



A measure for the impact of research

Alejandro M. Aragón

School of Architecture, Civil and Environmental Engineering (ENAC), École polytechnique fédérale de Lausanne (EPFL), CH-1015 Lausanne, Switzerland.

SUBJECT AREAS:

STATISTICS

COMPUTATIONAL SCIENCE

APPLIED MATHEMATICS

SCIENTIFIC DATA

Received

11 March 2013

Accepted

27 March 2013

Published

11 April 2013

Correspondence and requests for materials should be addressed to

A.M.A.

(alejandro.aragon@epfl.ch);

(alejandro.aragon@fulbrightmail.org)

The last few years have seen the proliferation of measures that quantify the scientific output of researchers. Yet, most of these measures focus on productivity, thus fostering the “publish or perish” paradigm. This article proposes a measure that aims at quantifying the impact of research de-emphasizing productivity, thus providing scientists an alternative, conceivably fairer, evaluation of their work. The measure builds from a published manuscript, the literature’s most basic building block. The *impact* of an article is defined as the number of lead authors that have been influenced by it. Thus, the measure aims at quantifying the manuscript’s reach, putting emphasis on *scientists* rather than on raw citations. The measure is then extrapolated to researchers and institutions.

The exponentially increasing number of publications¹, makes it increasingly hard for researchers to keep up with the literature. The problem of *paper inflation* presents newcomers the even more challenging exercise of finding those works that have made significant contributions, knowledge that researchers accumulate over years of experience. Part of the problem, noted almost half a century ago², stems from a structure of science that favors *productivity*. The current system puts pressure on researchers for the publication of scientific articles, for this is their only way to justify the received funding. The publication of results, states Margolis², “is probably the main means of accomplishing the almost impossible task of accounting for time and money spent on research. Inevitably, this puts a premium on quantity at the expense of quality, and, as with any other type of inflation, the problem worsens: the more papers are written, the less they account for and the greater is the pressure to publish more.” Thus, the *publish or perish* paradigm that besets scientists has, for the most part, inevitable implications on the quality of research published. At worst, many authors opt for publishing the same data, or even for the *minimum publishable unit* (MPU), in order to increase the apparent value of their work³. In the latter, research findings are divided into a series of articles instead of producing a single meaningful manuscript. This *modus operandi* has further implications, for the slight contribution to science of an MPU presents editors of journals the very subjective task of discerning if it even deserves publication. The problem of paper inflation may eventually place a tremendous burden on the entire scientific community in order to keep the infrastructure for the publication and dissemination of scientific works.

The problem has only been boosted by the appearance of recent quantitative measures that favor this productivity ideology. “A scientist”, states Hirsch⁴, “has index h if h of his or her N_p papers have at least h citations each and the other $(N_p - h)$ papers have $\leq h$ citations each”. The h index has been widely adopted since its introduction in 2005 by major online databases⁵, as it improves over former measures for the quantification of a researcher’s performance. Its simplistic computation is perhaps the responsible for its wide adoption, and Hirsch suggests that the index could even be used by decision makers for the granting of funding, tenure, promotion, and awards^{4,6}. While researchers are concerned about important decisions being made solely based on such measures⁷, the application of the index has even been extended to the productivity assessment of entire research groups, institutions and countries⁸. It has also been shown that the index is able to predict future scientific achievement⁹. On the downside, the measure carries with it many drawbacks as it appears from the many variants that emerged in the literature. One major criticism to the h index is its inability to compare scientists from different fields, for the citation practices vary significantly per discipline. Radicchi et al.¹⁰ have proposed a variant of the h index that takes this variability into account for a fairer comparison across fields, to the expense of a more involved calculation. In a recent publication¹¹, Bornmann shows that most of other 37 variants of the h index add no significant information. Another problem with the h index that has been observed, is that contrary to what its proponents may claim, it is not hard to manipulate its value in one’s favor^{12,13}. Lehmann et al.¹⁴ also show that, while the h index “attempts to strike a balance between productivity and quality”, the mean number of citations per paper is a better measure of research quality.

