Journal of Informetrics 10 (2016) 1079–1091

Contents lists available at ScienceDirect

Journal of Informetrics

journal homepage: www.elsevier.com/locate/joi

# Scientific influence is not always visible: The phenomenon of under-cited influential publications

Xiaojun Hu [a], Ronald Rousseau [b,c,*]

[a] School of Medicine, Zhejiang University, Hangzhou 310058, China
[b] Department of Mathematics, KU Leuven, Leuven B-3000, Belgium
[c] Faculty of Social Sciences, University of Antwerp (UA), Antwerp B-2000, Belgium

ARTICLE INFO

*Article history:*
Received 8 July 2016
Received in revised form 1 October 2016
Accepted 1 October 2016

*Keywords:*
Under-cited publications
Nobel prize winners
Follow-up research
Promoters of influential research

ABSTRACT

In this submission we introduce the notion of under-cited influential publications and show that these publications are like "wake-up switches" for significant follow-up research. To be considered an under-cited influential article we require an article to meet three requirements. One is on the level of received number of citations (first generation citations), while the two other ones take subsequent citation generations into account. In general terms these three conditions are:

1) The article is reasonably well-cited (a basic requirement to be influential)
2) Citations of citations (second generation citations) are rather highly cited, so that the original one is influential in an indirect way (a more refined token of influence);
3) Given condition two, the article received fewer citations than expected (being under-cited).

We claim that the phenomenon of under-cited influential publications is important and should receive more attention. Moreover, one may say that under-cited influential publications belong to the group of truly foundational scientific discoveries acting as promoters of influential research as shown by significant follow-up research.



## 1. Introduction

The publication of the first Science Citation Index (SCI), more than fifty years ago, can be considered as the result of a systematic effort to track citations over major scientific publication outlets (Garfield, 1963). This rich source of data describing the citation network of scientific information sparked the interest in citation analysis leading to thousands of articles. A recent review with emphasis on citing behavior can be found in (Bornmann & Daniel, 2008). The Science Citation Index and its companion indices have been used in numerous research management and research evaluation exercises (Koenig, 1983; Moed, Burger, Frankfort, & van Raan, 1985; Moed, de Bruin, & van Leeuwen, 1995; Zunde, 1971). Soon, however, it became clear that misguided use of citation counting led to many discussions and controversies (DORA, 2012; Seglen, 1997).

The Chinese scientist Youyou Tu received the 2015 Nobel Prize in Physiology or Medicine. Her case is a clear example that in science it is more important to make truly foundational discoveries rather than receive thousands of citations. Searching for Tu's publications in the Web of Science (WoS) one finds only four articles (of the 43 she wrote) and these are not heavily

* Corresponding author at: Department of Mathematics, KU Leuven, Leuven B-3000, Belgium.
*E-mail addresses:* xjhu@zju.edu.cn (X. Hu), ronald.rousseau@kuleuven.be, ronald.rousseau@uantwerpen.be (R. Rousseau).

http://dx.doi.org/10.1016/j.joi.2016.10.002
1751-1577/

cited. This phenomenon suggests that "success is not always what one sees", and indeed, many publications can be said to be under-cited by the current research evaluation system.

Most research leads to small, but necessary and useful contributions to scientific progress. Yet, every now and then a truly foundational contribution sees the light, such as Ruska's (1933) fundamental work in electron optics, or the discovery of restriction enzymes and their application to problems of molecular genetics (Arber & Linn, 1969; Danna & Nathans, 1971; Smith & Wilcox, 1970). Although the authors of these discoveries received the Nobel Prize and their original publications led to many significant follow-up research, they are not mentioned in *Nature*'s list of 100 most cited research of all time (based on Thomson Reuters' Web of Science) (Van Noorden, Maher, & Nuzzo, 2014). The most-cited ever article in the Web of Science is (Lowry, Rosebrough, Farr, & Randall, 1951), which received more than 305,000 citations. In the first ever *Citation Classic* (Lowry, 1977) wrote *"Nevertheless, although I really know it is not a great paper (I am much better pleased with a lot of others from our lab), I secretly get a kick out of the response. . .."*. Yet, Lowry's article surely was a useful one. These examples illustrate that performing top science and publishing articles that are extremely highly cited are different things, a point also made when discussing the most-cited articles of the 21st century (Sanz-Casado, García-Zorita, & Rousseau, 2016).

In this contribution, we focus on the type of phenomenon as illustrated by the work of Youyou Tu. Some publications are really influential but do not receive as many citations as one expects. This is not a blank statement, but one we will illustrate and try to quantify. A definition will be provided for such influential articles. Recently, informetricians have come to realize that in citation research not only direct (forward) citations or references (backward citations) of a record should be considered, but also its indirect citations and references (Atallah & Rodríguez, 2006; Fragkiadaki & Evangelidis, 2014; Hu, Rousseau, & Chen, 2011; Hu, Rousseau, & Chen, 2012; Kosmulski, 2010; Rousseau, 1987). Taking these observations into account we will use several generations of citations into account when defining the notion of an under-cited influential publication.

To be considered an under-cited influential article we require an article to meet three requirements. One is on the level of received number of citations (first generation citations), while the two other ones take subsequent citation generations into account. In general terms these three conditions are:

1) The article is reasonably well-cited (a basic requirement to be influential);
2) Citations of citations (second generation citations) are rather highly cited, so that the original one is influential in an indirect way (a more refined token of influence);
3) Given condition two, the article received fewer citations than expected (being under-cited).

These three general requirements will be operationalized further on.

In the next section we have a closer look on one of Tu's articles, introducing a concrete example of what one may call, at least in an intuitive way, an under-cited influential article. In the next section we operationalize the three requirements for being an under-cited influential article. We do this in two ways, one more suited for high profile fields, and one more suited for the 'average' field. Then we come to some case studies of fundamental work ahead of transformative research, and general articles in applied science. Finally, we discuss our approach and compare it with the scientific "gems" as found by (Chen, Xie, Maslov, & Redner, 2007) and present our conclusions.

## 2. A scientific contribution is not always what it looks as seen from a citation perspective

Although Youyou Tu began her work on the development of antimalarial drugs in the early nineteen seventies, the earliest publication of Youyou Tu included in the WoS is one published in the year 1982, entitled "Studies on the constituents of artesimia annua, Part II (Tu et al., 1982), while her second article in the WoS came much later. It is entitled 'The development of new antimalarial drugs: Qinghaosu and dihydro-qinghaosu', and was published in 1999 in the *Chinese Medical Journal* (Tu, 1999). These are key articles that led to a Nobel Prize. Yet, the first article received (up to March 31, 2016) 64 citations and the second one only 19. The first article has an h-index of 31, while the second one has only an h-index of 11. Recall that the h-index of an article A is defined as the largest natural number such that the h most cited articles citing article A received at least h citations (Schubert, 2009).

However, we observe an unusual phenomenon related to this second article: many of the citing articles have a better citation performance, and this, as they are younger, in a shorter citation window than Tu's article, referred to as article T, see Table 1. This table shows all articles in the h-core of article T. We observe the unusual phenomenon that most of them are cited more than article T.

There is nothing uncommon when an article, say article B, is cited by another one, that later turns out to become more highly cited than article B. Yet, when this is not just an exception but B is cited by a significant number of articles that each are more highly cited than B, then one may consider B to be a foundational article, that, for whatever reason is under-cited. It may even be under-recognized, although recognition and highly-cited are not the same, as illustrated again by Tu's work.