Influence of different journal publishing models in the presence and detection of scientific errors and misconduct

La influencia de diferentes modelos de publicación en la presencia y detección de errores y fraude científico

ABSTRACT

This study attempts to test how different journal publishing models can favor or reduce the presence of errors and misconduct articles, as well as to measure the response of journals to problematic articles according to these publishing models. For this, a new approach for the study of scientific misconduct in publications is proposed. Comments expressed in PubPeer about 17,244 troublesome articles were compared with the editorial response of journals (i.e. editorial notices). Journals of these publications were classified according to several publishing criteria: publisher type, access type, publication fee model and peer review type. The results show that in spite of scholar-published journals suffer more from problematic papers, they release the same editorial notices than commercial journals; open access journals react better to problematic articles than paywall journals; open access journals without APC has a special presence of Publishing fraud; and journals that use open review suffer less from misconduct, slightly releasing more editorial notices.

Keywords: PubPeer; scientific misconduct; scholarly publishers; open access; open peer review; editorial notices.

RESUMEN

Este estudio pretende comprobar cómo diferentes modelos de publicación de revistas científica pueden favorecer o reducir la incidencia de artículos erróneos o fraudulentos, a la vez que busca medir la respuesta de revistas a estos problemas en función de estos modelos. Para esto, se propone una nueva forma de estudiar el fraude científico en las publicaciones. Los comentarios expresados en PubPeer sobre 17.244 artículos problemáticos fueron comparados con la respuesta editorial de las revistas (i.e. notas editoriales). Las revistas de estas publicaciones fueron clasificadas en función de diferentes criterios editoriales: tipo de editor, tipo de acceso, modelo de financiación y tipo de revisión por pares. Los resultados muestran que a pesar de que las revistas editadas por la academia sufren más de artículos problemáticos, emiten el mismo número de notas editoriales que las revistas comerciales; las revistas de acceso abierto reaccionan mejor ante artículos problemáticos que revistas de pago; revistas de acceso abierto sin APC tienen una incidencia especial de Fraude en la publicación; y revistas que emplean una revisión en abierto sufren menos de fraude científico y ligeramente emiten más notas editoriales.

Keywords*: PubPeer; fraude científico; editores académicos; acceso abierto; revisión en abierto; notas editoriales.*

1. INTRODUCTION

The detection of errors and misconduct after the publication of research articles is a serious problem that reveals inefficiencies in the editorial control of publications (Marusic et al., 2007). The causes of this lack of control could be due to several factors: a research evaluation system that rewards more the venue than the output itself, provoking that researchers want to publish in a journal, instead to make public a scientific result (American Society for Cell Biology, 2012); opaque peer review systems that protect compromised peer review and make easy to sneak fraudulent manuscripts (Hadi, 2016); and a publishing system addressed more to obtain economic benefits than to ensure research integrity (e.g. recommended reviewers, pay for fast review) (Hawkes, 2015; Gao and Zhou, 2017). All these reasons would explain why even more journals need to correct articles already published.

The number of editorial notices released by a journal has been considered an indicator of incidence of misconduct, pointing out that journals with a high rate of retractions and errata are

journals that seriously suffer from misconduct and erroneous practices. This could be true if we assume that all the unreliable articles are identified and corrected. However, not all the journals correctly react to these cases and many articles with clear cases of manipulation or plagiarism remain published without any advice (William and Wager, 2011), contradicting the Committee on Publication Ethics's (COPE) guidelines (COPE, 2012). The reason could be that many journals lack of the sufficient editorial control to correct a posteriori erroneous publications. Even, this absence of reaction would hide the discredit of recognizing failures in the management of manuscripts.

An important factor that would inform us about the response ability of journals is their publishing model. The fact that publishers have a profit or nonprofit purpose; opt for an open access or paywall model; their publication fees are charged to authors, instead to readers; or they adopt innovations such as open peer review, could be elements that interfere in the ability of journals to react to unreliable articles. These different models could indicate what aspects could be associated to certain types of problems and how they are detected by journals' editorial boards.

Post publication peer review sites (e.g. Publons, PubPeer, F1000) could provide valuable information about unreliable cases, which could be contrasted with the reaction of journals. In this way, the number of editorial notices could be interpreted in a positive way, reflecting now the ability of a journal to detect problematic publications. This study aims to offer a new perspective about the study of scientific misconduct, matching, on one hand, the comments expressed in PubPeer about troublesome articles and, on the other hand, the response of journals to those articles through editorial notices. Considering different publishing models, this study attempts to indicate what type of journals react better to unreliable cases.

2. LITERATURE REVIEW

The study of the scholarly misconduct and its relationship with the publishing system has been dealt with differently. Many studies have addressed the limitations of journals in correcting the literature, concluding that the proportion of misconduct in publications should be much larger than the reported in editorial notices (Cokol et al., 2007; Stricker and Günther, 2019). This estimation has been different according to the referenced sample. Thus, comparing with already investigated cases, the response ratio was high (62%-83%) (Wager, 2007; Neale et al., 2007; Resnik and Dinse, 2013). However, these percentages drop when the editorial notices are compared with reports from web platforms. Brookes (2014), who analyzed anonymous complaints in a specialized blog, found that only 23% of the reported articles were later corrected or retracted. Ortega and Delgado-Quirós (2023) discovered that only 21% of articles reported of misconduct in PubPeer were subject of an editorial notice. Otherwise, a significant proportion of misconduct cases are not corrected with editorial notices, which add a new limitation to the use of editorial notices in scientific misconduct studies (Nath et al., 2006; Lei and Zhang, 2018; Vuong, 2020).

Nevertheless, the incidence of errors and misconduct in different publishing models have been treated in a dissimilar manner. The significant proportion of retractions in open access journals has thus been widely studied. Peterson (2013) compared the proportion of retractions in PubMed (.028%), finding that open access journals released almost the double of retractions (.049%). Using the same procedure, but limited to Chinese authors, Wang et al. (2019) detected that 11.6% of the retractions come from open access journals. Tripathi et al. (2019) explored the Web of Science database in 2008-2017 period, and they found that the proportion of retracted articles in open access journals (.522‰) was significantly higher than in subscription journals (.187‰). More recently, Shah et al. (2021) confirmed this perception when they obtained similar figures between open access (.28‰) and toll-based journals (.17‰).

Literature about the impact of open peer review in the detection of errors and misconduct has been much more scarce, and always from a theoretical perspective. Boldt (2011) suggested that open peer review models could make easy the notification of errors and misconduct. Schmidt et al. (2018) reviewed the literature on open peer review and they concluded that open reviews could improve the transparency of review processes, helping the identification of misconduct cases.
Contrarily, studies about the different roles of the scholarly and commercial publishers is almost inexistent. Resnik et al. (2010) studied the presence of misconduct policies and they did not find differences among publishers. Da Silva and Vuong (2021) warned about the ethical issues derived from the economic benefit of misconduct articles by commercial publishers.

3. OBJECTIVES

The aim of this study is to verify the influence of different journal publishing models in the presence of errors and misconduct cases, and how these journals react to these problematic cases. Scholarly vs. commercial publishers, open access vs. paywall journals, APC vs. non APC fees (diamond open access) and open vs. anonymous peer review are compared to find which model would suffer more from misconduct and could be more successful detecting suspicious publications. Four research questions were addressed in this study:

- Because their interest in profits may outweigh the editorial control on troublesome articles, could commercial publishers be more permissive to misconduct? And inversely, could scholarly publishers be stricter detecting misconduct, considering that they do not have an economic motivation?
- Are open access/paywall journals more prone to suffer from misconduct? Or, perhaps, are they more prepared to detect these practices?
- Do publication fees influence to any extent the appearance of misconduct cases? And could these cases be better detected by journals with/without APC?
- Could open peer review better filter erroneous and misconduct publications? And do journals with this peer review modality release more editorial notices?

4. METHODS

This study aims to present a new method to explore the presence of problem articles and how the publishing system reacts to them. To do this, comments that report errors and misconduct cases from PubPeer were extracted, processed and classified, to later be matched with editorial notices released by journals. In this form, we can observe in which proportion journals react to complaints of scientific errors and frauds.