Measures that are based on citations will carry over their intrinsic problems^{15,16}. Seglen states that “citations represent a measure of utility rather than of quality”, and gives a succinct summary of their main problems¹⁷. In



spite of current efforts that aim at measuring the outcome of scientific research not solely based on publications and citations¹⁸, the latter remain for now the only means to quantify the impact of a scientific article. According to Martin¹⁹, “the impact of a publication describes its *actual* influence on surrounding research activities at a given time. While this will depend partly on its importance, it may also be affected by such factors as the location of the author, and the prestige, language, and availability, of the publishing journal.” In a subsequent publication²⁰, he states that “citation counts are an indicator more of impact than of quality or importance”. Thus, citations in this article are implicitly assumed to measure impact, as to avoid criticism over their relation to both research importance and quality²¹.

This article provides an alternative measure of scientific achievement based on citations, but their negative effects are reduced to a minimum, thus yielding a measure that cannot be easily manipulated. The proposition does not aim at replacing current measures of productivity, but to complement them in order to provide the research community with an alternative evaluation of its scientific production. To measure the quality of scientific output, Bornmann and Daniel claim⁸, “it would therefore be sufficient to use just two indices: one that measures productivity and one that measures impact.” With the suggested measure it may be possible for a scientist to have considerable impact even if publishing a single article, contrary to what the *h* index and similar measures would suggest. As mentioned earlier, the measure is defined first for a scientific manuscript in an attempt to quantify its reach, and hence scientists are at the very core of its definition. Similarly, the impact of a scientist aims at determining the number of researchers that have been influenced by his or her work. The measure is further extrapolated so that the impact of an institution comprises the impact of its body of researchers. The impact measures for manuscripts, scientists, and institutions, whose formal definitions are given in the Methods section, are referred henceforth as Φ , ι , and I , respectively.

Results

For the computation of scientist impact, citation records are taken from the SciVerse Scopus database. Considering that this database lists records since 1996, impact values for scientists who have published earlier represent partial estimates. In addition, the number of citing documents are also obtained for some authors from the Web of Science citation index provided by Thomson Reuters.

The strengths of the proposed impact measure are evident when applied to eminent scientists who have not published considerably, but who have nevertheless produced substantial contributions to science. For instance, Kary B. Mullis won the Nobel price in chemistry in 1993 due to his contributions within the field of DNA-chemistry research. From a total of 8063 citing documents obtained from the Scopus database over a period of 18 years, it is found that 6435 of those records (roughly 80%) have a different first author who do not share authorship with Mullis. Consequently, with a modest *h* index of 15, Mullis has had an impact $\iota = 6435$ over that period of time. Yet, as mentioned before, this is a partial estimate and the true impact is expected to be much higher. The Web of Science lists a total of 24750 citing documents for this author, which is approximately 3 times the number of records used to compute ι . To put the impact of Mullis into perspective, seven authors taken from those 8063 citing documents were found to have the same *h* index. Yet, their impact values were {346, 404, 553, 555, 561, 680, 1284}, where the highest value is scarcely 20% of that of Mullis. Another example is the late Richard P. Feynman ($h = 37$), who won the Nobel price in physics in 1965 for his work on quantum electrodynamics. With a total of 8123 citing documents obtained from the Scopus database, an impact $\iota = 5175$ is assigned to Feynman. In this case only about 64% of the citing documents counted towards his impact. As before, the Web of Science lists

a total of 19629 citing documents, so the impact value above is a partial estimate.

Figures 1 and 2 show cumulative impact results, in semi-log plots, for distinguished scholars in physics and mathematics, respectively. Nobel laureates since 2007 are displayed in Figure 1 (except those from 2009 due to lack of data in the database), where impact values were obtained with data up to the year of their corresponding price as a means to only survey their research impact. The figure shows a clear trend of impact increase over the years, with the sudden increase in the first portion of the curves suggesting the lack of data prior to 1996. It is remarkable the curves of professors Andre Geim and Konstantin Novoselov, whose work on the graphene material has had an enormous impact over the last few years. The impact of Fields medalists does not seem to be largely affected by the receipt of the price, and thus their curves in Figure 2 are computed from data up to 2012. From this figure it is also noticeable the impact of Prof. Terence Tao, who has an order of magnitude difference with respect to his contemporaneous mathematicians.

