

■ Commentary

Sense and nonsense of bibliometrics

Pedro J. Aphalo, ORCID: [0000-0003-3385-972X](https://orcid.org/0000-0003-3385-972X)

ViPS, Organismal and Evolutionary Biology, University of Helsinki, Helsinki, Finland

DOI: [10.19232/uv4pb.2018.1.10](https://doi.org/10.19232/uv4pb.2018.1.10) © 2018 The Author, licensed under (CC BY-SA 4.0)



The facts

Several times during recent months, according to ResearchGate statistics I was the most read author from Finland, with dramatically increasing numbers of reads per week for one specific article. The article in question is approaching 100 000 reads in its life time of less than two years. This number of reads is 89% of the reads over their full lifetime accumulated by the 118 research items of my authorship tracked by ResearchGate since I created my profile several years ago. Because of this I had some time ago the exciting message shown in Figure 5.1 waiting for me when I visited the ResearchGate web site. The first of these messages alerted me on something unusual happening in my profile, and since then I have been receiving long strings of unlikely ResearchGate achievements week after week (Figure 5.2). Some of these apparent achievements reveal that fields of research are assigned relying on users' profiles—in the example lead astray by matching my expertise in “Mineral Nutrition” to an achievement in “Nutrition and Dietetics”.

Which of my publications could be so popular? and where could it have been published? A publication that is taking my bibliometric statistics to heights seen never before during my whole career! By further navigating the ResearchGate web site I found the answer: the article in question is a book review published in 2016 in this same *UV4Plants Bulletin* (Aphalo 2016)!

By plotting reads-per-month on a log scale

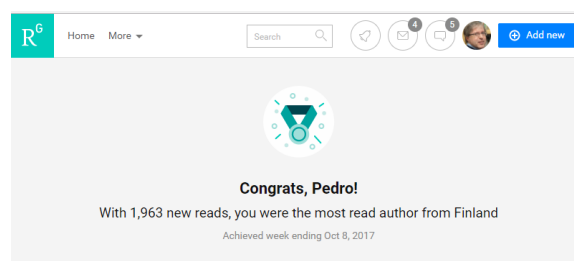


Figure 5.1: An exciting message from ResearchGate.

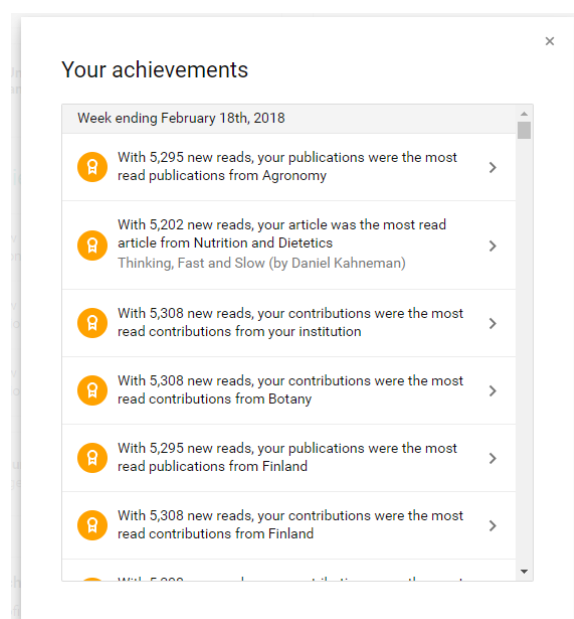


Figure 5.2: The top of my list of ResearchGate “achievements” for a recent week.

as a function of time we see that growth in the number of reads has been nearly exponential for a whole year (Figure 5.3). The initial lag was in part due the delay between

