

STI 2018 Leiden

*23rd International Conference on Science and Technology Indicators
"Science, Technology and Innovation Indicators in Transition"*

STI 2018 Conference Proceedings

Proceedings of the 23rd International Conference on Science and Technology Indicators

All papers published in this conference proceedings have been peer reviewed through a peer review process administered by the proceedings Editors. Reviews were conducted by expert referees to the professional and scientific standards expected of a conference proceedings.

Chair of the Conference

Paul Wouters

Scientific Editors

Rodrigo Costas
Thomas Franssen
Alfredo Yegros-Yegros

Layout

Andrea Reyes Elizondo
Suze van der Luijt-Jansen

The articles of this collection can be accessed at <https://hdl.handle.net/1887/64521>

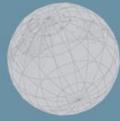
ISBN: 978-90-9031204-0

© of the text: the authors

© 2018 Centre for Science and Technology Studies (CWTS), Leiden University, The Netherlands



This ARTICLE is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License



STI 2018 Leiden

*23rd International Conference on Science and Technology Indicators
"Science, Technology and Innovation Indicators in Transition"*

Notice: preliminary findings

As authors of this abstract (Why (almost) Everything We Know About Citations is Wrong: Evidence from Authors), we wanted to give you a preliminary look at what we were currently exploring during the STI conference. However, we considered this work to be in a preliminary, exploratory state -not even a work in progress. Therefore, it was decided to keep the attached abstract out of the conference proceedings. Unfortunately, it was published by oversight in the frenzy of conference organization.

Since the conference proceedings were available online, many people have already read the abstract, and it has attracted some press. This is unfortunate but understandable. When, in the near future, a working paper on these preliminary results becomes available, the link will be made available, on [this page](#).

Why (almost) Everything We Know About Citations is Wrong: Evidence from Authors ¹

Misha Teplitskiy*, Eamon Duede**, Michael Menietti*, and Karim Lakhani*

* mteplitskiy@fas.harvard.edu; mmenietti@hbs.edu; klakhani@hbs.edu

Laboratory for Innovation Science, Harvard University, 1737 Cambridge St., Cambridge, 02138 (USA)

** eduede@uchicago.edu

Knowledge Lab, University of Chicago, 5801 S Ellis Ave., Chicago, 60637 (USA)

Notice: preliminary findings (see previous page).

Introduction

Analysis of citations has become increasingly common due to its many, seemingly straightforward applications. For instance, modern scientific research output has grown exponentially over the past several centuries and has resulted in something like an information overload for researchers and the institutions that support them. Moreover, intensifying specialization has made it more difficult for both evaluators to judge the promise and progress of research investments (funding), and for frontline researchers to select from the literature the prior knowledge upon which to build the next 'big idea'. All of these changes have suggested the importance of and given rise to automated approaches to recommendation, evaluation, and prediction that take into account the deluge of information that otherwise no individual researcher, program officer, dean, provost, or congressperson can master. A number of approaches to measure value in research and to recommend relevant or important literature in a given discipline or field have been created. These tools are now ubiquitous, used routinely by practitioners to search the literature, validate claims, or seek inspiration. Furthermore, institutions leverage similar tools and metrics such as the 'h-index' to make hiring and promotion decisions, and funding agencies use them to understand the 'impact' of their funding as well as to evaluate the promise of this or that researcher's proposal relative to the 'measured' value of his or her prior contributions. Yet, all such evaluations make heavy use of citations as proxies. While citations are widely used to evaluate research and allocate resources, the referencing decisions on which they are based are poorly understood.

The tremendous pace of scientific publishing outpaces individuals' abilities to thoroughly digest and evaluate each published work. Consequently, scientists, administrators, and policy makers often lean on quantitative metrics like citations to *value* scientific works. The more citations, the more quality, the more influence. Citations and metrics derived from them, like the *h-index*, are ubiquitous and routinely used to search the literature, validate claims, promote or hire individuals, allocate grant funding, and so on.

In this paper, we first distinguish two perspectives of citing decisions, the *normative* and the *social constructivist*. The normative view holds that scientists and scholars cite works that

¹ This work was supported by BIG Ideas Generator, University of Chicago

influenced their research choices, and that they consider to be of high quality. In contrast, the social constructivist view holds that individuals cite papers for rhetorical and strategic reasons that are independent of the individuals' *personal* perceptions of the works' quality. For example, under the social constructivist view, scientists and scholars will cite works that they do not know well and that did not influence their research choices, but that support claims they want to make and are familiar to the intended audience. Consequently, whatever the citation counts signal, they do not signal authors' judgments of the quality or the influence of the work.

Data and Methods.

To assess evidence for each of these views and rigorously determine precisely what can be inferred from citation counts, we fielded a web-based, intelligent, pilot survey² of scientists across 6 fields of science and humanities, in which we asked about specific references they made in their papers. While others have attempted to survey researchers about citation practices, none have attempted to survey broadly across disciplines and with systematic sampling of *cited* papers from the entire published literature. We rely on the unique blend of computational techniques with rich data from the complete Clarivate Analytics *Web of Science*, which enables our survey instrument to scale arbitrarily.

