

**Editorial:**

A warning for Chinese academic evaluation systems: short-term bibliometric measures misjudge the value of pioneering contributions*

Xiao-jun HU^{†‡1}, Jian-hong LUO¹,
Ronald ROUSSEAU^{2,3}

¹Zhejiang University School of Medicine, Hangzhou 310058, China

²KU Leuven, B-3000 Leuven, Belgium

³University of Antwerp (UA), Faculty of Social Sciences, B-2020 Antwerp, Belgium

[†]E-mail: xjhu@zju.edu.cn

<https://doi.org/10.1631/jzus.B1700569>

Publication citation-based research evaluation, even if only in support of peer review, is not everywhere, on every level, or for everyone suitable, because of differences in scientific research, patterns of research output, stages of scientific evolution, and merits—scientific or societal—of scientific results.

In 1985 Garfield wrote: “By basing funding or even scholarly tenure and hiring decisions on quantitative bibliometric data, there is always the potential for making two serious mistakes: one, in believing that mere publication or citation counting is equivalent to citation analysis; and, two, in believing that citation analysis, even when carefully performed by experts, is sufficient by itself to ensure objectivity” (Garfield, 1985).

In 2010, Anthony van Raan, the director of the Centre for Science and Technology Studies (CWTS) at Leiden University in the Netherlands, said: “If there is one thing every bibliometrician agrees [on], it is that you should never use the journal impact factor

to evaluate research performance for an article or for an individual—that is a mortal sin” (van Noorden, 2010).

Although plagued by many controversies, the current practice of using short-term bibliometrics to gauge scientific achievements is hard to change. In 2016 annual panel meeting of Zhejiang University School of Medicine, Ping Yi, director of the office for R&D management, said: “Although some experts are aware that the current evaluation system for funding is imperfect, the impact factor and the number of received citations are still taken as two standards to determine if applicants for grants are qualified.” This statement clearly reflects the current situation in China and the large gap between policy-makers, scientists, and users in understanding the purpose of indicators.

A characteristic of pioneering work

Inspired by discussion about measures of research merit (McNutt, 2014; Selvarajoo, 2015), we conducted a series of investigations to analyze citation characteristics of pioneering work based on the large-scale citation network of millions of articles in the Web of Science (WoS). As a result we found that many truly foundational discoveries, e.g. key articles leading to Nobel Prizes, do not behave as flames in the landscape of scholarly communication. They rather act like sparks, keeping the power to ignite and hence transform their field. This type of pioneering work, not having received large numbers of direct citations, needs subsequent papers to realize its innate potential.

For example, restriction enzymes have proved to be invaluable for the physical mapping of DNA and

[‡] Corresponding author

* Project supported by the National Natural Science Foundation of China (No. 71573225)

 ORCID: Xiao-jun HU, <https://orcid.org/0000-0001-8384-0221>

© Zhejiang University and Springer-Verlag GmbH Germany, part of Springer Nature 2018

the biotechnology industry would certainly not have flourished without it. Surprisingly, a key article of Danna and Nathans (1971), setting the stage for much of what is now routine and which led him to jointly receive the Nobel Prize in Physics in the year 1978 (<https://www.nobelprize.org>), received only 299 citations in total (all data shown in this contribution were updated on 18 August, 2017). However, this article acted as a “wake-up switch”, bringing about substantial highly-cited follow-up research (Fig. 1). The 15 most-cited papers citing Danna and Nathans (1971), each received a higher number of citations than their source of inspiration. Moreover, many items citing the citing articles (second generation citations), received even more citations and this in a shorter time span. In the fourth column of Fig. 1, we see eight papers citing the citing article C1 (published in 1977 and with 1531 citations), each having more than 1535 citations: the most-cited one among these eight even received 4555 citations.

Another typical example is Yoshinori Ohsumi, the 2016 Nobel Prize winner in Physiology or Medicine, who discovered the autophagy mechanism in

the early 1990’s. Among his key publications reported by RSAS (2016a), nine pioneering studies were published during the period 1992–1999. We found that all these articles have the same “sparkling” characteristics as the article of Danna and Nathans (1971). For instance, an article by Ohsumi’s research team (Kametaka et al., 1996), received only 76 citations. Yet, its 34 most-cited citing articles (sharing 45% of all received citations) each received more than 94 citations. Fig. 2 illustrates the citations of its 15 most-cited citing articles (all received more than 300 citations) and their significant follow-up papers. In the fourth column of Fig. 2, we point to another aspect reflecting the excellent performance of subsequent research citing the citing articles, namely the number of highly-cited papers (the 1% most-cited ones of their field and their publication year) included in the papers citing the citing articles. As shown in Fig. 2, most top citing articles of Ohsumi’s research (Kametaka et al., 1996) have a substantial number of highly-cited papers in follow-up research, this in stark contrast with the number of direct citations received by the original.