A department of the University of Illinois at Urbana-Champaign, the author's *alma mater*, is considered now for the survey of an institution's impact. More precisely, the Aerospace Engineering Department is examined. With a total of 23 actively working professors, the department's impact is $I = 7592$. This value takes into account the *impact overlap* that exists among the institution's scholars, i.e., the fact that they can have an impact on the same researchers. For the department the impact overlap is $\Omega \approx 18\%$. As long as there is a non-zero value of impact overlap between the scientists of an institution, the impact value will be smaller than the sum of their impacts, which in this case amounts to $\sum_i \iota_i = 9439$. The department has 6 assistant, 3 associate, and 14 full professors, and thus the same analysis can be made regarding their hierarchical positions within the institution. The impact of assistant, associate, and full professors are $I_a = 398$, $I_A = 443$, and $I_F = 7018$. Their corresponding impact overlap values are $\Omega_a \approx 1.3\%$, $\Omega_A \approx 0.7\%$, and $\Omega_F \approx 17\%$. For the formal definitions of the quantities used, the reader is referred to the Methods section.

Discussion

Instead of the total number of citations, which has been traditionally used as a measure of the impact of an article⁵, the proposed measure Φ aims at discerning the genuine number of *people* the paper has had an impact upon. In other words, Φ aims at measuring the manuscript's reach. Implicit to the definition of impact is the most basic assumption of citation analysis, that “references cited by an author are a roughly valid indicator of influence on his work”, in the words of McRoberts¹⁶. Yet, citations are used in a way that attempts the mitigation of any deviations from this assumption.

The impact of a manuscript Φ is defined in a way that excludes self-citations of any kind. Self-citations, states Schreiber¹², “do not reflect the impact of the publication and therefore ideally the self-citations should not be included in any measure which attempts to estimate the visibility or impact of a scientist's research.” Since only first authors are taken into account, Φ establishes a lower bound on the actual number of scientists that are influenced by the article.

One could argue that an article's impact should take into consideration not only first authors but also the entire authorship of citing papers, but in that case the measure could easily be over-bloated. There are fields where an article's authorship contains even thousands of names, so their inclusion would be not only superfluous but also meaningless. More often than not, first authors write the papers, and are responsible to give credit to those who have influenced their work.

The computation of impact values for manuscripts may help cope with the aforementioned problem of distinguishing those that have influenced the most people in their corresponding field. The impact of a manuscript has a monotonically increasing nature, but its

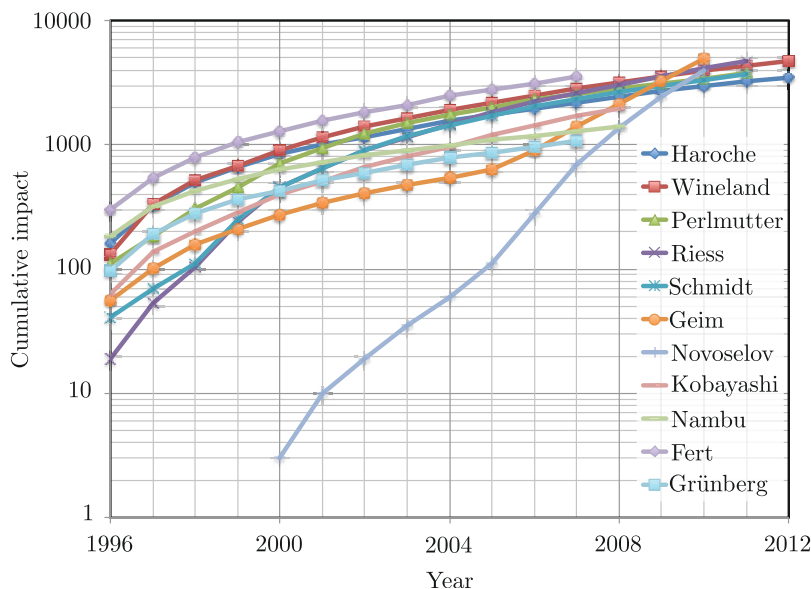


Figure 1 | Cumulative impact values for Nobel laureates in physics since 2007.