4.1. Sources

PubPeer is a journal club that discusses scholarly documents after being published or uploaded to the Web. This web forum was created by three neuroscientists in October 2012. The possibility of posting comments anonymously was the cause of the rapid success of this post-publication peer review site. This singular characteristic caused that the site was specialized in reporting misconduct and errors of the scientific literature. This fact is generating considerable controversy because many authors feel defenseless in the face of unknown accusers (Torny, 2018). On the contrary, the reporting of bad practices with no reprisals is benefiting the research integrity, bringing to the forefront a varied range of errors and misconduct.

4.2. Data access and extraction

PubPeer does not provide a public endpoint to extract their data (i.e. API, dump files), which

caused that information about publications and associated comments were directly extracted from the website (pubpeer.com) using web scraping techniques. For this study, two samples were retrieved in different moments. 32,097 threads and 65,179 posts were obtained in March 2019. Next, this sample was enlarged and updated with 7,659 threads and 21,200 posts in January 2020. Several queries using the first letters of the alphabet—a, b, and c—in the standard search box were launched to retrieve comments to publications, including the internal ID of each paper. Then, bibliographic metadata and information about the comments associated to those publications (user, text, date, etc.) were sequentially extracted using an *ad hoc* crawler designed with WebQL Studio (www.ql2.com). 26,133 research documents published after 2000 were selected, after a cleaning process (publications without user comments and comments generated by robots).

In spite that PubPeer alerts when a publication has been corrected, the list of publications was searched in several databases to verify or/and to enlarge those notifications. Retraction Watch database (retractiondatabase.org) and PubMed (pubmed.ncbi.nlm.nih.gov) were used to identify editorial notices.

4.3. Classification and selection criteria

A sub-sample of 17,244 (66%) articles were classified according to the content of the comments received from PubPeer. The rest of the publications (8,889, 44%) were rejected because the comments were not sufficiently explanatory or they do not fit with the classification scheme. The classification process was based on the extraction of keywords that described the content of the comments. Publications were then grouped into seven categories using these keywords (Ortega, 2022):

- Positive review: Comments that praise and highlight publications according to the reach and importance of the results.
- Critical review: Comments that discuss the methods and results and their interpretations. This group includes discussions about theoretical implications and scientific disagreements.
- Lack of information: Inside Critical review, this is a sub-category that addresses the problematic absence of information about how the study was performed, the availability of raw data, and lack of relevant bibliographic references.
- Honest errors (Resnik and Stewart, 2012): They could be rectifiable mistakes (e.g., erratum) due to confusion and oversight in the writing of the paper.
- Methodological flaws: They are motivated by a lack of awareness of statistical or other scientific techniques (e.g., western blots, spectroscopy) that throw up wrong results (e.g., correlation fishing, bar errors, loading controls). This category could be bordering on fraud, because this confusion could be intended to obtain the desired results. However, such intentionality is not always evident, and these issues are given the benefit of doubt.
- Publishing fraud: Interference with the publishing system to increase production and impact. It mainly includes plagiarism, reused text, ghost authorship, and fake peer review.
- Manipulation: Intentional edition and manipulation/fabrication of data and images to obtain better results than those expected and to corroborate the desired hypothesis.

Finally, to validate the accuracy of this classification procedure, a sub-sample of comments (4,000) was manually classified and compared with the original procedure. A confusion matrix showed high overall precision (88.1%), demonstrating that close to nine out of ten posts were correctly assigned (Ortega, 2022).

When a publication has generated several editorial notices, the most serious one was selected.

The importance goes from Erratum, Expression of Concern (EoC) to Retraction. Thus, whether an article has been corrected with an erratum, and later was finally retracted, we have then considered this paper as retracted.

To test how the editorial notices and the PubPeer comments are distributed according to different publisher models, different classification criteria were defined to group the publications:

Publisher type: Publications were classified in Commercial or Scholarly according to the journal's publisher. This distinction is established because we hypothesize that scholarly journals could be stricter reviewing and correcting publications and they could better detect erroneous or fraudulent papers. The reasons for this assumption could be that they are directly managed by scientists and they could then have only academic interests. While, commercial publishers could be more focused on economic benefits and could have a more relaxed attitude towards unreliable articles. Commercial was defined for publishers with a clear for-profit orientation (e.g. *Elsevier*, *Springer*, *Wiley*), while Scholarly category is for journals published by learned societies (e.g. *American Association of Medicine*, *Royal Chemistry Society*) or academic organizations (e.g. Universities presses, research centers). However, there is an important number of journals that are owned by scholarly institutions but they are managed by commercial publishers. For example, the *British Journal of Pharmacology* belongs to the *British Pharmacological Society* but is managed by *Wiley*. In these cases, journals were classified as scholarly because we understand that all the editorial responsibilities fall on the academic institution.

Access type: This typology distinguishes between open access and paywall journals. Open access means journals that do not require any subscription to read their content, whereas paywall refers to venues that require a payment to gain access to their articles. This classification was established according to the Directory of Open Access Journals (DOAJ). The rationale of this classification is to observe if different journal management models, could have some relationship with the presence and detection of problematic articles. Hybrid journals are considered paywall because they retain the same editorial process than a subscription-based journal.

Publication fee model: This is a dichotomous classification that defines if open access journals are supported by Author Processing Charges (APCs) or they opt for a different business model such as grants, sponsoring or advertising (*diamond open access*). This classification is defined because journals with APC could have a different relationship with the authors, who pay for the publication. This could cause low rejection rates and less severe reviews of manuscripts.

Peer review: this last classification distinguishes between traditional anonymous peer review and open peer review. The reason of this criterion is to observe if an open review system with more transparency would reduce compromised peer review and be more effective detecting errors and misconduct. This information was obtained from DOAJ and from the journals' websites.

Publisher type		**Publications**	**Publications %**
Scholarly		6,132	35.6%
Commercial		10,931	63.4%
Not classified		181	1.1%
Access		**Publications**	**Publications %**
Paywall		12,931	75.0%
Open Access	APC	3,981	95.3%
	Non APC (*diamond*)	195	4.7%
	Anonymous review	4,066	97.4%

	Open review	110	2.6%
	Total	4,176	24.2%
Not classified		138	0.8%
TOTAL		17,244	100%

Table I. Distribution of the sample according to different types of publisher, access and review.

Table I summarizes the number of publications analyzed according to different publishing, access and review models. This table shows that the number of cases in each category is very different, which could be misleading when these proportions are compared among them. To reduce this risk, each comparison is checked with the χ^2 test of proportions to confirm that the differences are statistically significant.

