ISSN: 2632-4091 History *of* Classical Scholarship 19 October 2020
Issue 2 (2020): 181–191

How and Why I Count(ed)
A Response to Ramsay MacMullen

— Nathan Pilkington —

Abstract

In a recent publication in this journal, Professor Ramsay MacMullen failed to correctly represent thoughts I had posted online, though not yet published in a peer-reviewed journal. I thus establish here my positions on the value of citation scores, both how to calculate them and why one would want to do so.

Keywords

citation scores, Google Scholar, Web of Science, performance-based research funding systems, ancient historians

Due to a proliferation of performance-based funding systems over the last three decades, particularly in Europe, citation scores, in 2013, when I first began to address them, were becoming an increasingly important measure of scholarly productivity.[1] At that time, Professor Walter Scheidel had already started the process of ranking United States-based ancient historians via the Web of Knowledge.[2] The Web of Knowledge has since morphed into the Web of Science and remains the dominant method by which citation scores are calculated. In an initial reply to Scheidel, I argued that the Web of Science, due to its focus on English language journals, provided a restrictive measurement

[1] For general discussion, see: Aagard, K. et al. "Impacts of Performance-Based Research Funding Systems: The Case of the Norwegian Publication Indicator." *Research Evaluation* 24 (2015), 106–117; Sile, L. and R. Vanderstraeten. "Measuring Changes in Publication Patterns in a Context of Performance-Based Research Funding Systems: the Case of Educational Research in the University of Gothenburg." *Scientometrics* 118 (2019), 71–91.

[2] Scheidel, W. "Citation Scores for Ancient Historians in the United States." Version 1.0. Princeton-Stanford Working Papers in Classics. February 2008. Updated in September 2011 and posted as "Updated Citation Scores for Ancient Historians in the United States." http://www.princeton.edu/~pswpc/pdfs/scheidel/091102.pdf, last accessed 28.09.20. Scheidel has also applied the same method to Moses Finley's career. "Measuring Finley's Impact." http://www.princeton.edu/~pswpc/pdfs/scheidel/041302.pdf, last accessed 28.09.20. Scheidel has since published his study of Finley in D. Jew, R. Osborne, and M. Scott, eds., *Moses Finley. An Ancient Historian and his Impact* (Cambridge, 2016), 288–297.

of citation scores.[3] Instead, I proposed the use of Google Scholar in order to capture a wider spectrum of published works and languages other than English. More recently, Professor Ramsay MacMullen, in this journal, has argued that citation scoring, generally, is an inadequate measure of a scholar's impact and proposed the use of *L'Année philologique*, a method that measures the frequency of publication rather than citation scores.[4] In making his argument, Professor MacMullen failed to both correctly represent my argument for the use of Google Scholar as well as my views on citation scores and their value generally.

My interest in citation scores is twofold. First, scholars should understand how they are being counted. Second, scholars may want to understand how to adapt their publication strategy in order use citation scoring methods to their own advantage.[5] It is for these reasons that in my initial writing, as cited by MacMullen, I considered the development of scholarly careers over time, in order to demonstrate the relatively slow accretion of citations in the Humanities.[6] In a second writing, uncited by MacMullen, I considered the distribution of Classical Studies scholars throughout departments at several universities in the United States and at two distinguished English institutions (Cambridge and Oxford). I argued, at that time, that an integrated Classics department, in which all subfields are included, offered the best chance to demonstrate a high rate of citation, be it total or annualized.[7] In sum, I am not interested in rankings per se, rather the tools with which we are ranked and the ways those tools can, and in certain geographies already do, influence decisions about scholarly careers and departmental structures.[8]

[3] For discussion of the Web of Science and the limitations of its index, see: Kulczycki, E. et al. "Publication Patterns in the Social Sciences and Humanities: Evidence from Eight European Countries." *Scientometrics* 116 (2018), 463–486.

[4] MacMullen, R. "Top Scholars in Classical and Late Antiquity." *History of Classical Scholarship* 2 (2020), 105–114.