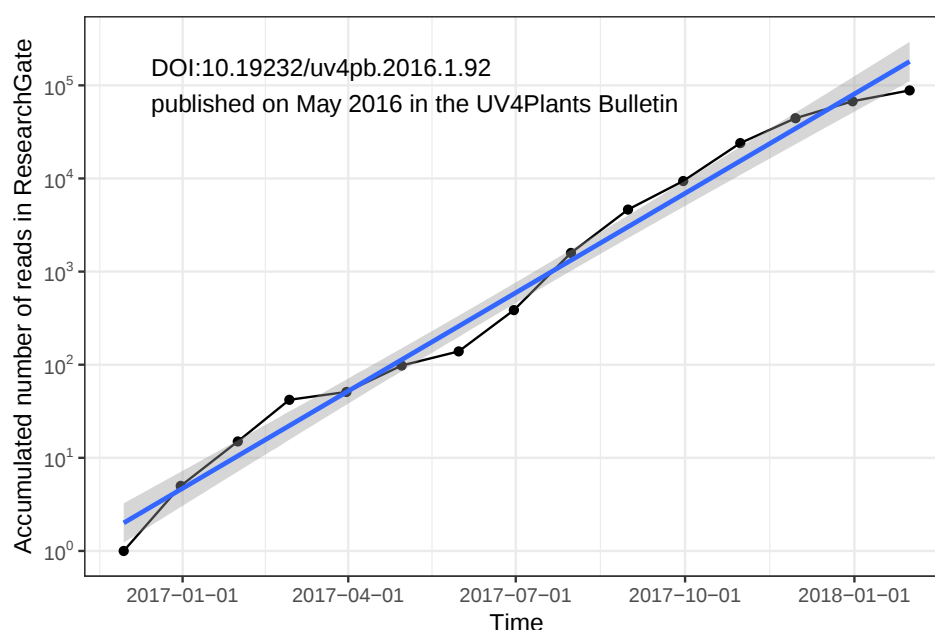


Figure 5.3: The exponential growth in ResearchGate reads of my book-review article [my book-review article](#).

publication in the Bulletin and upload of the offprint file to ResearchGate. An important question is: what kind of “snowball” process could be driving this unexpected pattern of growth in “interest” in a publication of minor significance? In the next section I discuss the possible causes of exponential growth in reads and the implications of the event described in relation to the use of “objective measures of research impact” in academic evaluations.

Pondering on the vagaries of bibliometrics

I think this event tells us a lot about the measurement of the impact of publications and research and the problems of using them for assessing performance. First of all it tells us, that a simple book review with no original research component can bias indexes in an extreme way. In particular, I suspect that “reads”, as opposed to citations, depend a lot more on the choice of title and keywords than on the content of articles.

When searching with Google and similar “engines” the ordering of search results we see is not random. There are two factors at play: payment of a fee by web site owners and the application of complex mathematical algorithms to guess which search results most closely match the expectations of the user who has submitted the search request. Those results “floated to the top” as a result of a fee paid to Google are discreetly labelled as “ad”. For other results Google uses the previous search history of the user submitting the current search and also the overall popularity of pages. Popularity of pages is inferred from number links to them in other pages and from how frequently they have been visited when they have been shown in earlier search results. A further source of information used is the popularity of keywords and similar content including words in titles. In other words, algorithms used by on-line search engines determine to a significant extent what internet content becomes popular or “goes viral”.

What web site developers call SEO, or

search-engine-optimization, is the manipulation of the contents of a web page to improve its positioning/ranking in the output from on-line search engines. Optimization ranges from careful selection of web page titles and URL texts, to embedding in pages hidden content unrelated to the visible content of a page to “trick” search engines. Numerous free resources on SEO are available in the internet (e.g. the guide at <https://moz.com/beginners-guide-to-seo>).

From the point of view of the UV4Plants Bulletin, events like this, increase visibility and may improve our Bulletin’s SEO through their impact on search results. From a practical point of view, it also shows that uploading all off-prints and pre-prints that copy-right allows us to, to a site like ResearchGate can boost our “numbers”. An improvement in apparent impact, even though possibly accidental in origin, can affect how our performance is assessed. It will also affect how many users visit our profile, find our other publications, and follow our project descriptions and then this may in turn affect the *real* impact of our work. As the current trend is for our work performance to be assessed by administrators, and even by peers, based on various impact statistics instead of by direct evaluation of our publications, we need to be aware of which factors unrelated to the quality of our work affect the data used in performance assessments.

I encourage every author and co-author of articles in our Bulletin to upload the reprints or even better make available links to their publications in our web site using DOI-based links as displayed in the Bulletin’s web site (<http://bulletin.uv4plants.org/>) through ResearchGate and similar sites. In addition, time and effort spent in “optimizing” publication titles and keywords for both human readers and search engines is very likely to pay back by making our work more visible in searches and improving ranking in performance evaluations.

Even less controversial measures of im-

pact like number of citations are not reliable measures of the value of scientific performance (reviewed by Bornmann 2016) and if used require very careful consideration of the context of the research being assessed. In my opinion impact statistics should be used only as a complement to actual expert evaluation of the scientific production of researchers and never by themselves in an automated manner. Based on the read statistics described above one could create the advertisement shown in Figure 5.4 relying on *actual hard numbers*; however, few readers would take it seriously. In contrast, a similar advert but with the name of a *big* journal, would be credible! The lesson is, that even *hard numbers* when taken out of context can drive us pitifully astray. Meanwhile, “blind” use of publication-impact- and similar indexes in performance assessments is too frequent to be ignored by those being evaluated. This suggests that well-guided *search engine optimization* or *SEO* of publications’ titles, keywords and contents could be a useful and morally acceptable tool for achieving success in academia.

