Cites & Insights

Crawford at Large/Online Edition

Libraries • Policy • Technology • Media

Volume 17, Number 8: September 2017 ISSN 1534-0937

The Front

Maybe I'm Doing it Wrong

Hat-tip to Randy Newman for the title, and the studio version's as good as any. It's a feeling I've been getting, especially about my investigations into Beall's lists and the number of articles in non-*DOAJ* open access journals—and maybe about my OA writing in general.

This is not a whine about not getting speaking invitations: I explicitly removed myself as a speaker on OA, at least for now. There are far too many younger, wiser, more relevant speakers; I've done my bit in the past.

Nor is it a whine about readership. *Cites & Insights* continues to get pretty good numbers. Here are the figures through August 31, 2017 for the six issues I looked at for this discussion:

THE SAD CASE OF JEFFREY BEALL, April 2014 (14.4): 18,757.
JOURNALS AND "JOURNALS"(2),Oct./Nov. 2014 (14:10): 2,912.
PPPPREDATORY ARTICLE COUNTS, January 2016 (16:1): 2,350.
"TRUST ME...," April 2016 (16:3): 2,929.
GRAY OA 2012-2016, January 2017 (17:1): 1,660.
THE ART OF THE BEALL, April 2017 (17:3): 1,635.

That first number is absurdly high, by far the highest since stats restarted (The LIBRARY 2.0 AND "LIBRARY 2.0" issue is still the overall champion, with more than 37,000 reads), and I regard the other numbers as entirely respectable. The stuff's getting read (or at least downloaded) by *my* audience and, in the SAD CASE case, by a much broader audience.

Inside This Issue

Perspective: Where I Stand: OA, "Predatory," Blackkists, The Bealllists and Thought Leadership.....4

So What's the Problem?

One problem is that THE SAD CASE OF JEFFREY BEALL seems to have become a token, a citation thrown in as a sign that authors of an article decrying

"predatory" journals (and generally praising Beall or at least the lists) have considered the "other side."

Consider: Google Scholar shows 29 citations to that essay—and *none* to any of the others, including "Trust Me," which is a much more cogent and compact reason why the lists are bad tools and should be ignored.

Consider: Google searches yield 53 hits for that issue and no more than 25 for any of the others, and if you exclude cases where I've cited the others in comments to articles (and the work of John Dupuis), there are precious few left.

Anybody want to suggest a ratio between articles citing the Shen/Björk "420,000 articles in predatory journals in 2014" nonsense and those citing my demonstrations that this figure is off by as much as an order of magnitude? I'm not sure what that ratio is, but offhand I remember few if any articles that cite the alarmist numbers and even hint that they might be exaggerated.

I'm not saying "Pay attention to Walt Crawford." I am saying it would be nice if authors paid attention to the facts. (Set aside *Gold Open Access Journals*—there, I believe, the SPARC sponsorship provides enough legitimacy that the books are read and paid attention to.)

Affiliation and Venue Matter More than Facts?

I hate to make such a suggestion, but it sure does look that way.

Recently, a reporter was writing a report about one of the "we're drowning in predatory journals and Something Must Be Done." Atypically, the reporter contacted me. Even more atypically, the reporter asked cogent questions. Almost uniquely, the reporter followed up my responses with additional requests.

After that, it was *entirely* typical: not only was I not quoted, but the *factual issues* I raised were ignored. So far, my score in that regard is perfect. I find that disappointing.

Did I mention that the followup questions also asked what institution I was affiliated with?

Do I believe the ultimate outcome might have been different if my response had not been, in essence, "None"?

I'm not sure.

Maybe That's Appropriate?

I'd love to say that it's inappropriate for reporters and scholars to pay attention only to writing affiliated with universities and the like and ignore writing from ~~the great unwashed~~ unaffiliated researchers.

But I'm not sure I can—especially because there's also the peer-review issue: My research doesn't generally appear in peer-reviewed journals for several reasons.

Given the sheer flood of material, maybe it's only prudent for people to only pay attention to stuff that arrives with the right credentials. I can appreciate that. (Albert Einstein wouldn't, but I'm no Albert Einstein.)

Why don't I submit articles for peer review?