computation over a time period could give an indication of the diffusion of the article in its field. Additionally, it may be expected that the impact follows the same rate of growth as that of the scientific literature, which has been shown to grow exponentially^{22,1}. Thus, non-contemporaneous manuscripts of the same relative importance may have very different values of Φ after t years. Still, this does not void its definition, for the newer article would still reach more people. Also, even though it is expected that the impact value is mostly obtained from positive citations, the measure carries no information about discrediting citations. And yet, even negative citations have influenced their authors, and thus this lack of information is immaterial to the definition of impact.

The impact for a manuscript bears no information about the reputation of the journal of its publication. Some may argue that the manuscript impact should be multiplied by a factor that takes this fact into consideration in order to separate the wheat from the chaff. Yet, in the view of the author this is not necessary, and it may even be harmful. The proposition would not only become elitist, but also

there is no current way to determine a fair value for a weight of a journal. Impact factors and similar measures should by no means be used as weights, as they are not representative of all manuscripts in a journal. It has been established that usually a very small number of publications are the responsible for the high impact factors of some journals, whereas the majority of manuscripts receive little citations^{17,23}. Conversely, the prestige of the journal where the article is published is already implicit in the very definition of impact, insofar as publishing in a reputable journal may provide higher visibility. As a result, it may be expected that an article published in a prestigious journal would reach more people, and would therefore have a higher impact value.

Regarding the impact of a scientist i , one immediately remarks that it is not only obtained from articles where the scholar is the lead author. The proposed asymmetry stems from the way scientific articles are written. Even though it is the first author who gets the most credit (except in fields that list authors alphabetically), the writing of a scientific paper is not a solo enterprise, and it usually comprises

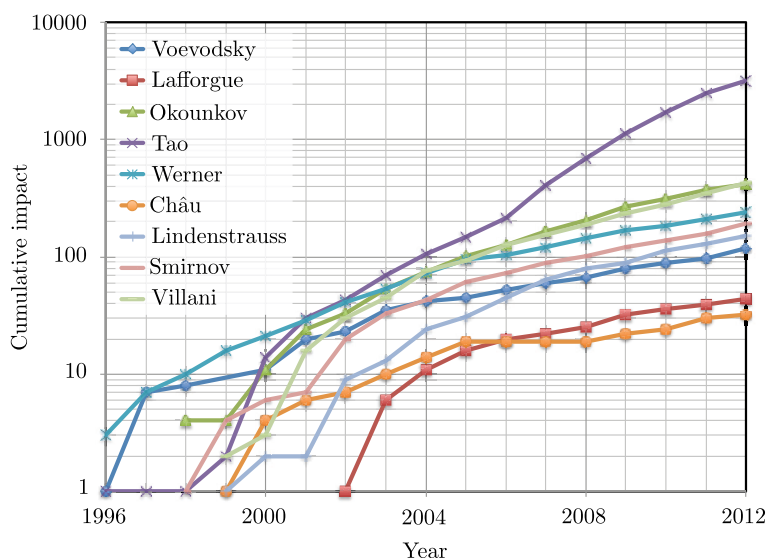


Figure 2 | Cumulative impact values for Fields medalists since 2006.



numerous contributions among those in the authorship. Even in the case that it is the first author who writes the manuscript, and who carries most (if not all) of the work needed to obtain its results, it is not uncommon that the main idea that originates the paper in the first place comes from someone else in the authorship (usually the student's advisor or manager of a project). Since it is impossible to discern credit from the authorship, the impact must account equally for all authors. Broad³ comments on a study where the scientists' judgements about their contributions in the research team summed up to a total of 300%. Besides, failing to give equal credit to all authors could prompt the researchers of higher-rank to take over first authorship, an unfortunate situation for the student. Defined this way, the impact can also be applied to those fields that use the Hardy-Littlewood rule to list authors. In spite of this, impact values in these fields are expected to be lower than if authors were not listed in alphabetical order.