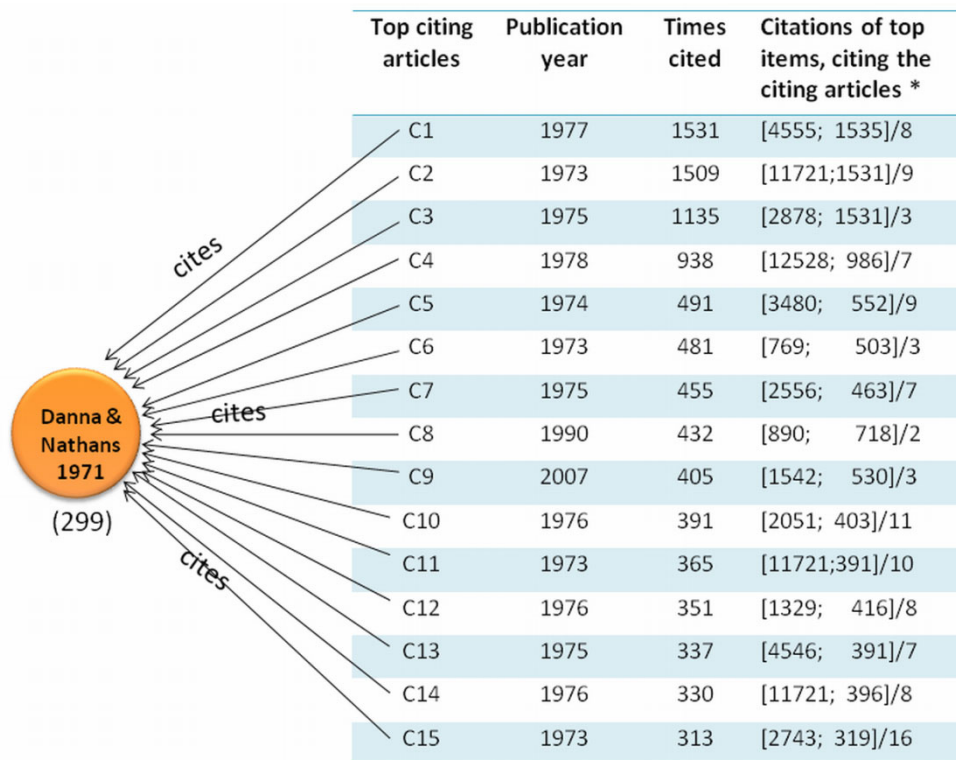


Fig. 1 Citations to an article by Danna and Nathans (1971) and its highly-cited follow-up research

The fourth column shows the data set of subsequent work for each citing article; all items received more citations than the corresponding citing articles. * This top “citations of citations” set is denoted as following: [the largest number of citations; the smallest number of citations]/the number of items in the set

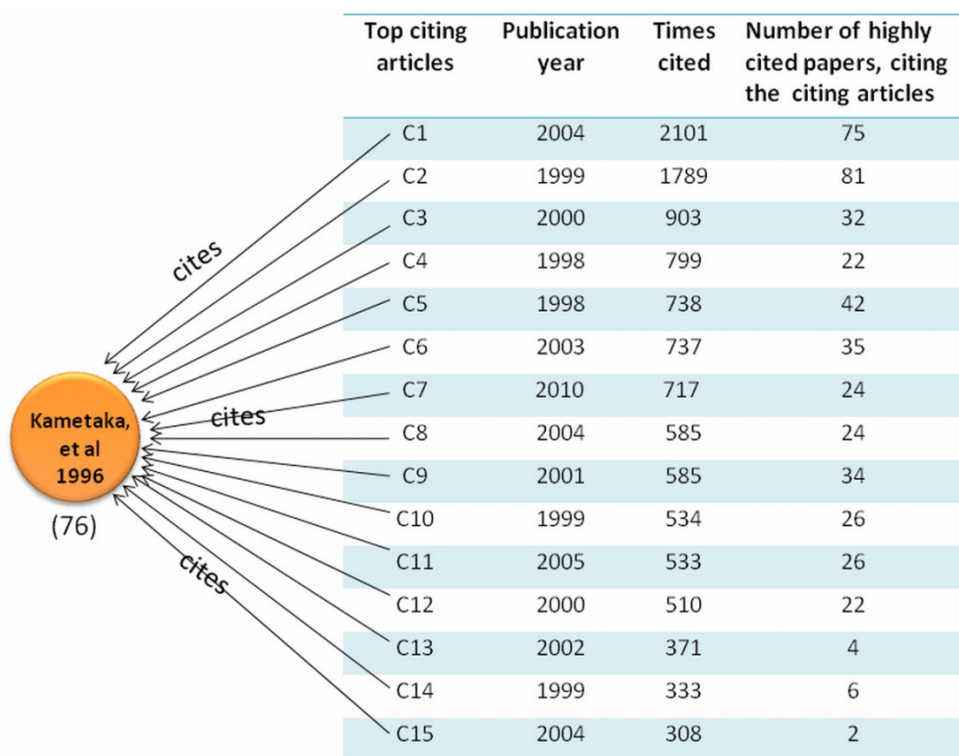


Fig. 2 Citations to an article by Kametaka et al. (1996) and its highly-cited follow-up research

