “poor predictor of publication productivity” overall (p. 1677). A meta-analysis of research productivity among men and
women in science identifies multiple factors contributing to gender differences, including self-citation rates, proportional representation of men and women in some disciplines and on evaluation committees [22]. In Sweden, an analysis of researcher productivity finds women’s output is disadvantaged by variables of age, lower academic rank, and positions in research teams, contributing to a “vicious cycle . . . and persistence of the glass ceiling” [10] (italics in original, p. 14).

2.2. The Openness Advantage

At this juncture, we turn away from the depicted gender bias in academia to explore the gender effects of open research dissemination. Open practices include preprints, open access publishing and open data-sharing, providing visibility and options in navigating academic promotion requirements, including for early career researchers [23]. Murphy et al. [24] analysed 879 journal and conference papers tagged as ‘Open Science’ in Microsoft Academic Graph published between 2010 and 2017 in eight fields of academic study: Analytical Chemistry, Bioscience, Computer Science, Engineering, Management, Medicine, Psychology and Statistics. Using the ropenscigender package in R to predict the gender of author names probability and semantic text and network analysis, the authors found a high positioning of women in authorship statements (first or last) in open science publications. Such positioning indicates authors are undertaking more tasks in a publication [25]. Murphy et al. considered the “collaborative, forward-looking focus of open science has the potential to facilitate greater diversity and inclusiveness” through the sharing of data and code to reduce barriers to entry and access and greater collaboration [24] (p. 24160).

Institutional evaluation and research assessment practices continue to promote and require publication in prestigious journals that are ranked using questionable measures such as the journal impact factor [26]. Evidence of the advantages of OA research output in terms of visibility and outreach in the scholarly community and media attention is documented in the literature [6,27,28]. However, awareness and understanding of the process and the benefits of open publishing are not widespread within all academic communities [29–32]. For example, an international study found that 56% of early career researchers interviewed indicated they published OA, but of the publications listed in the interviewees’ Curriculum Vitae, only 8.7% were open [33] (p. 9).

Exploration of the effect of open access on the gender narrative is limited [11]. To understand it further, we discuss a small body of research showing that within some disciplines and some geographic locations, women are making use of open access channels to share their research output with positive effects on the gender narrative. Zaveri notes the social and cultural constructions of gender limit the equal participation of men and women, even within “open processes” [34] (p. 89). However, the analysis below of gender distribution across Gold (publisher-based, DOAJ listed), Hybrid (Gold publisher not included in the DOAJ) and Green open access (repository-based) identifies the potential for change for women who embrace open access research in terms of output, citation advantage and leading authors. From the literature search profile outlined in Section 2.1, we identified a subset of items that explore open research output performance and productivity by gender.

Amy Atchison set out to understand if “Green OA is a gender-egalitarian publishing model” and if women self-archive at the same rate as men within the discipline of political science [35] (p. 450, italics in original). She identified 704 peer-reviewed articles published in non-Gold OA journals in the political science discipline (primarily North American) between 2007 and 2008 using Google Scholar. Atchison selected the articles with green OA versions (archived in a repository or website) and manually determined gender from author information on institutional websites and in author biographies. She used the “Wilcoxon–Mann–Whitney (WMW) test” to determine the position of each gender’s outputs from the mean (p. 451). The result showed a neutral gender citation advantage. Although women political scientists did not self-archive as much as men, they received the same rate of
citations as men for the Green OA (self-archived) articles. The effect was to negate the citation advantage men receive from publishing in closed-access political science journals. This is a positive outcome of open access for gender equality in academic environments where measuring citations features in promotion and tenure evaluations.

Analysing social science and humanities publications from 2008 to 2019, Vuong, Nguyen, Ho and Nguyen found some growth in OA publishing among women authors from Vietnam [36]. The authors used Bayesian analysis of 1201 open publications from the Vietnamese Social Science and Humanities database (SSHPA) from 2008 to 2019, determining gender from public profiles and emails, and used Unpaywall to establish OA status. They excluded Green repository OA from the analysis, focusing only on published articles (Gold, Hybrid or Bronze). Although men still dominated the first author position in OA articles, publications in Hybrid journals show the ratio of women and men as equal, providing some advantage to women researchers. The article also notes the high APCs of publishing can reduce Gold OA as a venue for women who lack resources and encouragement in this endeavour. In related research, Ngyuen et al. [11] analysed 3122 social science and humanities publications from the SSHPA, finding more articles with mixed authors (men and women) published as Gold OA than solo men or women authors.