The proposed measure ι establishes a lower bound on the *direct* impact of a scientist. However, if scientist B cites an idea from scientist A , the latter gets no credit by a scientist C who cites B for the idea. This is still congruent with the goals of establishing a lower bound, as scientist C never had to be exposed to the work of scientist A . Still, the measure could be combined with, e.g., network models² or with the modern PageRank citation ranking algorithm²⁴ to provide further insight on both direct and indirect impact. But the additional information may come at the expense of losing the impact's direct interpretation, and its computation would no longer be simple.

Another remark that can be made about a scientist's impact ι is that addition is not used to include citing first authors already taken into account in previous works, and with valid argument. In the view of the author, when addition is used as a means to quantify impact, the given measure can be abused as it promotes quantity over quality. Therefore, addition should only be used when measuring productivity, which is not the objective of the proposed measure. Although it could be argued that the impact ι is unfair to scientists who inspire the same scientific community with new ideas, there are several advantages that more than compensate for this drawback.

First, the direct interpretation of the impact ι , for it gives a realistic quantity on the number of people influenced by the research of a scientist. Second, an immediate consequence of the first point is that when the proposed measure is applied to eminent scientists, the impact values can give a rough indication of the sizes of their fields of research. Third, its computation is straightforward as it only requires information about the authorship of citing papers. Fourth and most important, the measure promotes *quality over quantity*. Authors gain nothing by either dividing their work, or by publishing the same data over and over, if they are to be cited by the same group of researchers. On the contrary, it is likely that a single meaningful work would have more visibility than that of an equivalent series of smaller manuscripts. Longer articles, states Laband²⁵, "receive more citations than shorter ones" (see also Bornmann and Daniel²², and references therein). Fifth, the impact encourages *innovation*. If researchers do not increase their impact with consecutive works, it means that either their articles are not being cited, or that they are cited by the same scientists. On the contrary, a continuous growth of ι may suggest that more and more scientists get engaged in their work. In order to increase their impact, scientists may look into exploring other areas within or even outside their field of expertise, encouraging multidisciplinary research and collaborations. Sixth, the impact may promote the evaluation of researchers not solely based on productivity. The impact is introduced not to compete with the h index or its variants, but to complement them. Seventh, the impact cannot be manipulated easily towards one's advantage, as its very definition excludes self-citations of any kind. Furthermore, networks of researchers that cross-cite their works in order to increase their citation count^{15,21}, practice known as cronyism, would have little

effect on ι . If nothing else, these networks would promote the appearance of new first authors. Also, it has been revealed that there are scientists who are eager to review articles whose citations would push towards the increase of their h index⁷. Needless to say that the *citation machinery*, referred to the addition of citations for "calling the attention or gaining the favor of editors, referees or colleagues"^{26,22}, cannot be avoided (see also¹²). But this is so for any type of measure based on citations. Finally, at worst the impact measure may have major implications on researchers that try to forge a career out of non-ethical work. For example, the plagiarism of articles and its publication in little known journals would have little to no effect on the impact of a researcher²⁷.

With respect to the impact of an institution I , departments would look into hiring faculty that would explore a different area to those already present. This is already implicit in hiring committees, as it makes no sense to recruit faculty whose work would overlap completely with existing professors, and who would therefore reach the same research community. By means of Eqn. (1), however, research institutions have a way of quantifying the added impact of prospective faculty. The institutions can also measure their impact overlap, given by Eqn. (2). The closer this value to zero, the more independent the fields of study among the institution's scientists. Yet, it is supposed that some degree of overlap is not only desirable, but also inevitable, for its direct implication on the collaborations among the scientists within the institution.