[5] For discussion of scholars' responses to performance-based research funding systems, see: Hammarfelt, B. and G. Haddow. "Conflicting Measures and Values: How Humanities Scholars in Australia and Sweden Use and React to Bibliometric Indicators." *Journal of the Association for Information Science and Technology* 69 (2018), 924–935.

[6] Pilkington, N. "Google Scholar and the Web of Knowledge: Citation Scores for Ancient Historians." https://www.academia.edu/3420110/Google_Scholar_and_the_Web_of_Knowledge_Citation_Scores_for_Ancient_Historians, last accessed 28.09.20.

[7] Pilkington, N. "Ancient Historians and Departmental Affiliations: The Value of Citation Scores?" https://www.academia.edu/3524452/Ancient_Historians_and_Departmental_Affiliations_The_Value_of_Citation_Scores_, last accessed 28.09.20.

[8] Hammarfelt and Haddow (2018).

How I Count(ed)

In a reply, posted in 2013, to Scheidel's initial ranking of US-based ancient historians, I argued that Google Scholar, as processed through the Publish or Perish software, offered a less restrictive measure of impact compared to the Web of Science.[9] In brief, at that time, the Web of Science did not index every journal in a field. The Arts and Humanities Index did not cover publications before 1975, inhibiting full access to many current scholars' careers. Finally, the database was heavily biased towards English language journal publications.

By contrast, at that time, Google Scholar indexed journals, books, dissertations, master's thesis, conference proceedings, and a whole host of other forms of scholarly communication. It provided a more comprehensive view of a scholar's penetration into the field, at the peer reviewed level of journal article and book, in addition to the humbler levels of graduate school work. It also included publications dating as far back as the 19th century. Finally, when searches were properly constructed using the Publish or Perish software, it was possible to access a scholar's citations in foreign language publications. Moreover, the software took account of translations of original editions, further demonstrating a scholar's degree of impact. To demonstrate the difference in method, I engaged with Scheidel's initial ranking of scholars. In 2013, the searches yielded what I felt was a substantial variance for certain scholars.

Table 1. Percentage Difference in Citation Scores (2013)

Scholar	Saller	Hall	Morris	Scheidel	Bagnall	Champlin	Matthews
% Difference GS/ WoS	76%	71%	70%	67%	53%	52%	27%

Over the past seven years, both methods of citation scoring have improved. Scheidel, in a more recently posted paper in 2019, has shown

[9] Harzing, A.W. (2007) Publish or Perish, available from: https://harzing.com/resources/publish-or-perish, last accessed 28.09.20. Harzing developed the software for the following reason, "The Social sciences, Arts and Humanities, and engineering in particular seem to benefit from Google Scholar's better coverage of (citations in) books, conference proceedings and a wider range of journals." I should note that I have often calculated my own annualized citation scores. Though the software is capable of doing this internally and provides this metric as part of its scoring system, reprints and new editions reset the date of publication for an individual work, somewhat skewing its citation per year count.

the utility of both in revising his rankings.[10] Scholars interested in the current state of both methods of counting should thus consult Scheidel's work. He concludes, "What matters is not the absolute number of citations but the relative ranking of scholars: in this regard, discrepancies between the two databases are fairly minor." A point to which I would now assent, though I continue to believe Google Scholar's broader coverage is superior for reasons discussed below.

Problematically, the capabilities of Google Scholar and thus my approach to citation scoring were incorrectly represented by MacMullen in this journal. He comments,

> What lies behind much of my criticisms even of Pilkington's choice of databases (better than Scheidel's choice pre-2019, as he concedes) which Pilkington found in "Google Scholar's citation Index processed through the Publish or Perish Software", is its deliberate limitations. Measurement of rank is sought "only in English language journals" (as later in Walter Scheidel 2019, 2, an "Anglo-only survey"). Yet no more than the 6% or so of the 980 periodicals pillaged by *Année philologique* are Anglophone (and additionally but also ignored by Scheidel, most European journals, such as *Historia* or *Epigraphica*, welcome English items along with other languages, beyond that of their own.[11]