Partly practicality: as an unaffiliated researcher, I'm not in a position to do the literature reviews I'd need for proper papers in the humanities and social sciences. Partly speed and length: my analyses tend to run long—and I'd like to get them out while they're still fairly current. Partly economics: I don't have funding for article processing charges, although that wouldn't be an issue for some of the OA journals I might consider.

Then again, I can attest that there are probably hundreds of OA journals (and, I'd guess, subscription journals) that *explicitly* rule out contributions from The Great Unwashed: They won't accept manuscripts where the contributing author isn't a PhD (or at least a doctoral student with their advisor as another author).

So, in a way, the fact that my stuff gets used at all is a minor miracle—even if the list-related stuff seems to be used as a token of considering the controversy.

Consistency?

Ah, but if you reject non-peer-reviewed material, how can you *possibly* support Beall's lists?

And if you think PhDs should be writing the scholarly papers—well, although some Indian media call him "Dr," Beall has no doctorate that I know of.

Of course, you could look at the evidence he provided for list entries and conclude that he was right—but it's abundantly clear that this hasn't happened, since any examination of the evidence would have to throw out seven out of eight listings.

I'm tired of writing about Beall, whose visibility seems to keep rising even after he decided to abandon the fatally-flawed lists. Which is another way of saying I'm tired of reading jeremiads about the flood of "predatory" publishing that inherently relies on one man's definition of "predatory" and requires taking his unsupported word on who's predatory. (I'm even more tired of apparently reputable scholars who decide to take his lists at face value and then base so-called research on them.)

Maybe I'm just tired: after all, this issue will appear just about the time I turn 72, and this summer was marked with an attack (either Bell's palsy or Ramsay Hunt syndrome 2) that has further reduced my energy.

Not quite tired enough to give up, which may be stupidity on my part. There will be a followup of sorts to *Gray OA*, a simpler piece looking at how these 18,961 journals and "journals" fared in the last half of 2016 and the first half of 2017, possibly followed with a series of brief profiles of prolific publishers and "publishers." It will probably get even less attention

than *Gray OA* did—and that's OK. The facts will be out there, and some people will use them.

Is Independent Scholarship Feasible?

That was a working title for part of this little piece—but over time I realized that I really don't have a useful perspective on that broader issue. I'll suggest that independent scholarship and research are more difficult than in the past, but "feasible" is too broad a word. I hope the answer is yes.

Perspective

Where I Stand: OA, "Predatory," Blacklists, The Beall lists and Thought Leadership

or why 140 characters is less than 1% of what I need to say about this cluster of topics.

The tl;dr Version

Blacklists in general don't work and are ethically questionable—and the Beall lists in particular are so fundamentally flawed that it's irresponsible of any librarian or scholar to advise using them. There are journals that fall short of good practices in different ways, some of which could be called "predatory"—but defining "predatory" continues to be difficult. While I can identify (and have identified) some characteristics that I'd say make a journal unsatisfactory for research papers, I am not qualified to issue more general edicts, even if I wished to do so. I'm not a thought leader on OA—because I don't believe in "thought leaders," because I'm more interested in digging into the facts than in guiding policy, and because I don't get listened to much in any case. And OA still has two colors, both of which require somewhat miraculous occurrences to become wholly successful, and a bunch of other terms and parties working to confuse the issues. OA should be first and foremost about access to research—but it should also, unless done badly, save money.

Prefatory Material

That's the short form: 1,065 characters or eight tweets—but the short form is too short to be worthwhile. In the 2,500 (or more) words following this prefatory material, I'll try to provide a bit of my current reasoning. Not that the tweeter who caused this piece to be written will read it, as he (*of*

course it's a he) seems to feel that I'm derelict in my duties as an OA advocate (which I'm not) if I don't immediately tweet short answers to difficult questions. The rest of you may find this useful, or not—and it's surely shorter than reading everything I've written about OA at *Cites & Insights*!

If you're wondering about the latter, I wrote enough to fill a 512-page paperback between 2001 and 2009. You can still download that PDF for free from Lulu. I stayed away from OA *in C&I* during 2010 through 2012 (with the exception of one tangential essay at the end of 2012), but I wrote *Open Access: What You Need to Know Now* as an ALA Special Report in 2011.