The fourth column shows the number of “highly-cited papers” following up each citing article; the “highly-cited papers” belong to the 1% most cited ones of their field and publication year

The examples shown here are not exceptional cases. Indeed, in previous investigations (Hu and Rousseau, 2016, 2017) we found that most fundamental work of Nobel Prize winners, such as the papers of Jean-Pierre Sauvage, J. Fraser Stoddart, and Ben L. Feringa (Nobel Prize winners in Chemistry in the year 2016) have the same “sparking” characteristics. Concretely, 23 of 29 publications reported by RSAS (2016b) have been shown to be influential in this way (Hu and Rousseau, 2017). Ruska’s fundamental work (Ruska, 1933) on electron optics entitled “The electron-electron microscopic image of irradiated Surfaces” (a key article that let him to receive a Nobel Prize in Physics in 1986; <https://www.nobelprize.org>), behaved similarly. Other examples include: a key paper of William A. Fowler (Marion and Fowler, 1957) that led him to receive the 1983 Nobel Prize in Physics (<https://www.nobelprize.org>) and You-you Tu’s article (Tu, 1999) that let her to receive the 2015 Nobel Prize in Physiology or Medicine (<https://www.nobelprize.org>). These three articles only received respectively 56, 51, and 25 citations, yet all brought about a significant amount of follow-up research (Hu and Rousseau, 2016).

Implications for research assessment

The evidence provided by our investigations suggests that it often happens that the value of pioneering work, even fundamental contributions leading to Nobel Prizes, is not immediately recognized and needs subsequent research to realize its innate potential. Obviously, if we only count direct citations from a short-term viewpoint, the value of this type of fundamental contributions may be severely misjudged. We recall that Figs. 1 and 2 illustrate this point clearly.

Users should become metric-wise

To remedy a flawed bibliometric-based assessment for scientists and improve our understanding of what constitutes good selection criteria, the key point to be emphasized is to make colleagues aware of the “invisible mortal sin” when they misuse short-term bibliometric indicators (van Noorden, 2010).

In an interview, Xiao-gang Peng, chemistry professor of Zhejiang University, listed eighth in

the Top 100 Chemists, 2000–2010 (<http://archive.sciencewatch.com/dr/sci/misc/Top100Chemists2000-10/>), says “Single quantitative indicators cannot reflect the true value of scientific research”, suggesting that quantitative indicators have limits and should not be used on their own. This is the true meaning of the word “indicator”: it just gives an indication but it most surely is not the absolute truth. Users must understand the limitations and proper use of bibliometric indicators and become metric-wise (Rousseau and Rousseau, 2015). For instance, scientists should realize that a quantitative indicator is a useful and convenient tool to make comparisons at the institutional or country level when a large amount of data is present, but is meaningless for measuring research merit and originality of thinking at the individual level, especially in the case of pioneering research.

Reducing the use of short-term bibliometric indicators

“Short-term insight and university rankings boost a tendency towards boasting and exaggeration”, says Shu-min Duan, dean of Zhejiang University School of Medicine, academician of the Chinese Academy of Sciences, “the essence of science is to trace the source of a scientific issue; most fundamental discoveries are not the result of a first attempt and their potential is often not seen in the short term. Yet, real significant work will bring about social progress sooner or later, perhaps even after several decades.” His comments are consistent with our findings that the value of many fundamental works of Nobel Prize winners would be misjudged when using a citation window of 5 years or less. Given the fact that citers are peers, we add that it is highly improbable that peer review (typically a small group of peers) would detect the value of this type of work.

Placing more emphasis on the “source of originality”

Pioneering work, particularly transformative research, is the fundamental source of originality in basic research. As such this type of investigation

plays an essential role in scientific progress. However, their results often do not fit within established models or theories and may initially be unexpected or difficult to interpret; their transformative nature and utility may not be recognized until years later. “The current evaluation system is not suitable for encouraging or supporting transformative research”, says Wei Yang, director of the National Natural Science Foundation of China, “it is high time to modify the review system to boost potential transformative research” (Yang et al., 2017). His insightful comments imply that besides short-term evaluations it is also important to perform longer-term evaluations, and this will be in a multi-dimensional and, perhaps, unconventional way. Creating more “objective” assessments is a hard task for the scientific community: in our opinion it needs intensive cooperation among different fields and a confluence of insightful perspectives coming from diverse stake holders.

