

“In numbers we trust?”

Kritische Betrachtung bibliometrischer Indikatoren
aus der Sicht eines Wissenschaftlers

Reinhard König
Leibniz-Institut für Neurobiologie
Magdeburg

20. Jahrestagung des Arbeitskreises Bibliotheken und Informations-
einrichtungen der Leibniz-Gemeinschaft
Magdeburg, 14./15. November 2019



LIVING SCIENCE

In numbers we trust?

Scientists go to great lengths to ensure that data are collected and analysed properly, so why, asks Eve Marder, do they apply different standards to data about the number of times research papers have been cited and viewed?

Marder. eLife 2014;3:e02791. DOI: 10.7554/eLife.02791

Eve Marder is an *eLife* senior editor, and is in the Department of Biology and the Volen National Center for Complex Systems, Brandeis University, Waltham, United States

Contribution to the Theme Section 'The use and misuse of bibliometric indices in evaluating scholarly performance'



Lost in publication: how measurement harms science

Peter A. Lawrence*

Department of Zoology, University of Cambridge, Downing Street, Cambridge CB2 3EJ, UK, and MRC Laboratory of Molecular Biology, Hills Road, Cambridge CB2 0QH, UK

“Measurement of scientific productivity is difficult. The measures used ... are crude. But these measures are now so universally adopted that they determine most things that matter [to scholars]: tenure or unemployment, a postdoctoral grant or none, success or failure.

As a result, scientists have been forced to downgrade their primary aim from making discoveries to publishing as many papers as possible—and trying to work them into high impact-factor journals. Consequently, scientific behaviour has become distorted and the utility, quality, and objectivity of articles have deteriorated. Changes to the way scientists are assessed are urgently needed.”

Nauka

Nauka w Polsce: Czas wrócić do rozsądku!

Andrzej Kajetan Wróblewski, fizyk, były rektor UW 17 września 2012 | 08:31

“I get the impression that when moving in the crazy circle of indicators, percentages, impacts and rankings, almost everyone has already forgotten what the purpose of science is. Therefore, it is worth recalling that **the purpose of science is, first of all, to search for the truth about the surrounding world, in other words, to discover new, previously unknown things ...**

However, it is not the purpose of science to publish articles, to bid for impact points, Hirsch, Egghe's or other indicators, nor is it the purpose of obtaining grants.”

Übersicht

1. Über die Motivation von Wissenschaftler*innen
2. Publikationen
3. Bewertung von Forschung / Wissenschaftler*innen:
Impact-Faktor und Hirsch-Index
4. Gibt es einen Ausweg?

1. Über die Motivation von Wissenschaftler*innen?

Was ist deren Motivation?

- Neugierde (die Welt zu verstehen, die Welt zu verändern)
- Kreativität
- Autonomie (um eigenen Ideen zu folgen)
- Anerkennung durch Kollegen (um in Erinnerung zu bleiben?)
-

Das Gehalt ist NICHT die primäre Motivation.

“Creativity in science, as in arts, cannot be organized. It arises spontaneously from individual talent. ... Discoveries cannot be planned, they pop-up like Puck, in unexpected places.”

Max Perutz, Preface to *I wish I'd Made you Angry Earlier: Essays on Science, Scientists, and Humanity*
Oxford University Press, 2002

Wie wird ‚Forschung‘ gemessen / bewertet?

- Produktivität (Anzahl an Publikationen)
 - ‚Impact‘ (Anzahl der Zitationen, Impact-Faktor)
 - Hirsch-Index (Quantität und Qualität?)
- ...
- Bewertung durch Experten

Rankings sind ein üblicher Ansatz zur Bewertung von Wissenschaft / Wissenschaftler*innen

- Sie sind günstig zu haben ... und ineffektiv und ungenau.
- Aus ihnen folgt häufig "Rankings First, Good Science Second".

Aber was hat dieser Rang eigentlich mit dem Prozess, der einer kreativen Entdeckung in der Wissenschaft zugrunde liegt, zu tun?