Since then, I've done several books on the gold OA landscape (some sponsored by SPARC, all available at lulu.com) and a handful of *Cites & Insights* articles and full issues: February and June 2013; April, May, August, October/November and December 2014; March, April, June, July, October and December 2015; January, February/March, April, May June and September/October 2016; and January, April, May, June and July 2017—so far. The total word count for those essays—*excluding* three cases where all or most of an issue presents what's also in a free ebook—is around 374,000 words. Which would make up something over 1,000 6" x 9" book pages. Of course, many or most of those words aren't mine. I haven't provided links here; the Tables of Contents page at *C&I* includes all essay titles since 2011 (with a link to the 2001-2010 contents) with links to either specific essays [2011-2012] or full issues [2013-].

I suspect most everything I say here is either said or hinted at somewhere in that mass of verbiage; I'm trying to provide a concise version here, without the argumentation. Three essays may be especially useful in expanding on what I say here: THINKING ABOUT LIBRARIES AND ACCESS, TAKE 2 [July 2015], ETHICS AND ACCESS 2015 (and especially "Blacklists: A Level Ground") [December 2015], and THE ART OF THE BEALL, including Cameron Neylon's expert takedown of blacklists in general [April 2017].

As to the tweetstorm that eventually resulted in this essay, I don't see much point in going over it in detail. It was, to my mind, a prime example of why some discussions just don't lend themselves to Twitter—especially when one party in the discussion seems determined to misinterpret what's being said. Will some of you find the [tweet*>100] that follows enlightening? Probably not if you read *C&I* on a regular basis, but one can hope.

The Longer Version

I'll use portions of the tl;dr paragraph as section headings.

Blacklists in general don't work and are ethically questionable

While I've stated my case in this regard previously, Cameron Neylon's discussion is so clear and eloquent that you should read his thoughts. Summarizing key points: blacklists are technically infeasible and can't be

complete; they're practically unreliable (partly because they're inherently discriminatory); and they're inherently unethical.

Here's more (from Neylon) on that last and most important point:

> Blacklists are designed to create and enforce collective guilt. Because they use negative criteria they will necessarily include agents that should never have been caught up… Blacklisting publishers seeking to experiment with new forms of review, or new business models both stifles innovation and discriminates against new entrants. Calling out bad practice is different. Pointing to *one* organisation and saying its business practices are dodgy is perfectly legitimate if done transparently, ethically and with due attention to evidence. Collectively blaming a whole list is not.
>
> Quality assurance is hard work and doing it transparently, consistently and ethically is even harder. Consigning an organisation to the darkness based on a mis-step, or worse a failure to align with a personal bias, is actually quite easy, hard to audit effectively and usually over simplifying a complex situation. To give a concrete example, DOAJ maintains a list of publishers that claim to have DOAJ certification but which do not. Here the ethics is clear, the DOAJ is a Whitelist that is publicly available in a transparent form (whether or not you agree with the criteria). Publishers that claim membership they don't have can be legitimately, and individually, called out. Such behaviour is cause for serious concern and appropriate to note. But DOAJ does not then propose that these journals should be cast into outer darkness, merely notes the infraction.

The Bealllists in particular are so fundamentally flawed that it's irresponsible of any librarian or scholar to advise using them

I'm calling out librarians here because at least a few have expressed sadness at the loss of the lists and opined that they might not be perfect but they were convenient, easy tools. That's irresponsible: they were fundamentally flawed tools, equivalent to offering a reference work where seven out of eight entries were generated out of thin air.

Setting aside some of the questionable criteria and Beall's known antipathy for OA journals, here's the key (quoted from a January 16 blog post, appearing in different form elsewhere):

> Beall's lists don't meet what I would regard as minimal standards for a blacklist *even if you agree with all of his judgments.*
>
> Why not? Because, in seven cases out of eight (on the 2016 lists), Beall provides no case whatsoever in his blog: the journal or publisher is in the lists Just Because. (Or, in some but not most cases, Beall provided a case on his earlier blog but failed to copy those posts.)

Seven cases out of eight: 87.5%. 1,604 journals and publishers of the 1,834 (excluding duplicates) on the 2016 versions have no more than an unstated "Trust me" as the reason for avoiding them.