The entire dataset of this study is publicly available on: https://osf.io/yf3dp/

5. RESULTS

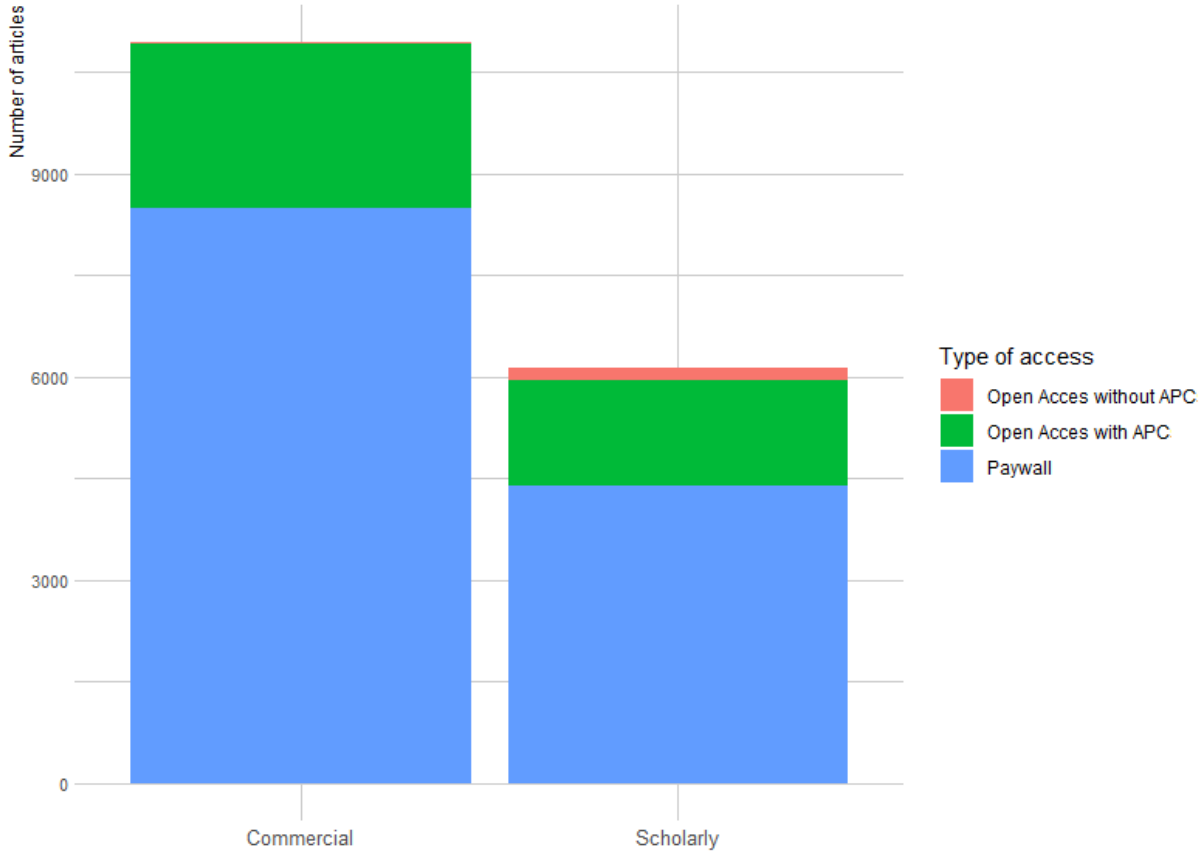

Figure 1. Number of research articles by type of publisher and type of access.

Figure 1 displays the proportion of publications according to the type of publisher and the access way of each journal. The aim is to describe the characteristics of the sample and observe differences between publishers when they establish open access journals. Two thirds of the publications are released in journals owned by commercial publishers (10,931, 64.1%) and a third in scholarly journals (6,132, 35.9%). This proportion evidences the ever more dominant presence of commercial houses in the current publishing system. The proportion of articles in open access journals by type of publisher is rather similar, with 22.3% for commercial publishers (N=2,433) and 28.3% for scholarly journals (N=1,734). However, the chi-square test detects significant differences (χ^2=76.8 p-value<.001), suggesting that scholarly publishers

could be more prone to adopt open access models. Almost the majority of papers in open access journals are published in venues with Author Processing Charges (APC) as publication fee model (3,972, 95.3%). According to the publisher type, there are more articles with APC in commercial (2,416, 99.3%) than in scholarly journals (1556, 89.7%), although these differences are barely significant (χ^2=4.92 p-value=.03).

5.1. Publishers

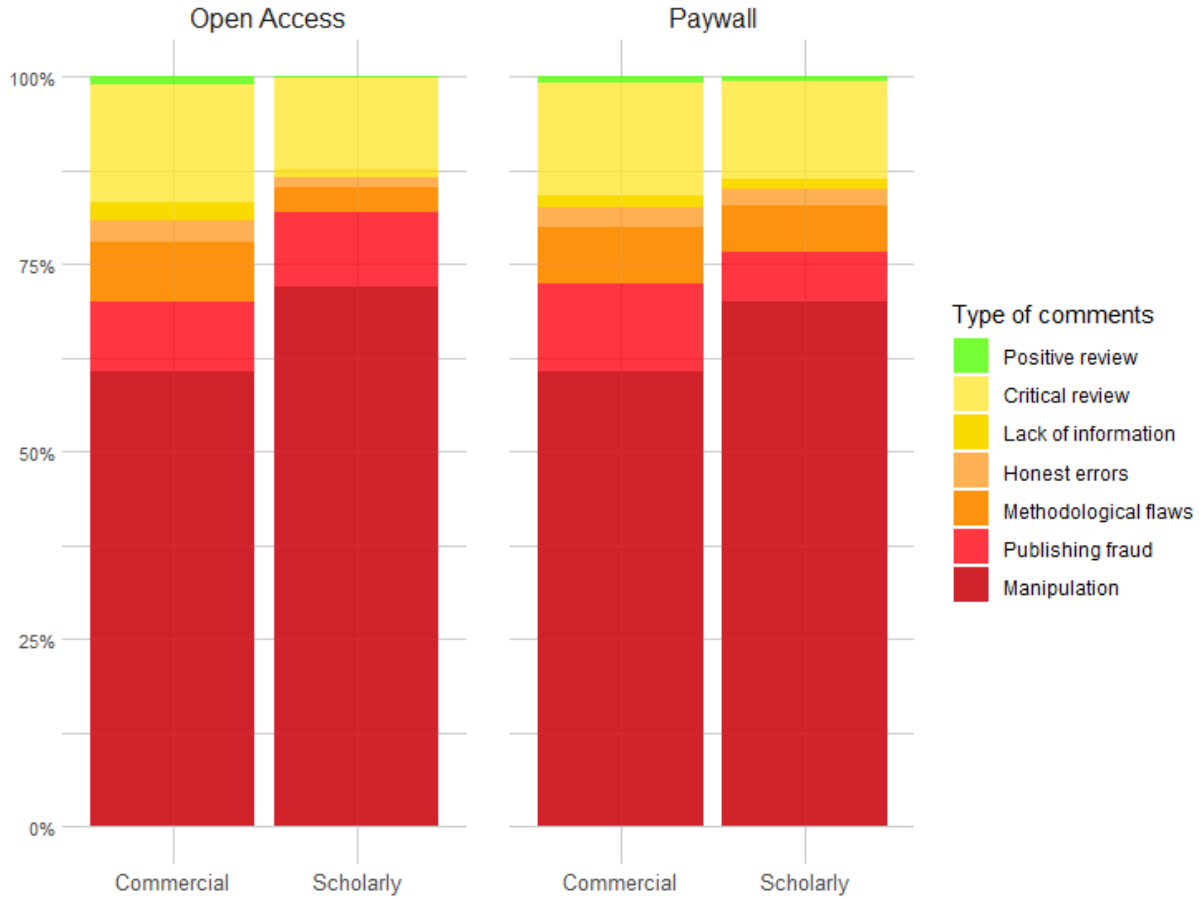

Figure 2. Distribution of articles by type of comment in PubPeer according to type of publisher and faceted by type of access.

Figure 2 depicts the proportion of comments in both type of publishers and according to access type. The purpose is to detect if the distribution of comments is influenced by publisher and access type. The results show an overall high proportion of comments about fraudulent practices (i.e. Publishing fraud and Manipulation) (Ortega, 2022). In comparison between types of publishers, scholarly journals show higher proportion of problematic publications than commercial publishers, both in open access (Publishing fraud=9.9%, Manipulation=72%) and paywall journals (Publishing fraud=6.6%, Manipulation=69.9%). These differences suggest that scholarly publishers suffer more from problematic articles than commercial ones, independently of the type of access (Open Access χ^2=92.43 p-value<.001; Paywall χ^2= 132.85 p-value<.001). With regard to access types, there are no significant differences between commercial open access and paywall journals (χ^2=16.61 p-value=.011), but indeed between scholarly open access and paywall journals (χ^2=47.24 p-value<.001), being scholarly open access journals the typology with the highest number of suspicious articles.

Figure 3. Proportion of publications commented in PubPeer with editorial notices by type of publishers and grouped by type of issue.

Figure 3 describes the percentage of articles that have been subject of an editorial notice according to the type of publisher and faceted by type of error or misconduct. The intent is to observe differences between scholarly and commercial journals when they respond to error or fraudulent papers. In general, the responsiveness of journals to errors and misconduct is rather low, since only 21% of suspicious publications receive an editorial notice (Ortega and Delgado-Quirós, 2023). It is possible to appreciate that scholarly journals release more editorial notices than commercial journals. This is evident in Honest errors (Scholarly=31.9%, Commercial=16.8% χ^2= 10.53 p-value=.0012), Manipulation (Scholarly=22.5%, Commercial=19.3% χ^2= 16.44 p-value<.001) and in less extent in Publishing fraud (Scholarly=19.3%, Commercial=27.7% χ^2= 7.31 p-value=.006). According to the type of editorial notice, Errata are more frequent for Honest errors (20.5%), while retractions are mainly released for Publishing fraud (22.7%) and Methodological flaws (14.2%).