In this day and age, all kinds of measures abound and will live on whether we like it or not. In spite of that, it is only fair to quantify scientific output not solely with measures that favor productivity. If the research of a scientist has a true impact, it should be feasible to measure it, even if that research is contained in a single publication. The suggested methodology to measure impact goes against the *publish or perish* dogma of modern science, putting forward an alternative ideology to quantify scientific achievement. If taken seriously, and scientists do take measures seriously⁷, the proposed measure may help cope with some of the problems caused by the modern structure of science.

Methods

Impact of a scientific manuscript. In the following definitions, bold and non-bold greek letters are used to represent sets, and elements of sets, respectively. The proposed measure builds from the impact of a single scientific manuscript μ , whose author set is denoted $\alpha(\mu)$. This set may contain a single author, in which case the set is a singleton. Let $\Pi = \{\pi_1, \pi_2, \dots, \pi_m\}$ be a set of m citing articles that do not have an author in $\alpha(\mu)$, expressed mathematically by $\{\bigcup_{j=1}^m \alpha(\pi_j)\} \cap \alpha(\mu) = \emptyset$. In other words, none of the authors of articles in Π can be found to be an author of μ . Furthermore, let $\varphi_i \equiv \varphi(\pi_i) \in \alpha(\pi_i)$ represent the *lead* or *first author* of the i th citing paper. Thus the set of all lead author scientists that cite μ is given by $\Phi = \bigcup_{i=1}^m \varphi(\pi_i)$. Bear in mind that sets do not contain duplicates. The impact of a manuscript is defined as $\Phi \equiv |\Phi|$, i.e., the cardinality of set Φ . Note that Φ belongs to the set of natural numbers, i.e., $\Phi \in \mathbb{N}_0$.

Impact of a scientist. Let $\iota = \bigcup_{j=1}^m \Phi_j$ denote the combined lead author set over the n manuscripts written by the scientist, where Φ_j refers to the lead author set (as defined above) for the j th article. The impact of a scientist ι is then defined as the cardinality of this set, i.e., $\iota \equiv |\iota| \in \mathbb{N}_0$. In words, the impact of a scientist comprises the number of all *first* authors who have cited any of the n published works regardless of the scientist's authorship position. Note that the impact of a scientist can be determined by using the inclusion-exclusion principle, and that addition is nowhere used. The scientist's average impact per article is ι/n .

Impact of an institution. By extension, the impact of an institution I is determined by the impact of its body of scientists. Given an institution with s scientists, its impact is thus defined using the inclusion-exclusion principle

$$I = \left| \bigcup_{k=1}^s \iota_k \right|, \quad (1)$$

where ι_k denotes the combined lead author set of scientist k . Again, no summation is involved in the definition. The institution's average impact per researcher is then given by I/s . An institution can measure its *impact overlap* as



$$\Omega = \frac{1}{I} \left| \bigcup_{i \neq j} I_i \cap I_j \right| \quad (2)$$

Algorithm. For the numerical calculations of research impact, it is assumed that there are no homographs, i.e., different authors sharing the same written name. This assumption may result in lower impact values, for those first author scientists who write their name exactly the same way are counted as one. Yet, some efforts are underway for the creation of unique researcher identifiers that could easily remove this obstacle^{18,28}. In spite of this, the determination of scientific impact is straightforward. Authors who write their names slightly different are detected using the Levenshtein distance, an algorithm that finds dissimilarities between two sequences of characters. When it is recognized that two authors are spelled slightly different, name initials are then compared. If up to this point no assurance about the uniqueness of the author can be made, one of the authors is marked as *dubious*. For the results presented in Figures 1 and 2, dubious authors are but a small percentage of the total number: 1.9% – 3.1% for physicists and 0% – 2.6% for mathematicians.