I believe that's inexcusable, and makes the strongest possible case that nobody should treat Beall's lists as being significant. (It also, of course, means that research based on the assumption that the lists are meaningful is fatally flawed.)

There are journals that fall short of good practices in different ways… Yes, of course there are. Some of those journals are OA; some are subscription. Some are from fly-by-night operators—and some are from the biggest names in the business. Quoting another post: Pretty much every publisher will occasionally publish a "bad" paper, possibly one that some scholars think is "obviously" bad, possibly one that's plagiarized. Pretty much every publisher will have at least one journal where at some point the editorial board or peer review may involve issues (excessive publication, editorial overrides, etc.).

A journal that lies about what it does is unacceptable.

Other and related examples of falling short:

- Charging an *unstated* article processing charge, one that's not defined until the article's been accepted. (Or, perhaps worse, billing for a charge that was never stated.)
- Journal sites with malware, where the prey is the reader. Usually, these are abandoned journals or journals attempting to gain ad revenue from anything-goes ad networks.
- Claiming to do peer review when no peer review has been done—but tell me there haven't been cases where subscription-journal editors overruled negative peer reviews, including cases on behalf of their personal friends.
- Charging for open access but not actually providing it, or providing it only when directly challenged (seems to be a "hybrid" journal failing).
- Creating journals that consist entirely (or mostly) of articles from other journals.
- Calling your journals "leading journals" or yourself a "leading publisher" when most or all of your journals are empty or have trivial numbers of articles and you have *no* established reputation.

There are doubtless others. Having too broad a scope? Paging *Nature* and *Science*, among others. Duplicating the scope of an existing journal? Paging probably 90% of all journals, including most newer subscription journals. I have trouble including things like this except under the "everything's pppredatory" label.

Defining "predatory" continues to be difficult

What makes a journal or a publisher predatory? *Who's the prey*?

Sometimes that seems clear enough—for example, when the prey are academic libraries compelled to pay outrageous online fees and take thousands of journals they don't want in order to meet their scholars' needs. The predators, in this case, are the big subscription journals.

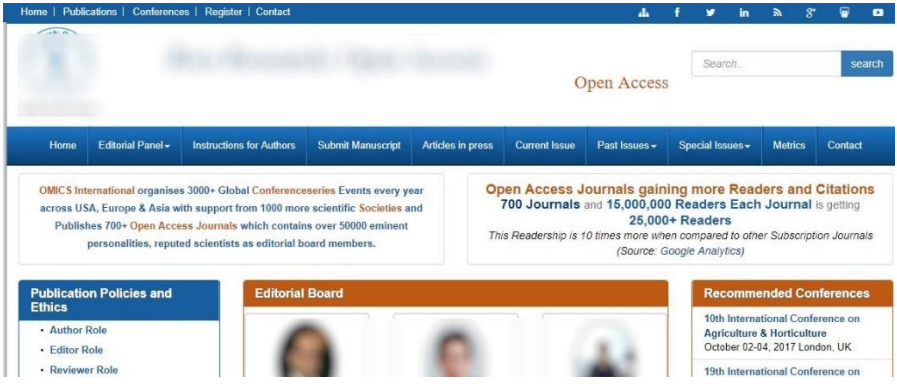
I've already mentioned hidden/unstated APCs, where the prey are scholars—and, to be sure, journals that don't do what they promise to do (whether peer review or editorial work).

Beyond those, I'm not sure. I don't believe that authors openly choosing low-fee or no-fee journals are being preyed upon: they're choosing a form of publication that suits their needs. (I honestly don't believe there are hundreds of thousands of *bad* science articles published each year because of all the more obscure OA journals: why would scholars knowingly publish *bad* science, as compared to the many small-science papers?)

Are readers and the public being preyed upon because of second-rate or badly-vetted papers? If so, you need to consider overall negative impact, and I'll suggest that one paper in a *highly* respectable subscription journal, falsely linking autism to vaccination, has probably caused more damage and deaths than all the badly-vetted OA articles put together.

I don't have answers here. I believe "predatory" is being used far too broadly and with far too little effort to define what makes something predatory and who the prey is.

That's not quite the same as finding journals that most sensible authors would avoid after visiting the site (and an author who submits an article *without* visiting the journal site does not deserve the word "sensible").