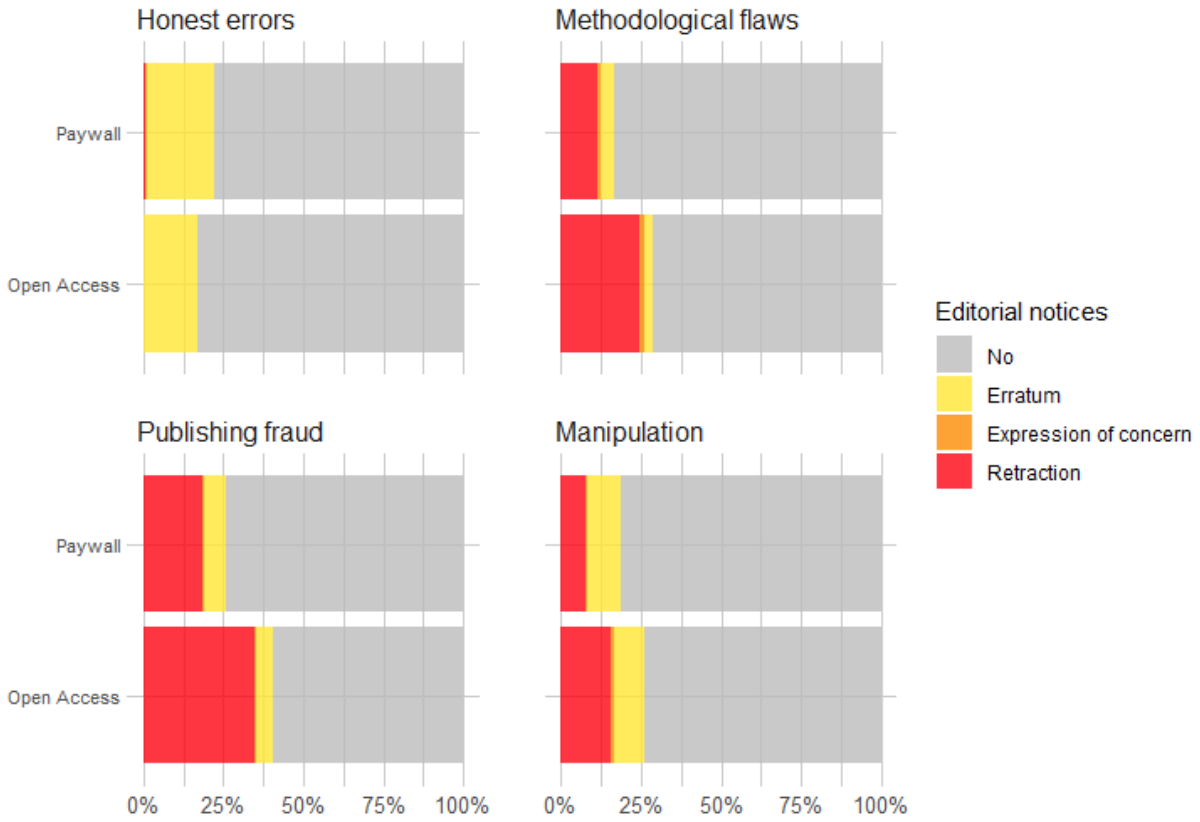

Figure 4. Percentage of publications commented in PubPeer with editorial notices by access type and grouped by type of issue.

Figure 4 presents the proportion of articles with an editorial notice grouped by open access or paywall and faceted by type of problem. The objective is to perceive if open access journals react more or less to problematic articles than paywall journals. The results show that, with exception of Honest errors, open access journals release more editorial notices than paywall journals. This is statistically significant in Methodological flaws (Open Access=28.6%, Paywall=16.7% χ^2= 17.34 p-value<.001), Publishing fraud (Open Access=40.8%, Paywall=25.9% χ^2= 31.84 p-value<.001), and Manipulation (Open Access=25.9%, Paywall=18.8% χ^2= 64.2 p-value<.001). In the event of Honest errors, the differences are not significant (Open Access=17.2%, Paywall=22.3% χ^2= .91 p-value=.34).

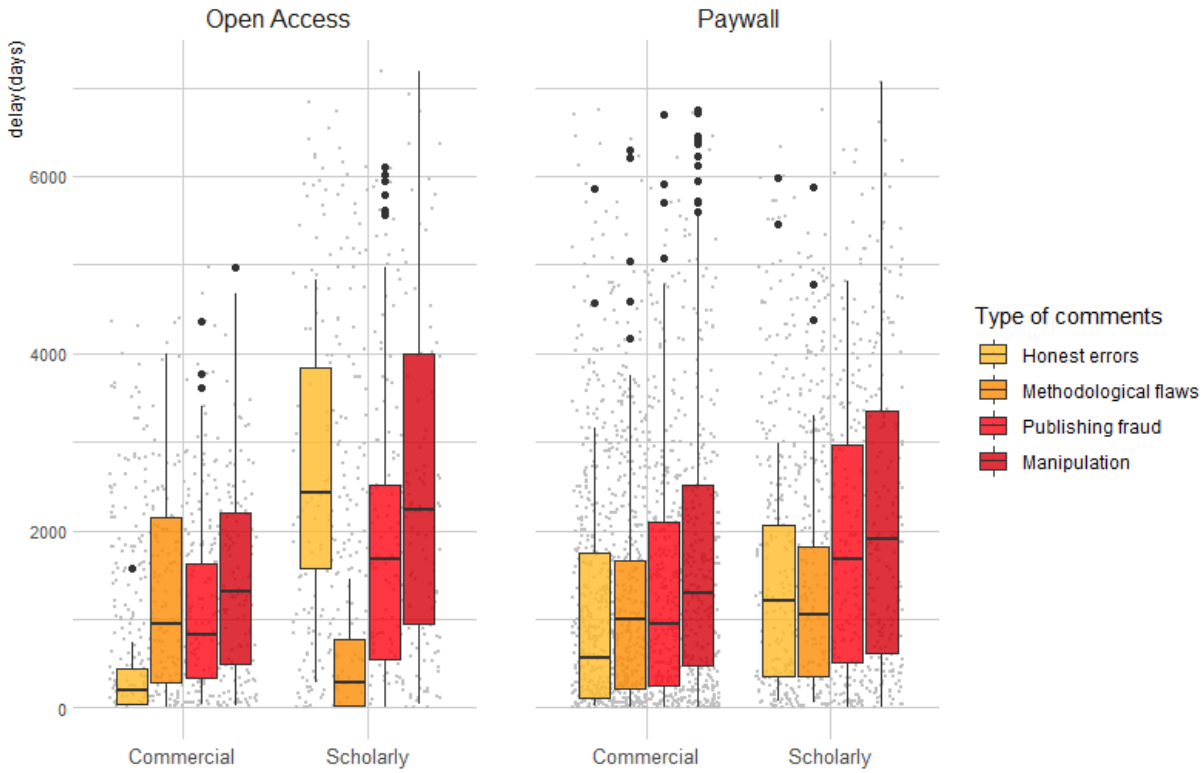

Figure 5. Box-plot of the time delay of editorial notices by type of comments in PubPeer and according to type of publisher and access.

Figure 5 depicts the time delay distribution between publication date and editorial notice release by complaint type in PubPeer. These distributions are faceted by publisher and access type. This measure would illustrate the diligence of journals reacting to problematic articles and inform us if this ability is associated to different types of publishers and accesses. In general, the reaction delay is inversely associated to the degree of seriousness of the comments. Papers reported of Manipulation (*Mdn*=1,552 days) take more time to be notified than Publishing fraud (*Mdn*=1,100 days, p-value<.001), Methodological flaws (*Mdn*=868 days, p-value<.001) and Honest errors (*Mdn*=799 days, p-value<.001). This reaction is also different according to the type of publisher. Scholarly journals (*Mdn*=1,843 days) tend to take more time to react to problem articles than commercial ones (*Mdn*=1,182 days) (p-value<.001). However, the type of access is not a significant factor that influences the delay of the editorial notices. Although Open Access journals take more time to release an editorial notice (*Mdn*=1,515 days) than paywall journals (*Mdn*=1,345 days), these differences are not statistically significant (p-value=.1144). Kruskall-Wallis' test and Dunn's pairwise test with the Bonferroni adjust were used to calculated the statistical differences.