Ruggieri, Pecoraro and Luzi [37] analysed 22,428 multi-disciplinary articles authored by at least one researcher from Consiglio Nazionale delle Ricerche (CNR), the Italian National Research Council, published from 2016 to 2018 and indexed in the Web of Science (noting its selection and language limitations). They found women chose to publish more OA compared to men (39% vs. 35.8% in the first by-line position and 36.3% vs. 34.0% as last-named authors), particularly in the natural and medical sciences. The number of women-authored publications was slightly higher than those authored by men in Gold OA (21.1% vs. 19.6%) and the same for Hybrid OA (non-DOAJ Gold and Bronze OA, with no reuse licence, 7.2% vs. 6.8%). CNR women researchers have selected open publishing more than their male colleagues, providing a positive increase in women’s research output and visibility and potential for greater impact.

Among the top 1205 elite researchers in Brazil who received highly prestigious Research Productivity Scholarships (PQ)1A from the National Council for Scientific and Technological Development (CNPq) in 2016, women made up only 25% of the researchers. However, dos Santos Costa, Weitzel, and Leta [38] found more women published in Gold OA journals listed in the DOAJ than men (14% vs. 6%). The publications were dated from 2000 to 2015 and included some journals with APCs. The authors note that women performed above men in this influential group with an overall low uptake of OA, suggesting they were more sensitised and motivated (“mais sensibilizadas e motivadas” [38] (p. 36)) to publish in OA journals. By contrast, Olejniczak and Wilson found that in the United States, publishing via Gold and Hybrid open access requiring APC payments in 11 disciplines is “skewed toward scholars with greater access to resources and job security” [39] (p. 1429). That is, high-ranked men who are employed in STEM disciplines in elite universities with substantial federal funding are more likely to choose Gold or Hybrid publishing options. It would be interesting to know whether individual researchers or their organisations paid the APCs. The charging of such fees can be a disincentive to publish Gold OA and a disadvantage for women who may receive less funding than men. However, the Green repository or self-archiving path involves no fees or charges.

In the literature examples discussed above, showing women academics’ use of open access to disseminate their research (Table 1), women feature as first authors in five articles [11, 24, 35, 37, 38] and as the corresponding author in one article [36]. This suggests that women are interested in and driven to undertake research about open access performance with a focus on gender. While not surprising, this demonstrates the roles and opportunities for women: to continue disseminating research through open access methods, to lead research to build further understanding, to transform their scholarly landscape, and to inspire others.
Table 1. Examples in the literature of open access publishing by women academic researchers counteracting the predominant deficit narrative.

<table>
<thead>
<tr>
<th>Authors</th>
<th>Geographic Location</th>
<th>Disciplines</th>
<th>Type of OA</th>
<th>Findings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nguyen, Nguyen, Le, Ho and Vuong (2021) [11]</td>
<td>Vietnam</td>
<td>Social Sciences and Humanities</td>
<td>Gold (fully OA, some APCs)</td>
<td>Mixed-gender (men/women) authors publish more in Gold</td>
</tr>
<tr>
<td>Murphy et al., (2020) [24]</td>
<td>International</td>
<td>Analytical Chemistry, Bioscience, Computer Science, Engineering, Management, Medicine, Psychology, Statistics</td>
<td>Open science literature</td>
<td>More women in high-status authorship statements</td>
</tr>
<tr>
<td>Atchison (2017) [35]</td>
<td>North America</td>
<td>Political Science</td>
<td>Green</td>
<td>Citations equivalent to men’s, neutralising existing gap</td>
</tr>
<tr>
<td>Vuong, Nguyen, Ho and Nguyen (2021) [36]</td>
<td>Vietnam</td>
<td>Social Sciences and Humanities</td>
<td>Hybrid</td>
<td>Equal ratio of women and men authors</td>
</tr>
<tr>
<td>Ruggieri, Pecoraro and Luzi (2021) [37]</td>
<td>Italy</td>
<td>Natural Sciences</td>
<td>Gold (DOAJ), Hybrid (non-DOAJ gold, bronze)</td>
<td>More women publishing OA than men</td>
</tr>
<tr>
<td>dos Santos Costa, Weitzel, and Leta (2020) [38]</td>
<td>Brazil</td>
<td>Science and Technology</td>
<td>Gold (DOAJ)</td>
<td>Women 14% vs. men 6% publishing OA</td>
</tr>
</tbody>
</table>