We sampled researchers using the following sampling frame: First, we selected one field from each of Web of Science's 6 major categories – the fields were Endocrinology, Ecology, Management, Analytical Chemistry, Religion, and Computer Science - Information Systems. Second, for each field we identified all publications published in 2010 and ranked them according to how many citations they accrued by 2015. Third, for each field, we randomly selected a paper from each percentile of the field's citation distribution and asked up to ten individuals who cited the paper in 2015 to evaluate its 'quality', 'validity', 'novelty' and other attributes, along with how much the paper influenced their research choices and how well they know their contents.

Additionally, we experimentally manipulated the information respondents observed when evaluating papers: the treatment group was shown how much the paper had been cited ("status signal") while the control group was shown no information regarding the paper's citations ("no status signal"). At the beginning of the survey, when respondents are shown the title and abstract of the target paper, half of the respondents (chosen at random) are given a 'Social Signal' of the target paper's status with regard to citation accumulation. Specifically, participants in the treatment group are shown a simple sentence which states the following: 'Our records indicate that this paper (the target paper) has been cited X times which ranks it in the top (or bottom) Y% among all papers published in the field in 2010.' Here, X and Y are the actual number of recorded citations and the actual citation percentile rank of the paper, respectively. Respondents in the treatment group only receive (i.e. see) the 'social signal' one time. Moreover, no attempt is made to highlight or draw attention to the treatment. It simply appears at the end of the abstract, immediately before the question asking whether the respondent remembers the target paper and never reappears at any future point in the survey.

Results.

We present two sets of findings, which combine data responses from all 6 sampled fields. First, authors know the content of the papers they cite less well when the references are highly

² We are currently collecting data from the full survey (~70,000 solicitations) and will be able to present it at the Conference.

cited (Figure 1). Specifically, after asking how well the participants know the content of the target papers that they cited in 2015, we find that at least 40% of all citations across all percentiles are known only 'Slightly Well' or 'Not Well' at all. Moreover, for the top quintile of cited papers (80th percentile and up), the fraction of responses that state that the authors who cited the target paper do not know the content of the target paper well (e.g., are only familiar with the main findings of the paper) jumps to over 20%. In fact, for all citation quintiles, the top quintile garners the highest proportion of respondents admitting that they are not well informed about the content of the papers they cited. That is, the most cited papers also have the highest proportion of authors who cite them claiming to not know more about the content of that paper than its main finding with an additional (roughly) 20% claiming to know them only slightly well (e.g., only familiar with the findings, data, and methods). The total fraction of respondents claiming to know the content of the most highly cited papers 'Slightly Well' or less is slightly higher than the fraction of responses stating the same for the most sparsely cited papers in my study. This observation is hard to square with the Normative View of citation practice.



Figure 1. Across citation quintiles, at least 40% of cited papers are 'Not Well' known or only 'Slightly Well' known by those who cite them. Furthermore, authors report knowing famous papers they cite less well than obscure ones.

Moreover, authors are influenced (*per capita*) equally by highly and lowly cited works (Figure 2).

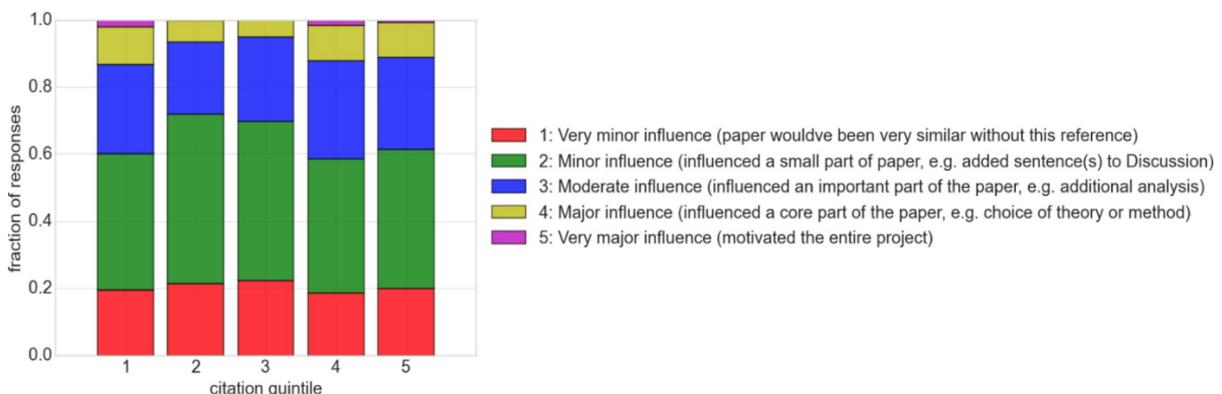


Figure 2. Across citation quintiles, including the very highest, at least 60% (sometimes as high as 70%) of citations are said to have been of either 'Minor Influence' or 'Very Minor Influence'.