5.2. Publication fee model

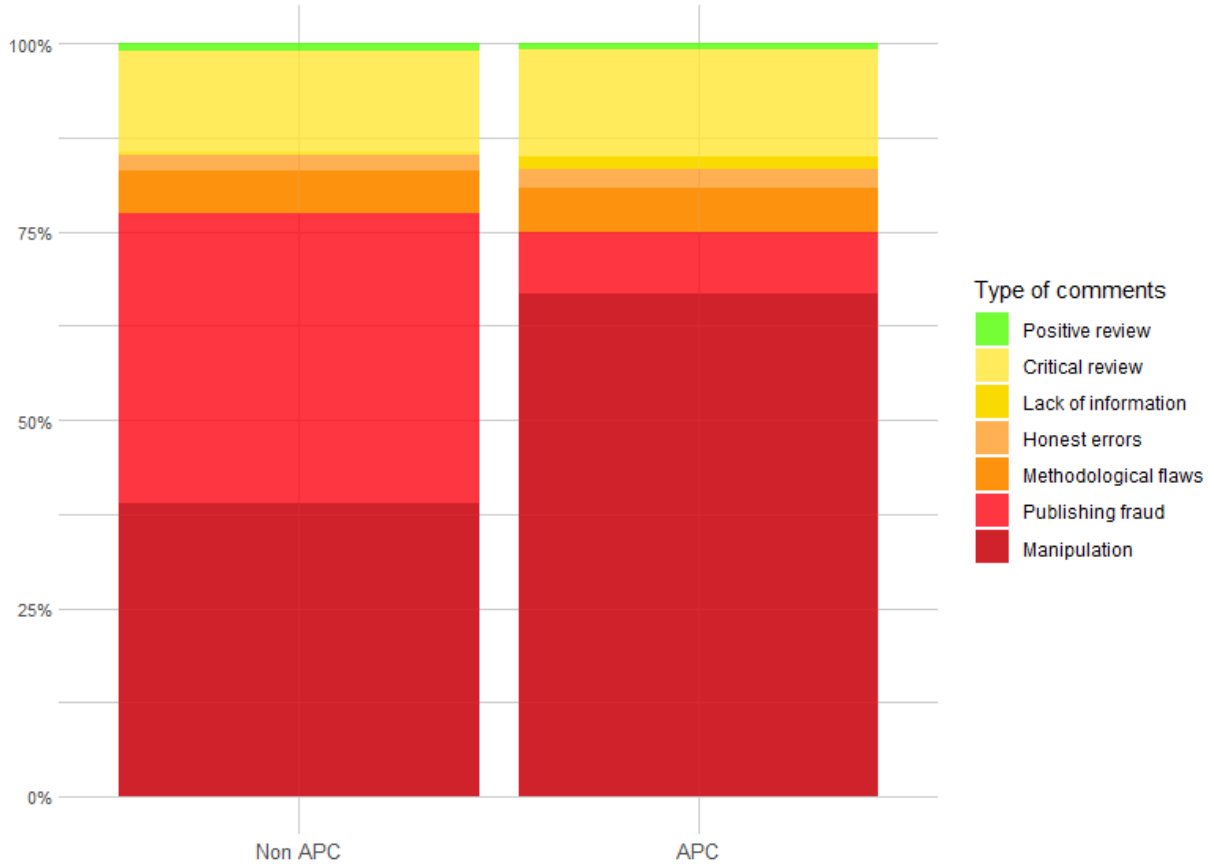

Figure 6. Distribution of articles by type of comment in PubPeer according to journals with or without APC.

Figure 6 displays the proportion of research articles by type of comment in PubPeer according to whether journals have or do not have APC. The aim of this analysis is to observe if this business model would influence the presence of certain types of issues. The results clearly show that there are not differences in the distribution of publications reported of misconduct and error (≈75%), showing similar proportions to Figure 2. This output could suggest that the APC is not a factor that fosters or reduces the research misconduct. However, it is remarkable the high proportion of Publishing fraud (38.5%) in diamond journals (without APC) in comparison with the 8.2% of APC journals (χ^2= 127.82 p-value<.001). Therefore, this result does show that diamond journals specially suffer from Publishing fraud. A possible reason of this fact could be due to these journals have less impact than journals with APC. Considering the distribution of journals by impact quartiles (Scimago Journal Rank, www.scimagojr.com), 88.2% of the journals with APC are located in Q1, while only 28.5% of diamond journals are in Q1. Whereas, 50.6% of diamond journals are in Q3 and .9% of APC journals are in Q3. A possible explanation for this connection would be that these practices, mainly plagiarism and reutilization, do not involve a new knowledge (as it does happen with Manipulation), and therefore they are addressed to increase more the production than the impact. In this sense, low impact journals with less publishing pressure (low rejection rates) could be more appreciated. In addition, this type of misconduct is more evident and easier to demonstrate, accordingly low impact journals with less editorial control could be less prone to detect these practices (Ortega and Delgado-Quirós, 2023).

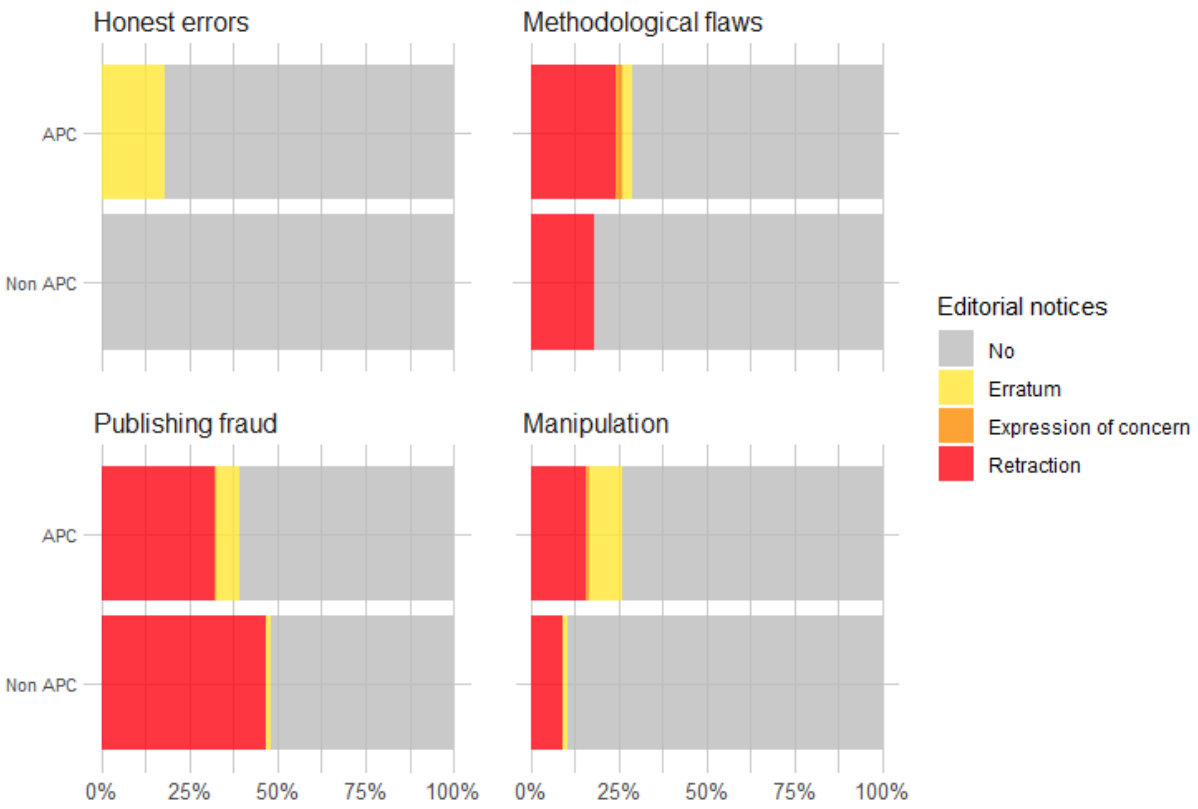

Figure 7. Percentage of publications commented in PubPeer with editorial notices by publication fee model and grouped by type of issue.