3. Methods and Data

In this article, we refer to those who self-identify as a woman or are of female gender and who may be cisgender, transgender or of other gender [40]. Globally, most higher education demographic workforce statistics collected and published are binary (women/men, female/male). The few country collections whose collections include non-binary data in categories such as ‘Unknown’, ‘Unclassified’, ‘Unspecified’, ‘No information’ and ‘Gender diverse’ often exclude such data from analysis for privacy reasons or because the numbers are too small [41]. We acknowledge the limitations of such binary analysis and the irony of using such binarised data when our argument is based on the breaking down of binaries.

We analyse workforce and revenue data for publicly funded universities sourced from the Australian Department of Education, Skills and Employment and the United Kingdom Higher Education Statistics Agency (HESA). Universities in these countries are required to report such data each year. To count people, we use staff numbers (headcount statistics). Australian statistics do not include staff on casual work contracts for whom reporting is in a different time frame and format [42]. The United Kingdom data do not include staff on “atypical” contracts, also counted in a different time frame [43]. We focus here on academics as producers of research. However, we recognise the role of professional or non-academic staff, such as library and research office staff, in facilitating and supporting academic research. We use demographic data analysis instead of the name-to-gender disambiguation approach employed often in bibliometric analyses of scholarly outputs.

OA research output analysis used Microsoft Academic bibliographic publication data obtained for each university through APIs (application programming interfaces, internet-based protocols for accessing information from a source). Every author has the full count of all their outputs, and for every institution, each distinct affiliated output is counted once. In our dataset, the number of OA items (with DOIs) published in 2020 was approximately 60,000 for Australia and 160,000 for the United Kingdom [44]. The OA status of each output was obtained from the Unpaywall database, and the number of OA research outputs for each university was counted and divided by the total output for that university. We analyse Crossref Events Data for each university, referring to the number of social media and online references to published research. Events data were normalised by the university’s total number of outputs [45].

4. Name-to-Gender Disambiguation

Analysis of gender productivity and research output to explore biases and disparities relies on the disambiguation of author names to determine their gender. Several methods exist to achieve this, including the census and national research databases such as the US Social Security Administration [46], the US baby names website [40], Wikipedia name lists, scraping or manual retrieval from websites, and applying face recognition software to web images. Software packages to undertake gender name disambiguation, including
ropensci/gender [47] and the Wiki-Gendersort algorithm [48], do not analyse successfully unisex or non-binary gender names and non-Western names, thus often excluding research by authors with these names, leading to incomplete analysis. For example, longitudinal research by Huang, Gates, Sinatra and Barabási [49] uses a commercial software package Genderize.io but states it excludes researchers from China, Japan, Korea, Brazil, Malaysia and Singapore. Olejniczak and Wilson used Genderize.io to “infer” the gender of faculty members listed in the US Academic Analytics database [39] (p. 1431). Transliteration of non-Anglo names in databases can present incorrect gender assignations, and lack of transparency and inherent algorithmic biases leading to the reinforcing of gender stereotypes create problems with such methods [50].

Analyses that utilise country, organisational and discipline-based databases as sources for name gender disambiguation may be more reliable. For example, Vuong et al. and Nguyen et al. used SSHPA, the Vietnamese social sciences and humanities database [11,36]. Ruggieri, Pecoraro and Luzi analysed the output from the CNR, where they are employed, and whose author names and gender are familiar [37]. Kwiek and Roszka determined gender from the Polish Science Observatory dataset they maintain in order to analyse collaboration patterns of Polish scientists, finding disparities by gender, age and position in terms of international and national collaboration [51].