“Research has to be evaluated
for rigour, originality and significance,
for the light it casts and
for the economic and heuristic value;
these qualities may be difficult to assess,
but we should try.”

Peter Lawrence: Lost in publication. How measurement harms science.
Ethics in Science and Environmental Politics 8, 2008

2. Publikationen

Warum ist es für Wissenschaftler*innen wichtig, ihre Arbeiten zu veröffentlichen?

Um neue Erkenntnisse/Informationen zu verbreiten, damit

1. andere daraus lernen können
2. andere Wissenschaftler*innen die Studien wiederholen oder mit zusätzlichen Beobachtungen oder Experimenten auf ihnen aufbauen können
3. die finanzielle oder anderweitige Unterstützung gerechtfertigt werden kann

...

Wie sieht es mit der Autorenschaft aus?

The screenshot shows a news article from the website of the journal 'nature'. The header features the word 'nature' in a large, white, serif font, with 'International weekly journal of science' in smaller text below it. A navigation bar above the article includes links for Home, News & Comment, Research, Careers & Jobs, Current Issue, Archive, Audio & Video, and Forum. Below the navigation bar is a breadcrumb trail: News & Comment > News > 2018 > May > Article. The main title of the article is 'Physics paper sets record with more than 5,000 authors'. Below the title is a subtitle: 'Detector teams at the Large Hadron Collider collaborated for a more precise estimate of the size of the Higgs boson.' The author's name, 'Davide Castelvecchi', is listed, along with the publication date, '15 May 2015'. The text of the article discusses the joint paper from the ATLAS and CMS experiments at the LHC, noting that while the first nine pages describe the research itself, the remaining 24 pages are dedicated to listing the numerous authors and their institutions.

NATURE | NEWS



Physics paper sets record with more than 5,000 authors

Detector teams at the Large Hadron Collider collaborated for a more precise estimate of the size of the Higgs boson.

Davide Castelvecchi

15 May 2015

Only the first nine pages in the 33-page article, published on 14 May in *Physical Review Letters*¹, describe the research itself — including references. The other 24 pages list the authors and their institutions.

The article is the first joint paper from the two teams that operate ATLAS and CMS, two massive detectors at the Large Hadron Collider (LHC) at CERN, Europe's particle-physics lab near Geneva, Switzerland. Each team is a sprawling collaboration involving researchers from dozens of institutions and countries.

“Ich habe ein *Nature*-Paper
und zwei *Cell*-Paper.”

Halten wir nun die Zeitschrift für wichtiger als die
wissenschaftliche Aussage?

Ist es so, dass wir, nur wenn wir in einem Top-Journal
veröffentlichen, tatsächlich auch „angekommen“ sind?

Zitations-basierte Metriken

„Simple citation analysis presupposes a highly rational model of reference-giving, in which citations are held to reflect primarily scientific appreciation of previous work of high quality or importance, and potential citers all have the same chance to cite particular papers ...”

B.R. Martin & J. Irvine, *Research Policy* 12, 1983

Zitations-basierte Metriken

Warum wird eine Publikation zitiert?

Was sagt eine Zitation über die Qualität der
zitierten Publikation aus?

Ist der damit verbundene
Bewertungsprozess objektiv?

“It is hard to escape the suspicion that many authors do not read every paper they cite, and instead tend to cite those papers that appear most often on other author’s references lists.”

“Deciphering impact factors”, *Nature Neuroscience* 6, Editorial, August 2003

“Some advisers even encourage young academics to publish the smallest possible slivers of their work to raise self-confidence and satisfy bean counters.”

Jevin D. West, *Nature* 465, Opinion, 2010

3. Bewertung von Forschung und von Wissenschaftler*innen: Impact-Faktor und Hirsch-Index

Der **Impact-Faktor** ist eine Möglichkeit, die durchschnittliche Anzahl an Zitationen zu berechnen, die die Artikel einer Zeitschrift über einen bestimmten Zeitraum erhalten.