Type of comment	Honest errors				Methodological flaws			
Publication fee	Non APC	Non APC %	APC	APC %	No APC	Non APC %	APC	APC %
No	4	100.0%	78	82.1%	9	81.8%	168	70.9%
Editorial notices	0	0.0%	17	17.9%	2	18.2%	69	29.1%
Erratum		0.0%	17	17.9%		0.0%	7	3.0%
EoC		0.0%		0.0%		0.0%	4	1.7%
Retraction		0.0%		0.0%	2	18.2%	58	24.5%
Total	4	100.0%	95	100.0%	11	100.0%	237	100.0%
Type of comment	**Publishing fraud**				**Manipulation**			
Publication fee	**Non APC**	**Non APC %**	**APC**	**APC %**	**No APC**	**Non APC %**	**APC**	**APC %**
No	39	52.0%	198	60.9%	68	89.5%	1956	73.6%
Editorial notices	36	48.0%	127	39.1%	8	10.5%	700	26.4%
Erratum	1	1.3%	21	6.5%	1	1.3%	255	9.6%
EoC		0.0%	2	0.6%		0.0%	22	0.8%
Retraction	35	46.7%	104	32.0%	7	9.2%	423	15.9%
Total	75	100.0%	325	100.0%	76	100.0%	2656	100.0%

Table II. Distribution of editorial notices by publication fee and type of comment.

Figure 7 and Table II present the proportion of publications with editorial notices by journal with or without APC and according to type of error or misconduct. The objective is to test if journals with different publication fee models react differently to errors and misconduct. Firstly, the number of articles from APC (3,981) and Non APC (195) journals is rather unbalanced.

This difference evidences how the APC model is prevailing over the diamond model, supported by commercial publishers (Crawford, 2021). The results show that, in general, APC journals react better to problematic publications, with the exception of Publishing fraud. Thus, the proportion of editorial notices in APC journals is higher in Honest errors (APC=17.9%, Non APC=0% χ^2=.06 p-value=.8), Methodological (APC=29.1%, Non APC=18.2% χ^2=.2 p-value=.66) and Manipulation (APC=26.4%, Non APC=10.5% χ^2=8.84 p-value=.002). Publishing fraud is the only case in which diamond journals release more editorial notices (APC=39.1%, Non APC=48% χ^2=1.66 p-value=.198). This could be due to diamond journals suffer more from Publishing fraud and they react more to this type of problem. However, these differences are not statistically significant and then we cannot completely conclude that the type of publication fee influences the reaction of the journals to errors and misconduct.

5.3. Peer review system

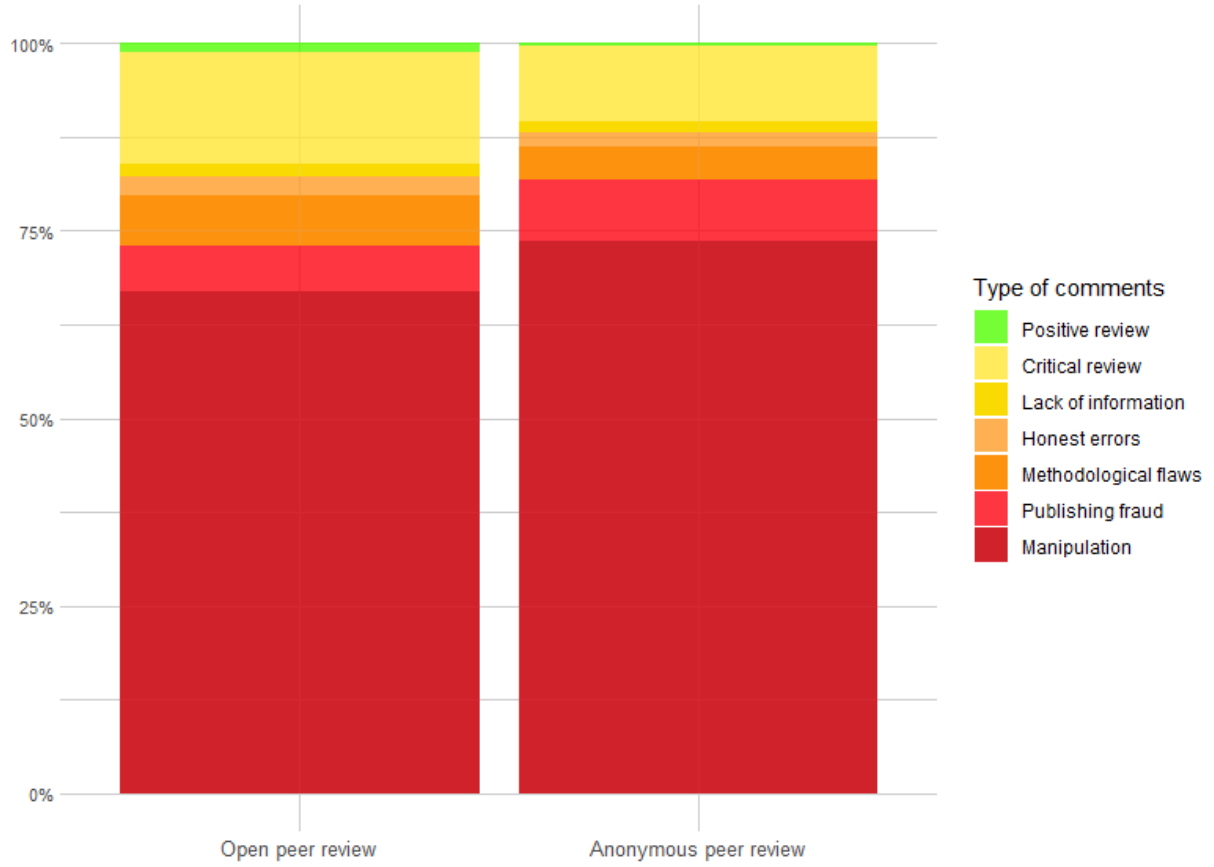

Figure 8. Distribution of articles by type of comment in PubPeer and split by peer review type.

Figure 8 plots the proportion of research articles by type of comment in PubPeer and according to whether journals use an open or anonymous review process. This graph attempts to show if the way in which peer review process is managed could influence the presence of errors and misconduct. The picture shows that journals with an anonymous peer review system have more cases of misconduct (Open peer review=72.9%, Anonymous peer review=81.7% χ^2=5.28 p-value=.02) than journals with an open peer review model. This difference is especially due to Manipulation (Open peer review=66.8%, Anonymous peer review=73.6% χ^2=246.27 p-value<.001) and Publishing fraud (Open peer review=6.1%, Anonymous peer review=8.1% χ^2=4.29 p-value=.04). These results suggest that transparent and open peer review processes would slightly reduce the number of fraudulent studies, avoiding conflict of interests and

compromised peer review. However, these statements have to be considered with caution because the observed differences are barely significant due to the low number of journals that implement this review process.