We are conscious that at a global level, gender disambiguation analysis and conclusions about gendered research, assessing productivity and performance, are problematic and best viewed as indicative. Gender behaviour differences are not fixed rigidly by category and vary contextually and situationally: “organisational culture, including formal and informal norms and incentives, clearly influences behavior” [52] (p. 40). More than this, gender is a question of diversity in self-identity, and algorithmically assigning gender in the service of addressing inequity appears to us a contradiction in terms. In academia, research publishing choices are often determined by the institutional infrastructures where decisions are made about best practice requirements for the institution (such as world university rankings and reputation) rather than in relation to sharing and disseminating research widely and within communities.

As our Open Knowledge Initiative project has a global focus, we choose not to apply name gender disambiguation methods because of the incompleteness discussed above. Rather, we examine gender as it relates to research production using public higher education workforce demographic statistics, which were available to us from specific countries and regions [41].

5. Correlation and Statistics Analysis

We collected and analysed data from a range of sources to develop a set of open knowledge institution (OKI) indicators through which to understand the performance of 43 Australian universities and 155 UK universities and their progress towards openness [45]. The indicators include types of OA research output (Gold—publisher, Green—repository); Crossref online and social media events. (https://www.crossref.org/services/event-data/, accessed on 2 April 2022); percentages of women by academic levels; and total revenue by the university. Data are primarily from the year 2020.

Pairing these indicators shows a weak positive correlation (Spearman rank of 0.27) between the Australian universities with higher percentages of women in academic roles (women employed as academics %) and Gold OA (Gold OA publications %) research output (Figure 1a). This aligns, for example, with research showing more academic women publishing Gold OA in Italy and Brazil [37,38].

However, in the United Kingdom (Figure 1b), there is a mild negative correlation between universities with a slightly higher percentage of women in academic roles and Gold OA output (Spearman rank correlation −0.12). We also note that the level of correlation between Gold OA and revenue/income is much higher for the United Kingdom universities (Spearman rank correlation 0.48) than for Australian universities (Spearman rank correlation 0.12); see Appendix A.
We collected and analysed data from a range of sources to develop a set of open knowledge institution (OKI) indicators through which to understand the performance of universities and regions [41].

Figure 1. Scatterplot shows correlations between ranks in percentages of women in academic roles and percentages of Gold (publisher) OA output across (a) 41 Australian and (b) 155 United Kingdom universities, 2020. The data are presented in terms of ranks. Data sources: Australia Department of Education, Skills and Employment, UK Higher Education Statistics Agency (HESA); Microsoft Academic, Unpaywall, Crossref. Image: Curtin Open Knowledge Initiative.

The analysis of Green OA, repository-based output (Green OA publications %) shows a slightly negative correlation with percentages of academic women in Australian universities (Spearman rank correlation $-0.05$) (Figure 2a) and UK universities (Spearman rank correlation $-0.10$) (Figure 2b). Levels of correlation between Green OA and income are also comparable between the two countries (Spearman rank correlation 0.40 for Australia and 0.39 for the United Kingdom); see Appendix A.

Figure 2. Scatterplot shows correlations between ranks in percentages of women in academic roles and percentages of Green (repository) OA output across (a) Australian and (b) United Kingdom universities, 2020. The data are presented in terms of ranks. Data sources: Australia Department of Education, Skills and Employment, UK Higher Education Statistics Agency (HESA); Microsoft Academic, Unpaywall, Crossref. Image: Curtin Open Knowledge Initiative.

We analyse events and mentions tracked by Crossref, which include the citation of a dataset or patent, discussion of scholarly content in a news article, a Wikipedia page, a blog, or social media [53]. For the 41 Australian universities, a positive, weak correlation (Spearman’s rho $= 0.13$) was found between the percentage of women employed...
as academics and the percentage of research outputs with at least one event or mention as recorded by Crossref (Figure 3a).

In the UK universities (Figure 3b), the Crossref events correlation is mildly negative (Spearman rank $-0.13$), similar to the correlation for Gold OA output. There is also a much higher correlation between event total and income in the United Kingdom (Spearman rank correlation $0.46$) than in Australia (Spearman rank correlation $0.08$); see Appendix A.

We analyse events and mentions tracked by Crossref, which include the citation of a dataset or patent, discussion of scholarly content in a news article, a Wikipedia page, a blog, or social media [53]. For the 41 Australian universities, a positive, weak correlation (Spearman’s rho = 0.13) was found between the percentage of women employed as academics and the percentage of research outputs with at least one event or mention as recorded by Crossref (Figure 3a).