Wie wird der Impact-Faktor berechnet?

Journal X's 2014 impact factor =

Citations in 2014 to all articles
published in Journal X in 2012 and 2013

divided by

Number of articles deemed to be “citable”
that were published in Journal X in 2012 and 2013

A dubious practice

The numerator in the equation can include citations to articles (such as news, comments, opinions, etc.) that do not appear in the list of ‘citable’ items — generally restricted to original research papers and review articles.

“The use of journal impact factors has gone much further [than just ranking journals], extending to the evaluation of individual institutes, departments and scientists. Many evaluating bodies look at scientists’ publication records and evaluate the quality of their output in terms of the impact factors of the journals in which their papers appear.”

D. Adam, “The Counting House”, *Nature* 415, News Feature, February 2002

“The use of journal impacts in evaluating individuals has its inherent dangers.

In an ideal world, evaluators would read each article and make personal judgments.”

Eugene Garfield, inventor of the impact factor

“... for a number that is so widely used and abused, it is surprising how few people understand how a journal’s impact factor is calculated, and, more importantly, just how limited it is a means of assessing the true impact of an individual publication in that journal.”

“The impact factor game”, *PLoS Medicine*, Editorial, August 2003



MathSciNet® Journals Member Directory Employment Services Giving to the AMS

PROCEEDINGS OF THE AMERICAN MATHEMATICAL SOCIETY

ISSN 1088-6826(online) ISSN 0002-9939(print)



MathSciNet® Journals Member Directory Employment Services Giving to the AMS

TRANSACTIONS OF THE AMERICAN MATHEMATICAL SOCIETY

ISSN 1088-6850(online) ISSN 0002-9947(print)

Kürzere Artikel (< 10 Seiten)

Im Zeitraum 2000-04:

2.381 Artikel

> 15.000 Seiten

2005 impact factor 0.434

Längere Artikel

Im Zeitraum 2000-04:

1.165 Artikel

> 25.000 Seiten

2005 impact factor 0.846

Die Veröffentlichung einer Arbeit in der Zeitschrift mit dem höheren Impact-Faktor (Transactions) wird als besser bewertet als die mit dem niedrigeren IF (Proceedings).

Sind die Arbeiten in den Transactions tatsächlich doppelt so gut wie die in den Proceedings?

Es geht nicht um **Durchschnittswerte**,
sondern es ist eine Frage von **Wahrscheinlichkeiten**:

Wie hoch ist die Wahrscheinlichkeit, dass eine zufällig ausgewählte Proceedings-Publikation ($IF = 0,43$) mindestens so viele Zitate enthält wie eine zufällig ausgewählte Publikation in Transactions ($IF = 0,85$)?

Die Antwort ist: 62%.

Das bedeutet, dass wir in 62 % der Fälle falsch liegen.

Wir liegen häufiger falsch als richtig.

Der zentrale Aspekt ist die schiefe Verteilung,

Dies ist eine typische Beobachtung für Zeitschriften, die für viele (die meisten?) überraschend ist.

Das ist lediglich eine Folge der **schiefen Verteilung** und des **engen Zeitfensters**, in welchem die Berechnung des IF erfolgt.

Dieses Ergebnis zeigt den Wert einer exakten statistischen Vorgehensweise, weniger das Ergebnis einer intuitiven Beobachtung.

Quelle: Citation Statistics, Joint Committee on Quantitative Assessment of Research
International Mathematical Union, 2008

“Foremost, using a journal’s impact factor to assess the quality of a particular paper or scientist published in the journal is akin to illiteracy in statistics.

Second, achieving scientific merit is not a popular contest, yet the careless use of citation numbers can make it seem so.”

Nature Biomedical Engineering, Editorial, July 2017

“Everything should be made
as simple as possible,
but not simpler.”

