Reviewer's report

Title: "Predatory" Open Access - An overview

Version: 2 Date: 29 May 2015

Reviewer: Jean-Claude GUÉDON

Reviewer's report:

A. Major compulsory revisions.

1. Language

The first important revision is to have the writing of this paper revised in depth, preferably by a native English speaker, or someone with near-native ability in English. Sometimes, the writing is downright opaque. Let me give a couple of examples:

a. "It is, however, very difficult to generalize from these figures to all predatory journals since more than half of the included journals were published in India, which might not be the case if all journals from publishers with multiple journals are included as well." This sentence makes little sense.

b. "Also generalized to all predatory articles, the overall average APC is only half as high (135 USD) as the average calculated over journals, indicating a clear author preference for lower priced journals, leading to higher publication volumes in these." This sentence is quite opaque.

c. "The share coming out of the authors own pocket would probably be much higher for predatory journals, since the threshold for asking for institutional or external funding for the APCs in highly questionable journals would seem quite high." What does this mean?

Conclusion: cleaning up the writing of this paper will begin to improve it greatly. It is a sine qua non condition.

2. Important questions:

a. I am no statistician, but the sampling of journals appears to have been done with some care. However, the determination of the sampling weight of each journal appears as potentially problematic in the following passage:

"This method may from a statistical viewpoint have some problems, but the quite diversified results that we obtained seem to warrant our choice."

A brief discussion of the validity of the method, and its limitations, would be useful. otherwise, we are faced with a certain degree of apparent statistical sophistication that does not seem sufficient to alleviate doubts about the sampling technique. How are we to trust the rest of the analysis?
b. All the results about article distribution, etc. should be compared to results taken from the Web of Science and from DOAJ. There is a clear lack of control metrics in this paper.

c. The 7% titles that belong both in Beall's list and DOAJ should have been scrutinized further: they might lead to interesting results about what might be called "border-line" cases.

B. Minor essential revisions

a. If it is difficult to see whether a publisher really is based in the USA, why should we trust any other location, short of saying that everyone wants to appear as if he/she works in the US. This might well be an artefact of Beall's behaviour as he seems to believe that publishers outside the North Atlantic are not to be trusted, almost by default. I remember his taking Hindawi to task, in part because it was based in Cairo.

b. The introductory section aiming at clarifying what "predatory journals" are could be shortened.

c. On the other hand, Beall's criteria for predatory journals could have been analyzed, scrutinized and criticized. This could have supported an analysis of inherent bias alluded to in the text, toward the end.

d. The explanation for the small presence of predatory journals in Latin America is very superficial. It should be noted that Scielo is not the only player in town: Redalyc also plays an important role in scholarly publishing in Latin America. Moreover, many journals are published within universities and appear in the web sites of these institutions. To some extent, this institutional branding will limit rogue behaviour because of concern for institutional reputations. Latindex provides a list of several thousand journals that have been vetted according to more than 30 criteria. Many countries in the region use quality lists derived from Latindex (e.g. Qualis in Brazil). Although not perfect, these national lists offer an added barrier to rogue behaviour. On the other hand, not all Scielo journals are in the WoS.

e. The control on articles attempted through "Scholar One" may introduce another form of bias that is not even acknowledged.

f. The variations in APC's are not adequately interpreted. It is quite difficult to draw any general lesson from what is said in the paper.

g. Waivers for APC's are generally applied by countries and according to GDP/person statistics. On such a scale, India is quite low, but India is sometimes excluded from such waivers on account of its overall size and the existence of a sizeable middle class. This point is not discussed, despite the importance granted to India in this study.

h. The number of authors in predatory journals compared to the number of authors in WoS-indexed articles does not tell the whole story, given the
well-established difficulties of "peripheral" scientists to publish in "core" journals. How about comparing that number with DOAJ articles, particularly with journals that are gratis for authors.

i. p 1: Non-commercial journals are not always run by volunteers (which, I assume, is mentioned to "explain" how such journals can be without fees. Many journals are subsidized, including subscription-based journals, and including journals in rich countries. The case is frequent, for example, in Canada where large programs exist to subsidize journals and monographs. There is no reason to intimate that journals that are gratis for authors are somehow "amateurish" or "unsustaintable" because dependent only on "volunteers".

**Quality of written English:** Not suitable for publication unless extensively edited

**Statistical review:** Yes, but I do not feel adequately qualified to assess the statistics.

**Declaration of competing interests:**

I declare that I have no competing interests