I never sought publications in English language journals only, as noted above. Further, both *Historia* and *Epigraphica* are indexed by Google Scholar, and thus were accessed through the Publish or Perish software in my initial writings. Of all articles ever published in *Historia*, MacMullen's 1980 contribution "Women in Public in the Roman Empire" is the fifth most cited. Moreover, Google Scholar reveals citations of this particular article in English, German, Italian, French, Spanish, Turkish, Dutch and Polish language publications. For *Epigraphica*, the most cited article in Google Scholar's index is M. Burzachechi, "Oggetti parlanti nelle epigrafi greche" from 1962, again cited by publications in multiple languages.

In response to Professor MacMullen's attempt to import frequency of publication into the discussion via *L'Année philologique*, I would argue that the proliferation of a particular scholar's writing has unclear utility. Multiple studies have shown that undifferentiated metrics focused only on the frequency of publication lead to a proliferation of publications in

[10] Scheidel, W. "Citation Scores for Greco-Roman Historians in North America, 2019." https://www.academia.edu/40416928/Citation_Scores_for_Greco_Roman_Historians_in_North_America_2019, last accessed 28.09.20.

[11] MacMullen (2020), 110.

lower-tier journals.[12] Further, many scholars counted in both my and Scheidel's lists have published articles that produce no record of citation in either the Web of Science or Google Scholar. Finally, certain sub-fields of Classics offer more opportunities for regular publication, most notably epigraphy, papyrology and archaeology, as MacMullen notes. *L'Année philologique* suffers also from many of the same problems as the Web of Science. It indexes journals and edited collections. It thus represents a more restrictive level of publication, typically requiring peer review and editorial scrutiny, and captures less of a scholar's impact compared to Google Scholar.

Why I Count(ed)

MacMullen presented my definition of impact as follows, "By 'impact' I understand whatever shapes people's ideas, values, and behavior — one would hope, beneficially. It is apparently what Scheidel and Pilkington intend, applied to the particular population of ancient historians." I would demur from such lofty notions. I view impact as nothing more than a stand-in word for citation, be it positive or negative.

Further, I believe that the Classics need to argue for the widest definition of impact possible. The Web of Science is likely an effective measurement of scientific and social scientific output, where journal articles are the dominant form of communication.[13] Because reproducibility matters in scientific studies, the more a publication is cited by journal publications, the more likely that publication has a strong effect on the present state of its field. Studies that fail to produce a significant result are unlikely to be cited in further publications. Science is by definition additive. It is also quick.[14] Consequently, the Web of Science would seem a useful tool for departments and administrators when making tenure decisions in these fields.

By contrast, citation in the Classics is not constructed additively or quickly, but diachronically. Thoughts are presented, become orthodoxy,

[12] Aagard et al. (2015).

[13] Verleysen, F. and T. Engels. "How Arbitrary are the Weights Assigned to Books in Performance-Based Research Funding? An Empirical Assessment of the Weight and Size of Monographs in Flanders." *Aslib Journal of Information Management* 70 (2018), 660–672 for discussion.

[14] To give one example, a paper published on 30 April 2020 related to Coronavirus has amassed 6,901 citations on Google Scholar as of 4 August 2020: Wuan, G.-J. et al. "Clinical Characteristics of Coronavirus Disease 2019 in China." *New England Journal of Medicine* 382 (2020), 1708–1720. Citation score accessed via Publish or Perish on 5 August 2020.

then are challenged, and finally become part of a literature review once discarded or modified. In my initial writing on the subject, I studied scholars whose careers were less than 28 years from the date of first publication indexed by Google Scholar. I demonstrated that it takes time for citation scores to mature in ancient history. In sum, scholars with under 28 years of experience, in the main, fall within an order determined by years of activity.[15]

In the two tables below, I revisit the scholars with the fewest number of years active in my initial 2013 list in order to further demonstrate that citation scores accrete over time.