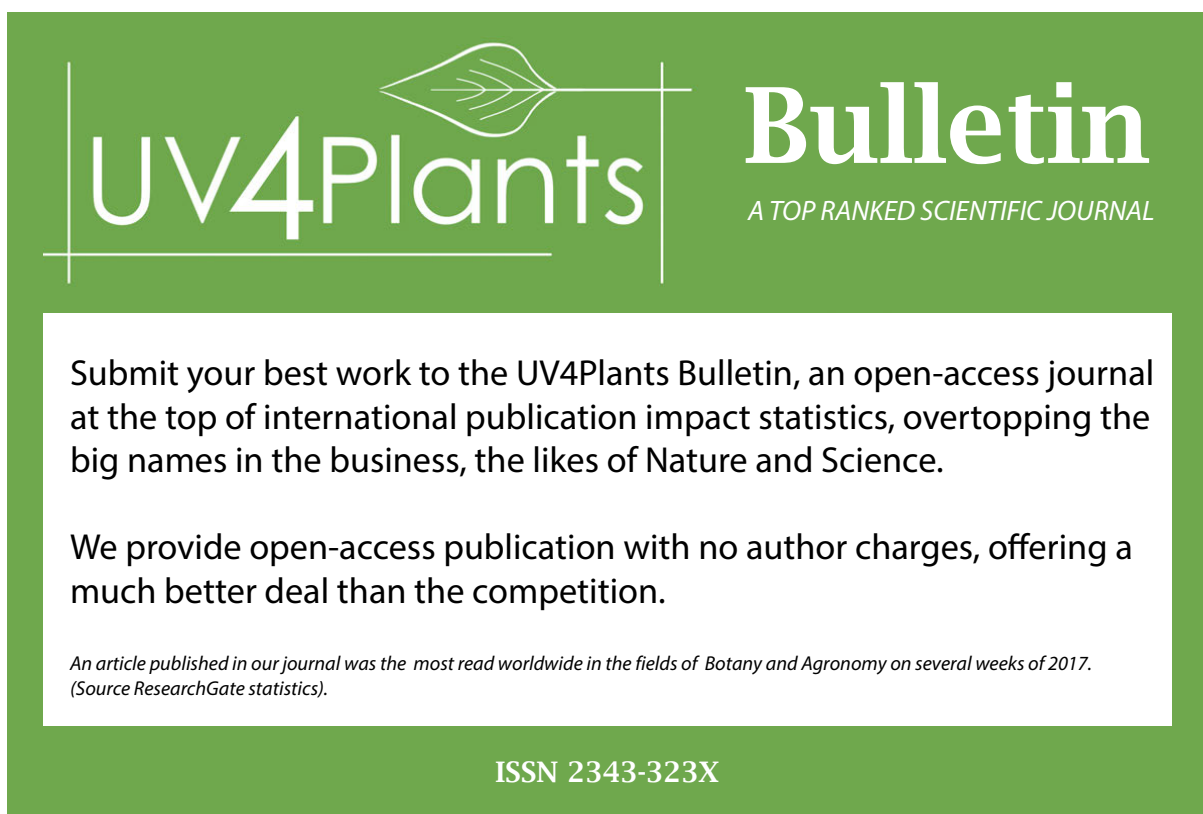
References

- Aphalo, P. J. (2016). “Thinking, Fast and Slow (by Daniel Kahneman)”. In: *UV4Plants Bulletin* 2016:1, pp. 43–44. ISSN: 2343-323X. DOI: [10.19232/uv4pb.2016.1.92](https://doi.org/10.19232/uv4pb.2016.1.92).
- Bornmann, L. (2016). “Measuring impact in research evaluations: a thorough discussion of methods for, effects of and problems with impact measurements”. In: *Higher Education* 73, pp. 775–787. DOI: [10.1007/s10734-016-9995-x](https://doi.org/10.1007/s10734-016-9995-x).

Editorial-board-reviewed article.

Published on-line on 2018-04-11.

Edited by: Titta K. Kotilainen.



UV4Plants

Bulletin
A TOP RANKED SCIENTIFIC JOURNAL

Submit your best work to the UV4Plants Bulletin, an open-access journal at the top of international publication impact statistics, overtopping the big names in the business, the likes of Nature and Science.

We provide open-access publication with no author charges, offering a much better deal than the competition.

*An article published in our journal was the most read worldwide in the fields of Botany and Agronomy on several weeks of 2017.
(Source ResearchGate statistics).*

ISSN 2343-323X

Figure 5.4: A fake advertisement for the Bulletin, which could have been real, had we not known the context of the readership statistics.