In the UK universities (Figure 3b), the Crossref events correlation is mildly negative (Spearman rank $-0.13$), similar to the correlation for Gold OA output. There is also a much higher correlation between event total and income in the United Kingdom (Spearman rank correlation $0.46$) than in Australia (Spearman rank correlation $0.08$); see Appendix A.

**Figure 3.** Scatterplot shows correlations between ranks in percentages of women in academic roles and percentages of Green (repository) OA output across (a) Australian and (b) United Kingdom universities, 2020. The data are presented in terms of ranks. Data sources: Australia Department of Education, Skills and Employment, Higher Education Statistics Agency; Microsoft Academic, Unpaywall, Crossref. Image: Curtin Open Knowledge Initiative.

In Figure 4, we analyse relationships between academic workforce gender and the nature of higher education institutions in terms of size, length of establishment, location, tradition and research focus. In Australia (Figure 4a) and the United Kingdom (Figure 4b), universities that may be smaller, younger, or located in regional areas have a lower total revenue or income but higher proportions of academic women. The older, more prestigious institutions, located in the bottom right quadrants of both graphs, have lower percentages of women academics but higher incomes. Australian universities (women academic %, total revenue) show a Spearman rank correlation of $-0.64$ (Figure 4a), and for UK universities (women academic %, total income), the rank correlation is $-0.34$ (Figure 4b). A note on terminology: ‘revenue’ is used in the Australian data source and ‘income’ in the United Kingdom data source.

We expand on institutional gender differences in relation to research profiles and reputations (Figure 5). In Australia, percentages of women academics compared to the total academic staff at universities in the Group of 8 (Go8) and the ATN (Australian Technology Network, which includes previous institutes of technology) are below gender parity. In contrast, most universities with academic women above 50 per cent are newer, regional and smaller from the Unaligned, IRU (Innovative Research Universities), IHEA (Independent Higher Education Australia) and RUN (Regional Universities Network) groupings.

A similar pattern emerges in the United Kingdom, where the majority of universities in the established Russell grouping with more research-intensive disciplines have less than 50% women academics compared to institutions in the newer, sometimes smaller Alliance, Million+ and GuildHE groupings with disciplines that are more diverse (Figure 6).
Figure 4. Scatterplot between ranks in percentages of women in academic roles and (a) total revenue in AU$ across Australian universities and (b) income in GBP across UK universities by university groupings, 2020. The data are presented in terms of ranks. Data sources: Department of Education, Skills and Employment, Higher Education Statistics Agency; Microsoft Academic, Unpaywall, Crossref. Image: Curtin Open Knowledge Initiative.

Figure 5. Percentages of women academic staff (headcount) compared to the total number of academics in the institution for 43 Australian universities by grouping, 2020. The full file is available at https://doi.org/10.5281/zenodo.6500293 (accessed on 24 June 2022). Data source: Department of Education, Skills and Employment. Image: Curtin Open Knowledge Initiative.
Figure 5. Percentages of women academic staff (headcount) compared to the total number of academics in the institution for 43 Australian universities by grouping, 2020. The full file is available at https://doi.org/10.5281/zenodo.6500293 (accessed on 24 June 2022). Data source: Department of Education, Skills and Employment. Image: Curtin Open Knowledge Initiative.

A similar pattern emerges in the United Kingdom, where the majority of universities in the established Russell grouping with more research-intensive disciplines have less than 50% women academics compared to institutions in the newer, sometimes smaller Alliance, Million+ and GuildHE groupings with disciplines that are more diverse (Figure 6).

Figure 6. Percentages of women academic staff (headcount) compared to the total number of academics in the institution for a subset of 165 United Kingdom higher education institutions by grouping, 2020. The full file is available at https://doi.org/10.5281/zenodo.6500293 (accessed on 24 June 2022). Data source: UK Higher Education Statistics Agency (HESA). Image: Curtin Open Knowledge Initiative.