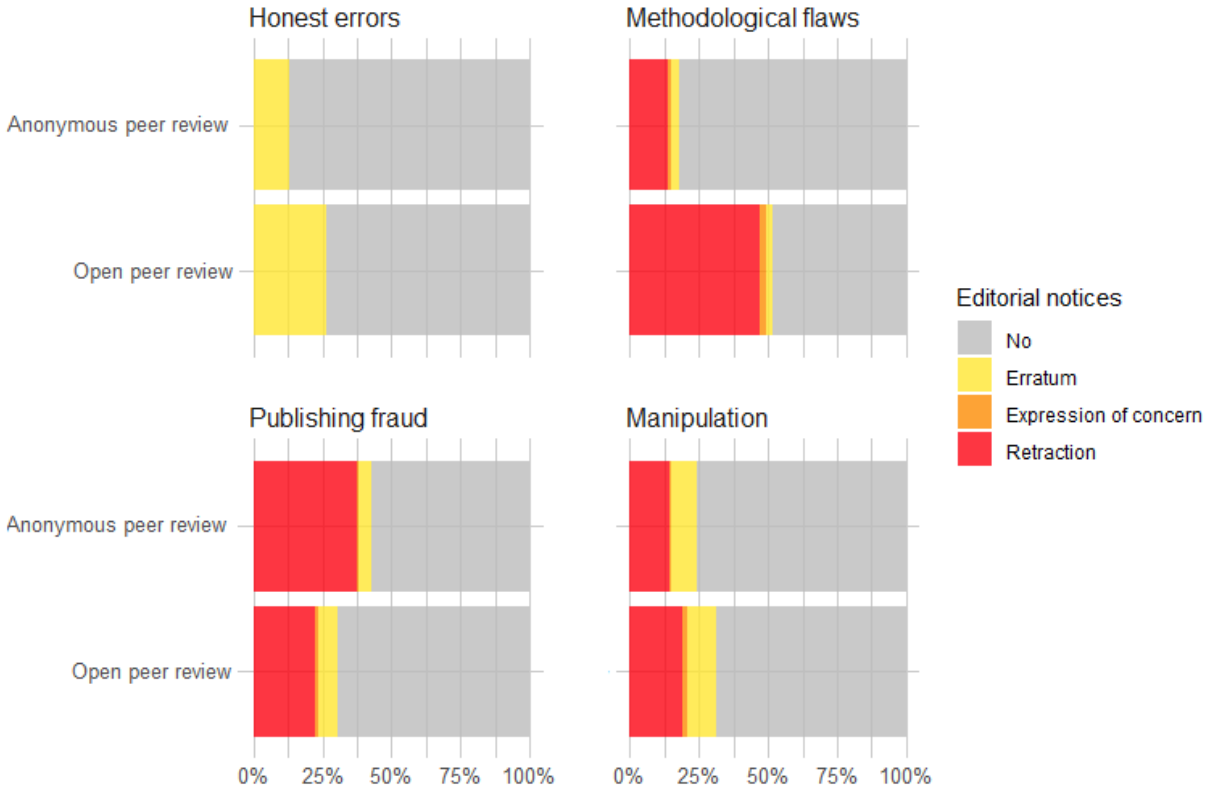

Figure 9. Percentage of publications commented in PubPeer with editorial notices by peer review model and grouped by type of issue.

Type of comment	**Honest errors**				**Methodological flaws**			
Peer review	**Anonymous**	**Anonymous %**	**Open**	**Open %**	**Anonymous**	**Anonymous %**	**Open**	**Open %**
No	75	75.8%	22	71.0%	190	81.5%	39	48.8%
Editorial notices	24	24.2%	9	29.0%	43	18.5%	41	51.3%
Erratum	24	24.2%	9	29.0%	11	4.7%	2	2.5%
EoC		0.0%		0.0%	2	0.9%	2	2.5%
Retraction		0.0%		0.0%	30	12.9%	37	46.3%
Total	99	100.0%	31	100.0%	233	100.0%	80	100.0%
Type of comment	**Publishing fraud**				**Manipulation**			
Peer review	**Anonymous**	**Anonymous %**	**Open**	**Open %**	**Anonymous**	**Anonymous %**	**Open**	**Open %**
No	254	60.0%	50	69.4%	3007	78.5%	552	69.8%
Editorial notices	169	40.0%	22	30.6%	823	21.5%	239	30.2%
Erratum	33	7.8%	5	6.9%	409	10.7%	89	11.3%
EoC	2	0.5%	1	1.4%	18	0.5%	12	1.5%

Retraction	134	31.7%	16	22.2%	396	10.3%	138	17.4%
Total	423	100.0%	72	100.0%	3830	100.0%	791	100.0%

Table III. Distribution of editorial notices by peer review and type of comment.

Finally, Figure 9 and Table III displays the proportion of articles commented in PubPeer with an editorial notice according to peer review model and grouped by type of problem. This result aims to show differences between peer review systems when they react to erroneous or fraudulent papers. The number of publications with anonymous peer review (3,830) is much higher than the ones with open peer review (423), which illustrates the strong reluctance to implement this new model in the publishing system (Thelwall, 2023). Open peer review journals have greater proportion of editorial notices than anonymous peer review in each type of issue, with the exception of Publishing fraud (Anonymous=40% Open=30.6% χ^2=.126 p-value=.72). In Honest errors (Anonymous=24.2% Open=29% χ^2=.002 p-value=.96), Methodological flaws (Anonymous=18.5% Open=51.3% χ^2=4.80 p-value=.03) and Manipulation (Anonymous=21.5% Open=30.2% χ^2=2.61 p-value=.11) open peer review journals tend to react better to errors and misconduct. However, these proportions are not statistically significant at 99%, and therefore we have to be cautious with this result and suggest that even if open peer review journals release more editorial notices, these differences could be small.

6. DISCUSSION

This analysis around the presence of errors and misconduct reported by PubPeer users and the late response of journals to these issues, has made possible to observe how different journal publishing models could be associated to a higher or lower degree of errors and misconduct. The results about the sample of discussed paper on PuPeer show that two thirds of the publications come from commercial publishers, which fits with previous studies (Morris, 2007; Ware and Mabe, 2015). This is the same with the percentage of open access articles (Martín-Martín et al., 2018; Piwowar et al., 2018). These similar figures could suggest that PubPeer is not biased according to type of publishers, which allows us to generalize the findings beyond this journal club.

Scholarly journals suffer more from misconduct, independent of they are open access or paywall, than commercial journals. These results refute our initial hypothesis that commercial publishers could be more permissive to misconduct because their interest in profits may outweigh the editorial control on unreliable articles. Rather, one possible reason would be more related to the quality of the journals than to the type of publisher. Scholarly journals have higher impact (92.7% journals in Q1) than commercial ones (74.9% journals in Q1), which could be more attractive for fraudulent studies. The importance of journals impact in the incidence of misconduct has been previously pointed (Steen, 2011; Fang et al., 2012: Ortega and Delgado-Quirós, 2023), and it suggests that the prestige of the journal would be the main driving force behind the publishing of fraudulent studies. Regarding to the editorial response of both types of publishers, scholarly journals react a little better than commercial ones, but this difference is barely significant and could be explained by a higher presence of errors and fraud. Then, we can conclude that the type of publisher is not a consistent criterion to explain the presence and response to misconduct. However, scholarly publishers do show a significant delay in the response of their editorial notices, which it could indicate that scholarly editorial boards could require more effort and guarantees before to release editorial notices.

Although there are not differences in the presence of misconduct between types of access, open access journals react better to problematic articles than paywall journals. This finding disagrees with the general opinion that open access model favors the publishing of troublesome articles

(Barreiro, 2013; Beall, 2013). This perception could perhaps be motivated by the fact that open access journals release more retractions than paywall journals (Peterson, 2013; Wang et al., 2019; Shah et al., 2021). This misinterpretation is due when assuming that the number of retractions would be an indicator of presence, when in fact is a measure of the ability of the journal to correct articles. Precisely, the methodology of this study allows to distinguish between presence (number of reports in PubPeer) and response (number of editorial notices).

With regard to differences between publication fee models, results have shown a special presence of Publishing fraud in diamond journals. As we have seen before, this fact could be motivated by the impact of the journal, a relationship previously detected by Fang et al. (2012) and Ortega and Delgado-Quirós (2023), and it exhibits that the presence and behavior of certain types of misconduct have different patterns. This could be linked to the fact that many of diamond journals are small and limited to a local audience (Bosman et al., 2021). However, we cannot state that the publication fee model could be an important factor in the detection of errors and misconduct.

A last question is related to the influence of the open peer review in the notification and detection of unreliable articles. The results have shown that journals that use an open review model suffer less from Manipulation and Publishing fraud, which would confirm the opinion that a transparent review process would reduce misconduct cases (Boldt, 2011), or at least the investigations would be more numerous (Schmidt et al., 2018). Our results would also support this last statement, because open peer review journals slightly release more editorial notices. Although this output should be considered with caution, because the statistical significance does not permit a categorical statement. More extensive studies would strengthen this hypothesis.

7. CONCLUSIONS

The main conclusion of this study is that there are differences in the presence and detection of error and misconduct cases according to the type of publisher. Scholarly journals suffer more from problematic papers possibly due to they gain more impact and therefore they are more attractive for misconduct. However, this fact does not cause that the editorial response of scholarly journals is greater.