Albert Einstein

Hirsch index

An index to quantify an individual's scientific research output

J. E. Hirsch*

Department of Physics, University of California at San Diego, La Jolla, CA 92093-0319

Communicated by Manuel Cardona, Max Planck Institute for Solid State Research, Stuttgart, Germany, September 1, 2005 (received for review August 15, 2005)

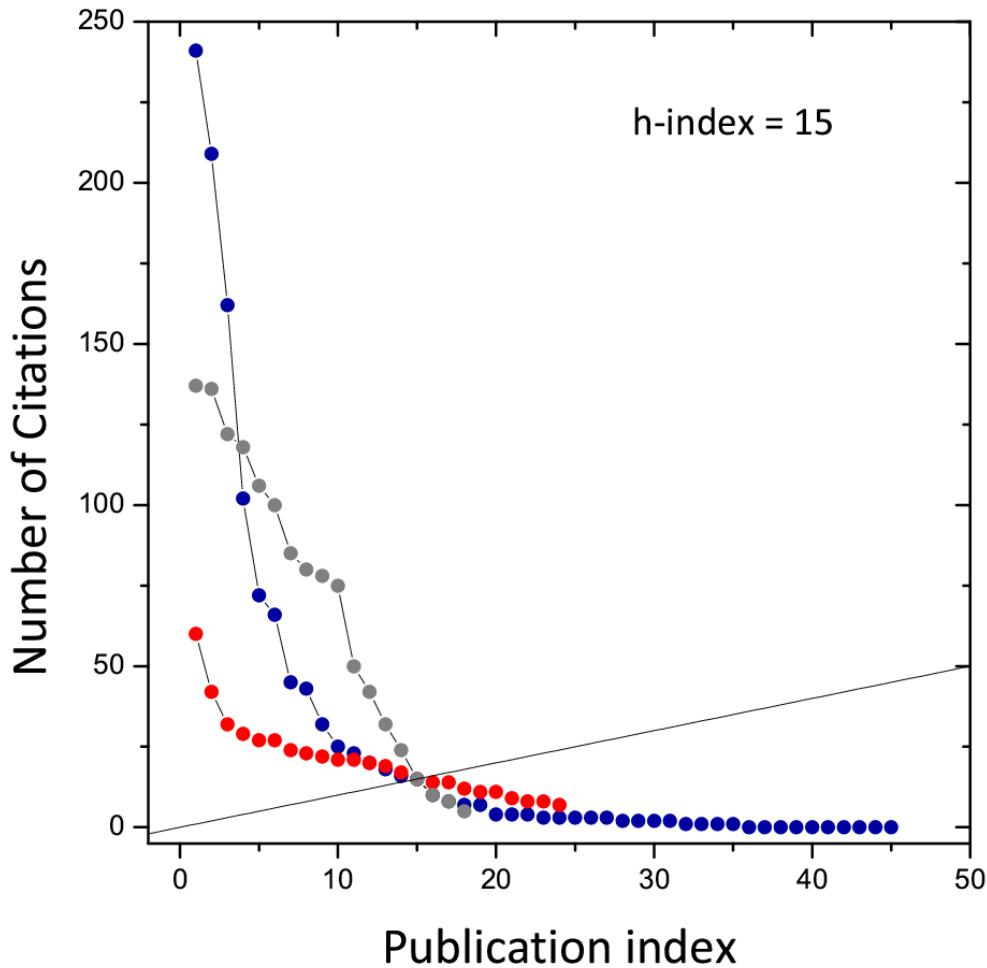
I propose the index h , defined as the number of papers with citation number $\geq h$, as a useful index to characterize the scientific output of a researcher.

citations | impact | unbiased

J.E. Hirsch, PNAS 102, November 2005

“A scientist has index h if h of his/her N_p papers have at least h citations each, and the other (N_p-h) papers have no more than h citations each.”

Hirsch index



Beispiel:

Drei konstruierte Fälle:
Wissenschaftler mit
18, 24 und 45
Publikationen,
die alle den gleichen
h-Index von 15 haben.

Hirsch about Hirsch index

“There are, however, a number of caveats that should be kept in mind.