6. Discussion

6.1. Findings

In relation to the first reference question, we find the small body of research analysing the uptake of open access by women authors and outcomes discussed in Section 2.2 identifies advantages and possibilities for changing the dominant academic gender narrative through open publishing. Across a range of disciplines and geographies, there are examples of women publishing OA achieving more than, or equivalent to, men in Gold, Green and Hybrid OA. Advantages noted include neutralising the gendered citation performance gap and, in open science, more women in higher authorship positions. The analysis also highlights opportunities for further investigation across disciplines, research environments and geographies. Questions may arise about the quality of the OA journals and if they may be considered ‘predatory’. The definition of ‘predatory’, however, is a contested “grey zone” [54] and the suggestion that OA equates with ‘predatory’ is attributed to the now-discredited Beall’s List [55]. All studies in Table 1 list their sources or use established bibliographic databases. While it is possible, some ‘predatory’ titles may be indexed in these sources, and this would apply to all analyses. We found no gender propensity to publish in journals with a questionable status. The small-scale analyses in the articles, focusing on research populations where full author identification is achievable, provide a more complete gender name analysis than many large-scale bibliometric analyses which utilise algorithmic name gender disambiguation.

Second, our analysis of correlations and relationships between percentages of women academics and open access publishing in Australian and United Kingdom universities finds that Australian universities with higher percentages of academic women have slightly higher levels of Gold OA output. Australia does not yet have a national OA policy, and the OA mandates of the two major research funders, the Australian Research Council (ARC) and the National Health and Medical Research Council (NHMRC), allow OA either via a
repository (Green) or via a publisher (Gold). This suggests the researchers’ choice of OA mode. In UK universities, the same correlation is mildly negative. This may be a reflection of a stronger OA policy in the United Kingdom, with funding for Gold OA publishing fees (APCs) provided through transformative agreements with publishers [56], encouraging more use of Gold OA across both genders and reducing the impact in universities with more academic women. UK institutions have taken different OA paths depending on the amount of OA funding they received from UK research councils [57]. The difference may also be due to a factor of institution size, with a greater number of universities in the UK scatterplot (155) and research outputs compared to Australia (41 universities). In both countries, we find a slight negative correlation between the percentages of academic women and Green OA. Although surprising, because self-archiving research output into repositories (Green OA) involves no fees, authors are not always aware of the option or may find it complex to navigate the process of repository deposit, which includes ascertaining journal policies [58].

Factors affecting the motivation for self-archiving include researchers’ disciplinary context, peer pressure, technical skills and age, as well as copyright concerns [59]. Women may appear to self-archive less than men, but the reasons are multifactorial. Gender disparities contribute to this imbalance. For example, fewer women in most science disciplines [60], more women hold lower or junior positions in some disciplines [35]; women publish less or may have greater risk aversion to using a perceived less acceptable publishing venue [61]. However, as we discussed above, self-archiving or Green OA neutralised the gender citation advantage in Political Science [35]. This positive message suggests publishing research via OA can change the predominant gender narrative of women’s deficit research performance.

Crossref events refer to mentions of scholarly output in social media, blogs and news outlets and provide opportunities for online engagement and dissemination of research to a broader audience [62]. Analysis by Fortin et al. [63] demonstrated a gender-neutral effect from the Altmetrics attention score (measuring similar events to Crossref) for research output in Nature, PNAS, PLOS One, New England Journal of Medicine, Cell, and bioRxiv from 2011 to 2018. The slightly positive correlation we find between Crossref events and percentages of women in academic roles in Australian universities suggests efforts by them towards making their research visible and accessible beyond traditional scholarly sources.

We also suggest correlations exist between percentages of academic women and income or revenue in universities with different locations and traditions. In Australia, the younger, regional, less prestigious universities have higher proportions of women academics. These institutions feature more disciplines such as nursing, education, social sciences and humanities, where community engagement and outreach focus is strong and in which women can reach higher academic ranks. Older, wealthier, metropolitan, and more prestigious institutions have strengths in the research-intensive disciplines of science, technology, engineering and medicine, where men tend to dominate. This gender disparity aligns with the historical profile of research output in science disciplines globally, except for Argentina and Russia [49]. The regional Australian universities may attract lower profile researchers because of their geographic location and reputation, although they receive less research funding [64]. However, these institutions offer opportunities for women who may have different career patterns and flexibility, levels of research output and grant funding, and who do not necessarily fit the institutional promoted norm of research ‘excellence’ [65]. Similarly, our analysis of UK universities indicates members of the prestigious Russell Group have lower percentages of women academics than some newer and more diverse discipline-based institutions (Alliance, Cathedrals, GuildHE, Million+).