According to the type of access, open access journals react better to unreliable articles than paywall journals, in spite of that the presence is similar in both cases. This finding contradicts the overall notion that open access journals have more problems due to their high proportion of retraction notices.

Another conclusion is that open access diamond journals have a special presence of Publishing fraud, also motivated by the small size and low impact of this type of venues. Nevertheless, the publication fee model does not show evidences that confirm that APC journals could be more engaged in the detection of errors and misconduct.

Finally, journals that use open review suffer less from Manipulation and Publishing fraud and they slightly release more editorial notices, which could be considered a peer review process more suitable for the detection and reduction of problematic publications.

8. ACKNOWLEDGEMENTS

This work was supported by the research project (NewSIS) “New scientific information sources: analysis and evaluation for a national scientific information system” and funded by the Spanish State Research Agency (grant number PID2019-106510GB-I00). We also want to thank Kim Eddy the proof reading of the manuscript.

9. REFERENCES

American Society for Cell Biology. (2012). San Francisco declaration on research

Assessment (DORA). February 23, 2023 from: https://sfdora.org/
Barreiro, E. (2013). Open access: is the scientific quality of biomedical publications threatened?. *Archivos de bronconeumologia*, 49(12): 505-506.
Beall, J. (2013). The open-access movement is not really about open access. *TripleC: Communication, Capitalism & Critique*, 11(2): 589-597.
Boldt, A. (2011). Extending ArXiv.org to achieve open peer review and publishing. *Journal of Scholarly Publishing*, 42(2): 238-242.
Bosman, J., Frantsvåg, J. E., Kramer, B., Langlais, P. C., & Proudman, V. (2021). The OA diamond journals study. Part 1: Findings. Science Europe. February 23, 2023 from: https://munin.uit.no/handle/10037/22224
Brookes, P. S. (2014). Internet publicity of data problems in the bioscience literature correlates with enhanced corrective action. *PeerJ*, 2, e313. https://doi.org/10.7717/peerj.313
Crawford, W. (2021). *Gold Open Access 2015–2020 Articles in Journals (GOA6); Cites & Insights Books*. Livermore, CA, USA, p. 245
Cokol, M., Iossifov, I., Rodriguez-Esteban, R., & Rzhetsky, A. (2007). How many scientific papers should be retracted? *EMBO reports*, 8(5): 422-423. https://dx.doi.org/10.1038/sj.embor.7400970
COPE (2012). Code of Conduct and best practice guidelines for journal editors. Retrieved February 23, 2023 from: http://publicationethics.org /files/Code_of_conduct_for_journal_editors.pdf
Fang, F. C., Steen, R. G., & Casadevall, A. (2012). Misconduct accounts for the majority of retracted scientific publications. *Proceedings of the National Academy of Sciences*, 109(42): 17028-17033. https://doi.org/10.1073/pnas.1212247109
Gao, J., & Zhou, T. (2017). Stamp out fake peer review. *Nature*, 546(7656): 33-33.
Lei, L., & Zhang, Y. (2018). Lack of improvement in scientific integrity: An analysis of WoS retractions by Chinese researchers (1997–2016). *Science and Engineering Ethics*, 24(5): 1409-1420. https://doi.org/10.1007/s11948-017-9962-7
Hadi, M. A. (2016). Fake peer-review in research publication: revisiting research purpose and academic integrity. *International Journal of Pharmacy Practice*, 24(5), 309-310.
Hawkes, N. (2015). Nature journal's plan to fast track paper for a fee prompts resignation from editorial board. *British Medical Journal*, 350: h1761.
Martín-Martín, A., Costas, R., van Leeuwen, T., & López-Cózar, E. D. (2018). Evidence of open access of scientific publications in Google Scholar: A large-scale analysis. *Journal of informetrics*, 12(3): 819-841.
Marusic, A., Katavic, V., & Marusic, M. (2007). Role of editors and journals in detecting and preventing scientific misconduct: strengths, weaknesses, opportunities, and threats. *Medicine & Law*, 26: 545.
Morris, S. (2007). Mapping the journal publishing landscape: how much do we know?. *Learned Publishing*, 20(4): 299-310.
Nath, S. B., Marcus, S. C., & Druss, B. G. (2006). Retractions in the research literature: misconduct or mistakes? *Medical Journal of Australia*, 185(3): 152-154. https://doi.org/10.5694/j.1326-5377.2006.tb00504.x
Neale, A. V., Northrup, J., Dailey, R., Marks, E., & Abrams, J. (2007). Correction and use of biomedical literature affected by scientific misconduct. *Science and engineering ethics*, 13(1): 5-24. https://dx.doi.org/10.1007%2Fs11948-006-0003-1
Ortega, J. L. (2022), Classification and analysis of PubPeer comments: How a web journal club is used. *Journal of the Association for Information Science and Technology*. https://doi.org/10.1002/asi.24568
Ortega, J. L., & Delgado-Quirós, L. (2023). How do journals deal with problematic articles? The editorial response of journals to articles commented in PubPeer. *El profesional de la*

Información, 32(1), https://doi.org/10.3145/epi.2023.ene.18
Peterson, G. M. (2013). Characteristics of retracted open access biomedical literature: A bibliographic analysis. *Journal of the American Society for Information Science and Technology*, 64(12), 2428-2436.
Piwowar, H., Priem, J., Larivière, V., Alperin, J. P., Matthias, L., Norlander, B., & Haustein, S. (2018). The state of OA: a large-scale analysis of the prevalence and impact of Open Access articles. *PeerJ*, 6: e4375.
Resnik, D. B., Patrone, D., & Peddada, S. (2010). Research misconduct policies of social science journals and impact factor. *Accountability in research*, 17(2): 79-84.
Resnik, D. B., & Dinse, G. E. (2013). Scientific retractions and corrections related to misconduct findings. *Journal of medical ethics*, 39(1), 46-50. https://doi.org/10.1136/medethics-2012-100766
Shah, T. A., Gul, S., Bashir, S., Ahmad, S., Huertas, A., Oliveira, A., ... & Chakraborty, K. (2021). Influence of accessibility (open and toll-based) of scholarly publications on retractions. *Scientometrics*, *126*(6), 4589-4606.
Schmidt B, Ross-Hellauer T, van Edig X and Moylan EC. Ten considerations for open peer review [version 1; peer review: 2 approved]. *F1000Research* 2018, **7**:969
da Silva, J. A. T., & Vuong, Q. H. (2021). Do legitimate publishers benefit or profit from error, misconduct or fraud?. *Exchanges: The Interdisciplinary Research Journal*, *8*(3), 55-68.
Steen, R. G. (2011). Retractions in the scientific literature: do authors deliberately commit research fraud? *Journal of medical ethics*, *37*(2), 113-117.
Stricker, J., & Günther, A. (2019). Scientific misconduct in psychology. *Zeitschrift für Psychologie*.
Thelwall, M. (2023). Journal and disciplinary variations in academic open peer review anonymity, outcomes, and length. *Journal of Librarianship and Information Science*, https://doi.org/10.1177/09610006221079345
Tripathi, M., Sonkar, S., & Kumar, S. (2019). A cross sectional study of retraction notices of scholarly journals of science. *DESIDOC Journal of Library & Information Technology*, *39*(2), 74-81. https://doi.org/10.14429/djlit.39.2.14000
Vuong, Q. H. (2020). The limitations of retraction notices and the heroic acts of authors who correct the scholarly record: An analysis of retractions of papers published from 1975 to 2019. *Learned Publishing*, *33*(2), 119-130. https://doi.org/10.1002/leap.1282
Wager, E. (2007). What do journal editors do when they suspect research misconduct. *Medicine and Law*, *26*(3), 535.
Wang, T., Xing, Q. R., Wang, H., & Chen, W. (2019). Retracted publications in the biomedical literature from open access journals. *Science and engineering ethics*, *25*(3), 855-868.
Ware, M., & Mabe, M. (2015). The STM Report: An overview of scientific and scholarly journal publishing. *Copyright, Fair Use, Scholarly Communication, etc*. 9. http://digitalcommons.unl.edu/scholcom/9
Williams P, Wager E (2011) Exploring why and how journal editors retract articles: findings from a qualitative study. *Science and Engineering Ethics*, 19(1), 1-11, doi:10.1007/s11948-011-9292-0