So, for example, if the top half of a journal's home page looks like the one captured here (with key identifying elements obscured), I think that the publisher is a lot more interested in tooting its own horn and making odd or questionable claims (why would 3,000 global conferences be a *good* thing? should I really believe that a typical OA article will get anything like 25,000 readers?) than it is in informing me about *this particular journal*. Predatory? I can't say—but I wouldn't submit there.

Similarly, if I see this paragraph as the only text (other than journal name, publisher name, menu and ISSN "tbd") on a home page:

> Generic Journal is open access Journal. Generic Journal is aim to published journal monthly basis and Generic Journal is covering full length research article, review article, case study, short communication and etc. Generic Journal is covers all the fields of Generic research related topics. Generic Journal is inviting you to submit your valuable unpublished research paper and paper will publish without delay.There is no page limitation in this journal.

That's taken directly from one of several hundred templated home pages, identical in every regard except for the words I've replaced with "Generic" in this case. Between "is open," "is aim to published journal monthly basis," "is covers all the fields of" and "paper will publish without delay," there's so much bad English there—especially since in this case all these journals are all called American—that I'd steer clear.

Incompetent? Perhaps—and in that latter case, it's probably worth noting that 370-odd "journals" with that template published a *total* of 16 articles in 2014, six in 2015, two in 2016 and *none* in the first half of 2017. If these are "predatory" journals, they're ineffectual predators.

I am not qualified to issue more general edicts [on sketchy journals]

Is Journal X in any field other than, possibly, media or librarianship, questionable or predatory? Other than the obvious signs—some mentioned above—I can only say "How would I know?"

I am 100% unqualified to say whether a biomed journal has sufficiently high standards for study reporting and upholds those standards, or whether its articles are good science. I am not qualified to judge the coherence and practices of articles in a journal on physics, herbal medicines, optics, engineering, ecology, mathematics…

Other than your own subfields, isn't the same true for you?

For that matter, I'm not qualified to say whether a new journal is truly redundant or meets a new need. Neither are you, I'll venture.

I'm reluctant to say that one or two bad articles mean that a journal is rotten—and even more reluctant to say that one or two questionable journals mean that a publisher (or a publishing platform!) should be denounced in its entirety.

So if someone says "Name five predatory journals," after arguing about the word "predatory" I'll say "How should I know?"

Sure, I can probably go through 20 or 30 journal sites for OA journals that aren't in *DOAJ* (I'm visiting nearly 19,000 of them at the moment) and come up with four or five where I'd be uncomfortable submitting an article (if the journal was in a field I write about)—but that's a little different. I believe any scholar who thinks they have an article worth publishing should be willing to use a little sense and a little time to vet a possible journal. I've offered some quick guidelines for doing that in the past, but they're mostly just sensible:

Check *DOAJ*. If the journal's listed, that's a strong positive, but the other steps may still make sense. Look at the journal's home page. Evaluate the information you're provided. Look at the recent publishing history. Skim through an article or two—would you publish where you wouldn't read? If you're acquainted with an editorial board member, maybe drop them an email asking about the journal. Follow your instincts: If in doubt, there are always other venues. You don't need me to tell you that.

I'm not a thought leader on OA

Getting past "predatory" journals, a few brief notes in some other areas.

I don't believe in "thought leaders" in general; it strikes me as an elitist old-white-boy concept.

More to the point, I'm a poor candidate to give you The Truth about OA as it stands today. I believe my 2011 ALA Special Report *Open Access: What You Need to Know Now* (still available, but a bit pricey for its length) did a good job of providing a quick overview—but this is 2017, not 2011.

I write commentaries that cite interesting articles and posts about OA and comment on them. I also do detailed research on factual aspects of OA—specifically, the nature and extent of gold OA publishing.

I believe that both are useful functions, but they're quite different from either giving The Word or providing stirring keynotes on moving OA forward. I guess that's what "thought leaders" would do, and if that's what you need, I'm not your man. I do have facts—and get frustrated when they're ignored. But that's another essay, which you may have just read.

OA still has two colors…

Gold OA, which has *nothing* to do with article processing charges, and green OA, which has everything to do with digital repositories.

That's it.