- Larsen, P. O. & von Ins, M. The rate of growth in scientific publication and the decline in coverage provided by Science Citation index. *Scientometrics* **84**, 575–603 (2010).
- Margolis, J. Citation indexing and evaluation of scientific papers. *Science* **155**, 1213–1219 (1967).
- Broad, W. J. The publishing game: getting more for less. *Science* **211**, 1137–1139 (1981).
- Hirsch, J. E. An index to quantify an individual's scientific research output. *Proceedings of the National Academy of Sciences* **102**, 16569–16572 (2005).
- Noorden, R. V. Metrics: A profusion of measures. *Nature* **465**, 864–866 (2010).
- Hirsch, J. E. An index to quantify an individual's scientific research output that takes into account the effect of multiple coauthorship. *Scientometrics* **85**, 741–754 (2010).
- Abbott, A. *et al.* Metrics: Do metrics matter? *Nature* **465**, 860–862 (2010).
- Bornmann, L. & Daniel, H.-D. The state of *h* index research. Is the *h* index the ideal way to measure research performance? *EMBO Reports* **10**, 2–6 (2009).
- Hirsch, J. E. Does the *h* index have predictive power? *Proceedings of the National Academy of Sciences* **104**, 19193–19198 (2007).
- Radicchi, F., Fortunato, S. & Castellano, C. Universality of citation distributions: Toward an objective measure of scientific impact. *Proceedings of the National Academy of Sciences* **105**, 17268–17272 (2008).
- Bornmann, L., Mutz, R., Hug, S. E. & Daniel, H.-D. A multilevel meta-analysis of studies reporting correlations between the *h* index and 37 different *h* index variants. *Journal of Informetrics* **5**, 346–359 (2011).
- Schreiber, M. A case study of the Hirsch index for 26 non-prominent physicists. *Annalen der Physik* **16**, 640–652 (2007).
- Bartneck, C. & Kokkelsmans, S. Detecting *h*-index manipulation through self-citation analysis. *Scientometrics* **87**, 85–98 (2011).
- Lehmann, S., Jackson, A. D. & Lautrup, B. E. Measures for measures. *Nature* **444**, 1003–1004 (2006).
- Thorne, F. C. The citation index: Another case of spurious validity. *Journal of Clinical Psychology* **33**, 1157–1161 (1977).
- MacRoberts, M. H. & MacRoberts, B. R. Problems of citation analysis: A critical review. *Journal of the American Society for Information Science* **40**, 342–349 (1989).
- Seglen, P. O. The skewness of science. *Journal of the American Society for Information Science* **43**, 628–638 (1992).
- Lane, J. Let's make science metrics more scientific. *Nature* **464**, 488–489 (2010).
- Martin, B. R. & Irvine, J. Assessing basic research: Some partial indicators of scientific progress in radio astronomy. *Research Policy* **12**, 61–90 (1983).
- Martin, B. R. The use of multiple indicators in the assessment of basic research. *Scientometrics* **36**, 343–362 (1996).
- Phelan, T. J. A compendium of issues for citation analysis. *Scientometrics* **45**, 117–136 (1999).
- Bornmann, L. & Daniel, H.-D. What do citation counts measure? A review of studies on citing behavior. *Journal of Documentation* **64**, 45–80 (2008).
- Campbell, P. Escape from the impact factor. *Ethics in Science and Environmental Politics* **8**, 5–7 (2008).
- Page, L., Brin, S., Motwani, R. & Winograd, T. The PageRank citation ranking: Bringing order to the web. Technical Report 1999–66, Stanford InfoLab (1999).
- Laband, D. N. Is there value-added from the review process in economics?: Preliminary evidence from authors. *The Quarterly Journal of Economics* **105**, 341–352 (1990).
- Vinkler, P. A quasi-quantitative citation model. *Scientometrics* **12**, 47–72 (1987).
- Broad, W. J. Would-be academician pirates papers. *Science* **208**, 1438–1440 (1980).
- Editorial. Credit where credit is due. *Nature* **462**, 825 (2009).

Acknowledgements

The author would like to thank Dr. Vladislav Yastrebov and Prof. Jean-François Molinari for their suggestions on the article.

Additional information

Competing financial interests: The authors declare no competing financial interests.

License: This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivs 3.0 Unported License. To view a copy of this license, visit <http://creativecommons.org/licenses/by-nc-nd/3.0/>

How to cite this article: Aragón, A.M. A measure for the impact of research. *Sci. Rep.* **3**, 1649; DOI:10.1038/srep01649 (2013).