Over 60% of respondents indicate that the papers they cited had only "minor" or "very minor influence" on their research choices. Moreover, this overt reporting of a widespread lack of

influence holds across every citation quintile including, again, the highest (80th percentile and up). In fact, the fraction of respondents claiming that the target paper that they cited had 'Very Minor Influence' (e.g., that the source paper would have been very similar without the reference) is the same (roughly 20%) in both the bottom quintile (the most sparsely cited papers) and top quintile (most highly cited papers)

Furthermore, without an explicit signal of a paper's status in the citation distribution (control condition), respondents perceive the quality, influence, validity, novelty and significance of highly and lowly cited papers to be equal, on average. With an explicit status signal (treatment condition), a positive correlation appears between a paper's citation count and its citers' perceptions of 'quality', 'influence', 'significance' and other attributes of the papers (Figures 3 and 4).

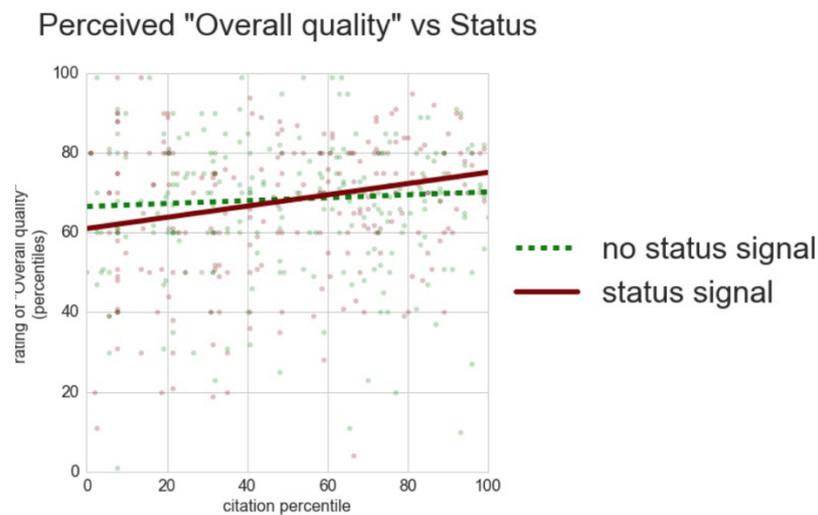


Figure 3. OLS regression shows a statistically significant relationship ($p < 0.1$) between 'Perceived Quality' and 'Citation Percentile' for the *treatment* group (explicit information on how often the paper has been cited) but not the *control* group (no status signal).

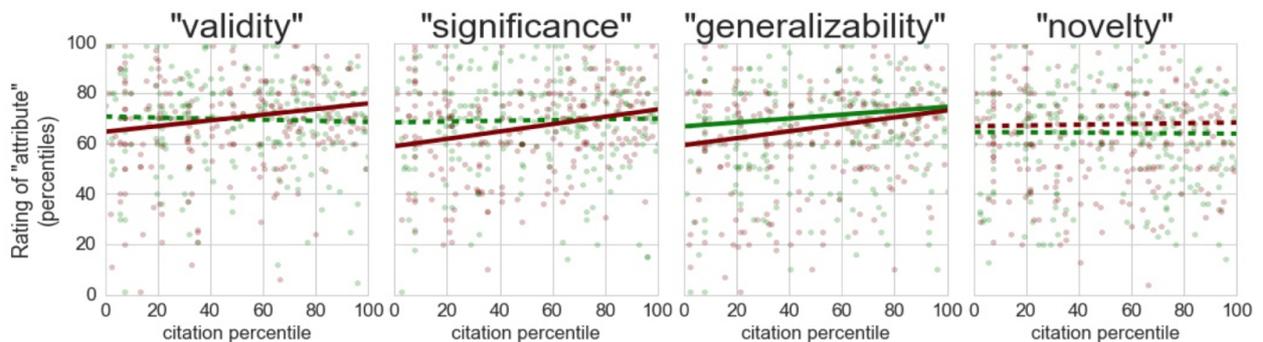


Figure 4. OLS regressions for the *control* group show no statistically significant relationship between a paper's 'Citation Percentile' and its perceived 'Validity', 'Significance', or 'Novelty'. However, one regression does show a slight, positive relationship ($p < 0.1$), that between 'Generalizability' and 'Citation Percentile'. In contrast, the *treatment* group shows positive, statistically significant relationships between 'Citation Percentile' and all attributes, with the exception of perceived 'Novelty'.

Positive correlations between citations and perceptions of the quality of a paper, like its validity or significance, are thus explained entirely by status signals. Nevertheless, scientists do rate the works they cite as being above a certain threshold of quality.

Conclusion.

We argue that the evidence is most consistent with a “citation decision function” that combines normative and social constructivist elements. Authors do not cite works they perceive to be below a minimum threshold value of quality, supporting the normative view. However, above this threshold, frequency of use is unrelated to quality. Instead, usage is determined by social constructivist elements: scientists tend to cite works they are not influenced by and that they do not know particularly well. Although normative considerations play a role, the threshold-nature of the role makes it invalid to infer differences in perceived quality between highly and lowly cited items. In sum, our findings elucidate what drives citation decisions, severely undermine the normative view of citation practices, and require a radical reassessment of the role of citations in evaluative contexts.