Gold OA consists of journals that make all peer-reviewed articles (and articles subject to post-publication peer review) freely *and immediately* available for any user with internet access to read in full, download, and pass along to others, without registration or fee. Most gold OA journals that appear in *DOAJ* do not require article processing fees, being supported through other means. Journals that *do* charge APCs tend to publish more

articles than those that don't, at least at the moment. (Beyond *DOAJ*, numbers get confused by whether or not something can be considered a journal.) An overlay journal can be an entirely proper gold OA journal.

"Diamond" and "platinum" are just fancy ways of saying non-APC gold, and mostly confuse the picture. My own term "gray OA" is probably unfortunate: it consists of those gold OA journals that are *not* in *DOAJ*.

Green OA consists of articles that have been or will be peer reviewed, deposited in some version to a personal, institutional or thematic internet repository, and freely *and immediately* available to anyone with internet access to read in full, download, and pass along to others, without registration or fee. The distinction is that the article may not be in the precise form in which it has been or will be published in a journal. In my mind, if a publisher can require that an article in a repository be withdrawn or made unavailable even for a temporary period, *it's not green OA*.

Two colors are enough.

…both of which require somewhat miraculous occurrences to become wholly successful

Here's where it gets tricky and where I admit that I think it unlikely that we'll have 100% OA (either for prospective or retrospective) research material in my lifetime. Why is that?

For gold OA to become the universal practice in all fields where peer-reviewed work is published, we'd either need a massive "flip" (which I believe would be economically unfortunate and a bad thing in general) or for all scholars to move their work to cost-effective gold OA platforms, which might not be journals as such. That might provide universal OA or the future, but it would also strike me as unlikely to the point of being miraculous. Maybe I'm wrong. But I doubt it.

For green OA, the case is simpler. As long as green OA consists of variant versions of papers appearing in subscription journals, where the journal's version is the citable version *and the journal name provides the paper's prestige*, green OA will do *nothing* to reduce the costs of scholarly publishing or its domination by a handful of big publishers. Until, to be sure, A Miracle Occurs: scholars, libraries, funders and all other parties agree that journals are passé and that existing publishers should, at most, be paid modest sums to facilitate peer review, copyediting and layout. To my mind, that miracle is even more unlikely than the gold-OA miracle—and, of course, does equally little to free the existing literature.

Again, I could be wrong: I'm mostly an observer.

OA should be first and foremost about access to research

I am not now and never have been a professional librarian. I call myself a library professional: while I haven't worked in a library since 1979, I've always worked *for* and *with* librarians, and my professional interests have

been in the library field. I will tell you this about my own *very* early involvement in OA and, I believe, the involvement of most library people: It's been first and foremost about making research accessible.

Accessible to those in institutions too small or poor to be able to pay for every journal their students and scholars could possibly want (and since Harvard's been in that category for years, that includes most institutions). Accessible to those *without* institutional access.

Because access to research and information is a core library value.

But [OA] should also, unless done badly, save money

When I see non-librarians claim that librarians only support OA because they're out to save money, it raises my hackles: I don't believe that's ever been true for most librarians.

But barring the idiocy of universal "flips" at current income rates and with publishers allowed to keep raising rates as they see fit, or the equal idiocy of thinking green OA will do just fine and Magic Will Happen, I believe it's fair to say that a mostly-OA future *should* save money for libraries and others, money that can be put to good use in supporting monographic purchases, preservation and conservation, and all the other things libraries do.

I believe that's a strong *secondary* reason to support OA, and to do so strategically, not with "take all the money you want, but make the articles open" approaches. Because it's becoming clear that, if you do the latter, the average cost per article will start at $5,000 and go up each year.

140 is Less than 1%...

I was aiming for *slightly* over 14,000 characters or 100 tweets for the longer version here. As it stands, it's 17,195 characters or 123 tweets. That's not bad, given what a wordy old coot I can be.

Does this leave you better informed on where I stand and why I don't pretend to be an OA guru? I hope so.

Masthead

Cites & Insights: Crawford at Large, Volume 17, Number 8, Whole # 207, ISSN 1534-0937, a periodical of libraries, policy, technology and media, is written and produced by Walt Crawford.

Comments should be sent to waltcrawford@gmail.com.

URL: citesandinsights.info/civ17i8.pdf