6.2. Limitations

We acknowledge the limitations of our focus on binary gender and the need to explore intersectionality and inequalities [66] within the construct of research ‘excellence’ in academia to understand the strategies underrepresented groups are using to create positive change in academic institutions. Our correlational analysis is limited to output and performance from two countries for a specific year. We plan to extend the analysis to other
countries where workforce demographics are available publicly and to analyse longitudinal data in order to observe trends in gender balances within open research output. Further investigation by discipline will provide a deeper understanding of situations in which open access advantages for women may appear.

7. Conclusions

In this paper, we have investigated the relationship between gender and open access publishing through a review of existing literature and our own correlational analysis examining links between academic women and ways of sharing research openly. This takes a different approach to bibliometric analyses such as those discussed in Section 2.1, in which women are almost always depicted as deficit research performers compared to men. The use of correlational analysis presents a macro-level perspective on women in academia and OA practices and a means of exploring the narrative around gender balance to suggest, although not necessarily to prove, a causal relationship. It signals possibilities for further in-depth analysis and indicators to map and understand gender-based research performance: exploring the effects of institutional contexts, notions of research excellence, and thinking about how to effect change in the scholarly space [45]. In this way, we open the discussion and draw attention to the possibilities of changing the academic gender narrative.

To return to our first question of OA publishing countering the existing deficit narrative, two strands emerge. First, we suggest women can achieve wider dissemination of research by moving beyond traditional, existing scholarly communication methods and challenging the publication, promotional and funding practices that continue to impose gender-blind assumptions [67]. Second, research such as our correlational analysis examines the context and situatedness of knowledge production [68] and draws attention to the importance of institutional and cultural change; to extend and apply fully intersectional equality and diversity policies and for leadership to recognise, acknowledge and welcome the different perspectives and experiences that gender diversity contributes to knowledge production [9]. The barriers and factors affecting women’s research productivity need to be acknowledged, highlighted and counteracted in output analysis and institutional assessment and evaluation practices.

Kiesewetter proposes an intersectional “genealogy for critical OA publishing” that incorporates the feminist and decolonising methods to challenge the hierarchical patriarchies within academia [69] (p. 62). She highlights successful collective, non-commercial and open repositories and invokes the methodologies of publishing projects that challenge and cross geographical, structural borders and academic hierarchies. This critical approach aligns with the aims of our Open Knowledge Coalition: we believe information is a community asset and an opportunity to move beyond commercial systems toward the construction of resources that are governed by the higher education and research community. The infrastructure supporting research and scholarship is an activity in which we can all participate and be involved in to bring about change and expand the diversity of knowledge production [70].


Funding: The project is funded by the Research Office of Curtin University through a strategic grant, the Curtin University Faculty of Humanities and the School of Media, Creative Arts and Social Inquiry.

Data Availability Statement: The data and codes used to generate the figures in this article are available at Zenodo: Katie Wilson, Chun-Kai Huang, Lucy Montgomery, Cameron Neylon, Rebecca N. Handcock, Alkim Ozaygen and Aniek Roelofs (2022). Data and codes: Changing the Academic Gender Narrative through Open Access. https://doi.org/10.5281/zenodo.6672931 (accessed on
The levels of correlation between Gold OA (Gold OA publications %) and revenue or income (Total revenue, Total income) are more comparable between the two countries (Spearman rank correlation 0.12 for Australia, Figure A1a) and income (Total income) for United Kingdom universities (Spearman rank correlation 0.48, Figure A1b). This suggests that universities with greater income provide support for the payment of Gold OA fees or APCs, and such support and output is higher in the United Kingdom.

Additional analysis explores the relationship between open access (OA) research output in 2020 and the revenue/income of universities in Australia and the United Kingdom. We include these figures in the Appendix A because the correlations do not present results relating to gender but link indirectly to the main text. Figure A1 shows the correlation between Gold OA (gold OA Publications%) research output and revenue (Total revenue) for Australian universities (Spearman rank correlation 0.12, Figure A1a) and income (Total income) for United Kingdom universities (Spearman rank correlation 0.48, Figure A1b). This suggests that universities with greater income provide support for the payment of Gold OA fees or APCs, and such support and output is higher in the United Kingdom.