Table 2: 2013 Citation Scores

Author	Total Citation Score	Highest Publication Score	Years Active
S. Forsdyke	134	61	15
J. Ma	98	31	15
C. Kelly	96	57	17
C. Noreña	83	52	13
B. Holmes	45	16	9
W. Riess	40	15	14
G. Ruffini	38	14	10

Table 3: 2020 Citation Scores

Author	Total Citation Score	Highest Publication Score	Years Active
S. Forsdyke	960	358	22
J. Ma	1,808	587	22
C. Kelly	827	325	24
C. Noreña	737	227	20
B. Holmes	826	241	16
W. Riess	389	87	21
G. Ruffini	559	102	17

[15] Pilkington, “Google Scholar”.

Additionally, I argued that most scholars who achieve high citation scores tend to benefit, with reference to total number of citations, from a well-cited monograph, the first or second for the majority of scholars.

Table 4: Most Cited Work Overall versus Annualized Most Cited Work (2013)

Author	Top Ind Pub Score	Book # (Name)	Year of Career	Highest annual cite rate (2013)	Cites/year
Richard Saller	404	1 (Pers Pat)	2	Same	13.03
Ian Morris	368	1 (Bur Anc)	2	Why the West	27.25
Victor Hanson	365	2 (West Way)	9	Carnage Culture	17.75
Ronald Syme	1,317	1 (Rom Rev)	11	Same	17.79
Josiah Ober	742	2 (Mass Elite)	11	Same	29.68
Walter Scheidel	71	4 (Death Nile)	11	Same	5.46
Keith Hopkins	526	1 (Conq Sla)	15	Same	14.61
Fergus Millar	627	3 (Emp RW)	15	Roman Near East	18.33
Moses Finley	1,123	8 (Anc Econ)	22	Same	28.07
William Harris	832	3 (Anc Lit)	24	Same	34.6
Peter Brunt	775	1 (It Man)	24	Same	18.02
Roger Bagnall	342	17 (Egy Lat)	25	Same	19
Peter Brown	1,571	7 (Body Soc)	27	Same	60.42
A.H.M. Jones	1,201	8 (Lat Rom)	36	Same	42.89
Russell Meiggs	654	2 (Ath Emp)	42	Same	15.57

The same pattern would hold true for the scholars listed in Tables 2 and 3. Changes in citation scores over the past seven years are primarily driven by the first monograph. If these monographs were judged in 2013, when they were less than a decade old, the citation scores would have appeared paltry, as would the overall citation score for the scholar.

In sum, classical scholars benefit from time and monographic publications. If citation scores were to affect tenure decisions, classicists would need to be cognizant of the relatively slow rate of accumulation. Because most first books are published near to tenure review, citation scores, at least in the short term, have little predictive value with reference to the impact of a particular monograph over the long term.

An additional concern is the present state of Classics departments. In a second writing, posted also in 2013, I studied various departmental structures, ranging from the fully integrated Classics department (languages, literature, history, art, archaeology and philosophy) to the

Classics department with exclusively scholars of literature and language.[16] I wanted to understand how the inclusion or exclusion of ancient historians, archaeologists and other sub fields affected the citation scores of various departments. I focused on full professors only, due to that fact that these scholars had sufficient time for their citation scores to mature. In an initial ranking, I considered Classics departments as they existed at the time. I produced a set of rankings based on average number of citations and average annual citation rate.

Table 5: Comparative Ranking of Classics Departments (2013)

School	# of Full Profs	Total # of Citations	Avg. # of Citations	Avg. # of Years
Cambridge	16	28,598	1,787	38
Stanford	12	16,991	1,416	30
Oxford	23	19,736	858	29
Cornell	12	9,914	826	32
Harvard	11	7,911	719	28
Brown	9	6,449	717	30
UC-Berk	11	7,835	712	34
Chicago	8	5,016	627	31
Princeton	11	6,787	617	29
Yale	10	5,489	549	31
Penn State	7	3,724	532	33
UCLA	10	4,883	488	32
Michigan	14	6,682	477	32
NYU	9	4,172	464	33
Penn	7	3,009	430	28
Duke	6	2,388	398	30
Columbia	6	2,087	347	30
UNC	6	1,950	325	31