Compliance with ethics guidelines

Xiao-jun HU, Jian-hong LUO, and Ronald ROUSSEAU declare that they have no conflict of interest.

This article does not contain any studies with human or animal subjects performed by any of the authors.

References

- Danna K, Nathans D, 1971. Specific cleavage of simian virus 40 DNA by restriction endonuclease of *Hemophilus influenzae*. *Proc Natl Acad Sci USA*, 68(12):2913-2917. <https://doi.org/10.1073/pnas.68.12.2913>
- Garfield E, 1985. Uses and misuses of citation frequency. *Current Contents*, 43:3-9.
- Hu XJ, Rousseau R, 2016. Scientific influence is not always visible: the phenomenon of under-cited influential publications. *J Informetr*, 10(4):1079-1091. <https://doi.org/10.1016/j.joi.2016.10.002>
- Hu XJ, Rousseau R, 2017. Nobel Prize Winners 2016: igniting or sparking foundational publications? *Scientometrics*, 110(1):1053-1063. <https://doi.org/10.1007/s11192-016-2205-x>
- Kametaka S, Matsuura A, Wada Y, et al., 1996. Structural and functional analyses of *APG5* a gene involved in autophagy in yeast. *Gene*, 178(1-2):139-143. [https://doi.org/10.1016/0378-1119\(96\)00354-X](https://doi.org/10.1016/0378-1119(96)00354-X)
- Marion JB, Fowler WA, 1957. Nuclear reactions with the neon isotopes in stars. *Astrophys J*, 125(1):221-232. <https://doi.org/10.1086/146296>
- McNutt M, 2014. The measure of research merit. *Science*, 346(6214):1115. <https://doi.org/10.1126/science.aaa3796>
- Rousseau S, Rousseau R, 2015. Metric-wiseness. *J Assoc*

Inform Sci Technol, 66(11):2389.

<https://doi.org/10.1002/asi.23558>

- RSAS (the Royal Swedish Academy of Sciences), 2016a. Scientific background discoveries of mechanisms for autophagy. http://www.nobelprize.org/nobel_prizes/medicine/laureates/2016/advanced-medicineprize2016.pdf [accessed on Oct. 5, 2016]
- RSAS (the Royal Swedish Academy of Sciences), 2016b. Scientific background on the Nobel Prize in Chemistry 2016—Molecular Machines. http://www.nobelprize.org/nobel_prizes/chemistry/laureates/2016/advanced-chemistryprize2016.pdf [accessed on Oct. 5, 2016]
- Ruska E, 1933. The electron-electron microscopic image of irradiated Surfaces. *Zeitschrift für Physik*, 83(7-8):492-497 (in German).
<https://doi.org/10.1007/BF01338960>
- Selvarajoo K, 2015. Measuring merit: take the risk. *Science*, 347(6218):139-140.
<https://doi.org/10.1126/science.347.6218.139-c>
- Tu YY, 1999. The development of new antimalarial drugs: qinghaosu and dihydro-qinghaosu. *Chin Med J*, 112(11): 976-977.
- van Noorden R, 2010. A profusion of measures. *Nature*, 465:864-866.
<https://doi.org/10.1038/465864a>
- Yang W, Zheng YH, Dong C, 2017. How to review a potential transformative research? *Bull Natl Nat Sci Found China*, 31(4):1-8 (in Chinese).

中文概要

题目:对中国学术评价体系的警示——短期文献计量指标误判科学先驱性工作价值

概要:虽然用短期文献计量指标评估科学家成就的现行做法在学术界受到许多争议和质疑,但是这一“根深蒂固”的观念和评估方法在现实中一直难以改变,主要原因是长期以来在政策制定者、科学家和使用者之间存在着很大裂缝。为纠正这种有明显缺陷的“文献计量”评估方法,我们对一些诺贝尔奖获得者的主要论文,进行了基于大数据的引用特征系列研究。通过对观察节点在科学引文索引数据库(Web of Science)引用网络中百万篇论文节点间的深度关系分析,我们发现:许多真正的奠基性工作,并没有像耀眼的火焰在科学星空中闪烁出灿烂光芒;相反,它们更像是导火索上的星星火花,点燃大量后续研究火焰继而引发了科学界的转型。结合资深专家访谈,我们呼吁科学同盟者们应该意识到,“误用短期文献计量指标”等同于“道义上犯罪”,评价体系应该更强调“创新源头”,助推潜在的颠覆性与转型性研究。

关键词:学术评价;科学先驱性工作;诺贝尔奖;文献计量指标;短期评估