### Figure A1
Scatterplot shows correlations between percentages of Gold OA publications and revenue or income across (a) 41 Australian and (b) 155 United Kingdom universities, 2020. The data are presented in terms of ranks. Data sources: Australia Department of Education, Skills and Employment-U Cube Higher Education Data cube; UK Higher Education Statistics Agency–Higher Education Staff Data.

**Acknowledgments:** The authors thank the anonymous reviewers and editors for their generous and constructive comments that have improved this paper.

**Conflicts of Interest:** The authors declare no conflict of interest.

### Appendix A

Additional analysis explores the relationship between open access (OA) research output in 2020 and the revenue/income of universities in Australia and the United Kingdom. We include these figures in the Appendix A because the correlations do not present results relating to gender but link indirectly to the main text. Figure A1 shows the correlation between Gold OA (gold OA Publications%) research output and revenue (Total revenue) for Australian universities (Spearman rank correlation 0.12, Figure A1a) and income (Total income) for United Kingdom universities (Spearman rank correlation 0.48, Figure A1b). This suggests that universities with greater income provide support for the payment of Gold OA fees or APCs, and such support and output is higher in the United Kingdom.

The levels of correlation between Green OA (Green OA publications %) and revenue or income (Total revenue, Total income) are more comparable between the two countries (Spearman rank correlation 0.40 for Australia, Figure A2a, and 0.39 for United Kingdom, Figure A2b). Australian universities produced a higher percentage of Green repository-based research output (green OA %) than Gold (publisher) OA in 2020, while the percentages of Green and Gold output from UK universities are more similar.

In Australian universities, this analysis indicates there seems to be limited connection between events or mentions of research recorded by Crossref (Publications with Crossref events %) and revenue (Total revenue) (Spearman rank correlation 0.08), see Figure A3a. This is very different for the UK where the Spearman rank correlation between events (Publications with Crossref events %) and income (Total income) is higher at 0.46 (Figure A3b).
Figure A1. Scatterplot shows correlations between percentages of Green (repository) OA publications and total revenue/income across (a) Australian and (b) United Kingdom universities, 2020. The data are presented in terms of ranks. Data sources: Australia Department of Education, Skills and Employment, UK Higher Education Statistics Agency (HESA); Microsoft Academic, Unpaywall, Crossref. Image: Curtin Open Knowledge Initiative.

Figure A2. Scatterplot shows correlations between percentages of Gold OA publications and revenue or income across (a) Australian universities (Total revenue) and (b) United Kingdom universities (Total income), 2020. The data are presented in terms of ranks. Data sources: Australia Department of Education, Skills and Employment, UK Higher Education Statistics Agency (HESA); Microsoft Academic, Unpaywall, Crossref. Image: Curtin Open Knowledge Initiative.

Figure A3. Scatterplot shows correlations between percentages of publications with Crossref events (%) and revenue or income across (a) Australian universities (Total revenue) and (b) United Kingdom universities (Total income), 2020. The data are presented in terms of ranks. Data sources: Australia Department of Education, Skills and Employment, UK Higher Education Statistics Agency (HESA); Microsoft Academic, Unpaywall, Crossref. Image: Curtin Open Knowledge Initiative.

Note
The data were transformed to their ranks by orders of magnitude, e.g., the university with the highest percentage in “Women employed as academics” in the cohort is given the rank of one. This allows us to use the Spearman rank correlation coefficient to measure the strength and direction of the monotonic relationship between the two variables while catering for non-normality and outliers.

References


10. van den Besselaar, P.; Sandström, U. Vicious Circles of Gender Bias, Lower Positions, and Lower Performance: Gender Differences in Scholarly Productivity and Impact. *PLoS ONE* 2017, 12, e0183301. [CrossRef]


38. dos Santos Costa, E.H.; Weitzel, S.d.R.; Leta, J. Adherence of the Brazilian elite of researchers to open access journals and its relation to gender, region and area of knowledge. *Em Questão* 2020, 26, 15–42. [CrossRef]


51. Kwiek, M.; Roszka, W. Gender-Based Homophily in Research: A Large-Scale Study of Man-Woman Collaboration. *J. Inf.* 2021, 15, 101171. [CrossRef]


70. Cressman, C. Trust in Infrastructure. *Commonplace* 2021. [CrossRef]