School	Avg. Annual Cit. Rates
Stanford	47.20
Cambridge	47.03
Oxford	29.59
Cornell	25.81
Harvard	25.68
Brown	23.90
Princeton	21.28
UC-Berk	20.94
Chicago	20.23
Yale	17.71
Penn State	16.12
Penn	15.36
UCLA	15.25
Michigan	14.91
NYU	14.06
Duke	13.27
Columbia	11.57
UNC	10.48

[16] Pilkington, N. "Ancient Historians and Departmental Affiliations: The Value of Citation Scores?" https://www.academia.edu/3524452/Ancient_Historians_and_Departmental_Affiliations_The_Value_of_Citation_Scores_, last accessed 28.09.20.

I then integrated ancient historians, archaeologists and any other scholars of the Classics housed outside of Classics departments into their Classics departments. To give an example, Columbia University Classics was integrated with scholars of antiquity housed at that time in the departments of History, Philosophy, and Art and Archaeology. Departments already integrated have the same citation score as in Table 5.

Table 6: Integrated Classics Departments (2013)

School	# of Full Profs	Total # of Citations	Avg. # of Citations	Avg. # of Years
Cambridge	16	28,598	1,787	38
Stanford	12	16,991	1,416	30
Oxford	23	19,736	858	29
Cornell	12	9,914	826	32
Harvard	12	8,734	728	29
Penn	9	6,470	719	32
Brown	9	6,449	717	30
UC-Berk	11	7,835	712	34
Chicago	12	7,688	639	29
Princeton	12	6,910	576	29
NYU	12	6,636	553	33
Yale	10	5,489	549	31
Penn State	7	3,724	532	33
UCLA	11	5,242	477	34
Columbia	12	5,601	467	28
Michigan	16	7,363	460	33
UNC	10	4,054	405	34
Duke	6	2,388	398	30

Integrated School	Rank	Actual School
Cambridge	1	Cambridge
Stanford	2	Stanford
Oxford	3	Oxford
Cornell	4	Cornell
Harvard	5	Harvard
Penn	6	Brown
Brown	7	UC-Berk
UC-Berk	8	Chicago
Chicago	9	Princeton
Princeton	10	Yale
NYU	11	Penn State
Yale	12	UCLA
Penn State	13	Michigan
UCLA	14	NYU
Columbia	15	Penn
Michigan	16	Duke
UNC	17	Columbia
Duke	18	UNC

School	Annualized Cit. Rates
Stanford	47.20
Cambridge	47.03
Oxford	29.59
Cornell	25.81
Harvard	25.10
Brown	23.90
Penn	22.47
Chicago	22.03
UC-Berk	20.94
Princeton	19.86
Yale	17.71
NYU	16.76
Columbia	16.68
Penn State	16.12
UCLA	14.03
Michigan	13.94
Duke	13.27
UNC	11.91

As the data indicated at the time, integration was important. Ancient historians and archaeologists have on average higher citation scores compared to other disciplines within Classical Studies. Such a finding suggests that Classics departments, especially if counted by citation scores, would benefit from additional scholars within the most cited fields.

Table 7: Average Number of Citations by Discipline (2013)

Field	# of Scholars	Total # of Citations	Avg. # of Citations
Archaeology	25	26,297	1,052
History	57	58,981	1,034
Philosophy	25	18,795	752
Literature	85	49,788	586
Art	11	3,145	286
Language	10	2,755	275

To conclude, I continue to believe that citation scores will increase in their importance over the next decade. As Classics departments are already under pressure, it behooves the field to make positive arguments about the value of the Classics. Citation scores can be part of that positive argument, but only if classicists understand fully how and why they are counted.

Nathan Pilkington
University of North Carolina Wilmington
pilkingtonn@uncw.edu