Obviously, a single number can never give more than a rough approximation to an individual’s multifaceted profile, and many other factors should be considered in combination in evaluating an individual.

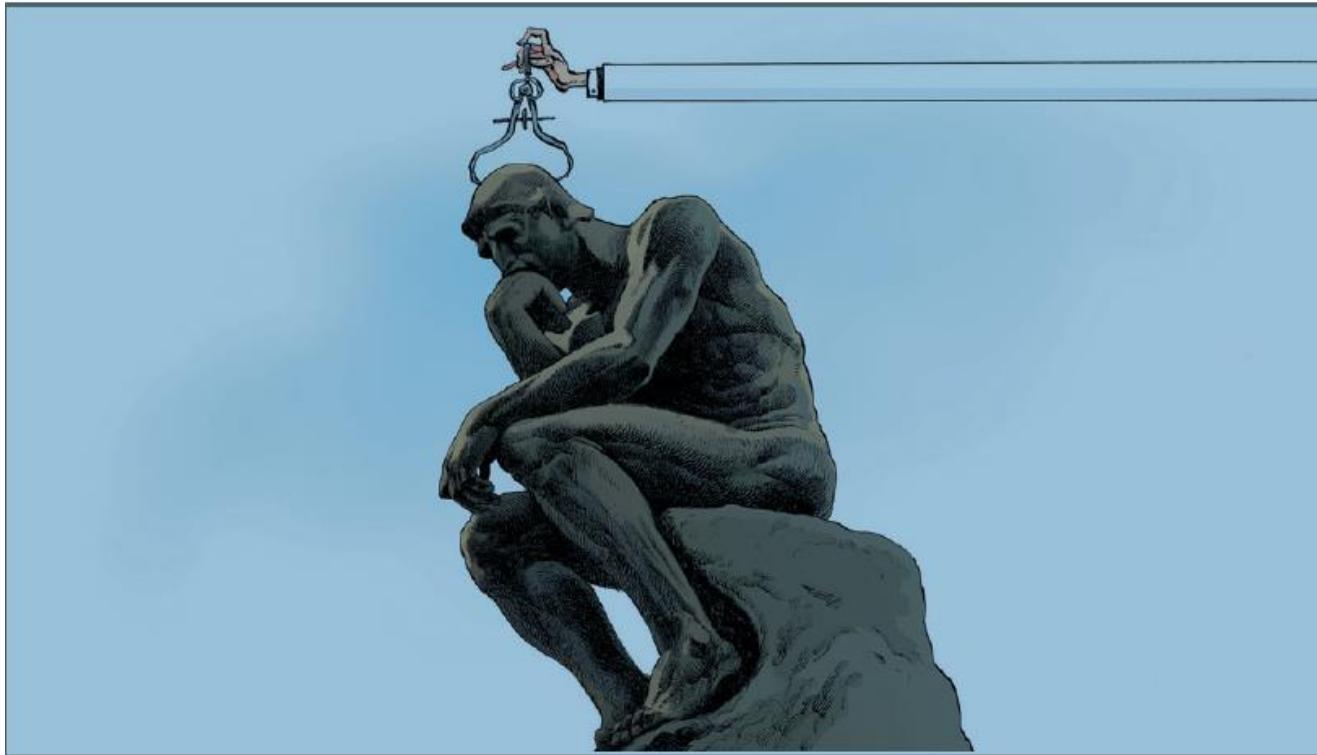
Furthermore, the fact that there can always be exceptions to rules should be kept in mind, especially in life-changing decisions such as the granting or denying of tenure.”

J.E. Hirsch, PNAS 102, November 2005

4. Gibt es einen Ausweg?

“Research usually has multiple goals,
and it is therefore reasonable
that its value must be judged
by multiple criteria.”

Joint Committee on Quantitative Assessment of Research
International Mathematical Union (2008)



The Leiden Manifesto for research metrics

<http://www.leidenmanifesto.org/>

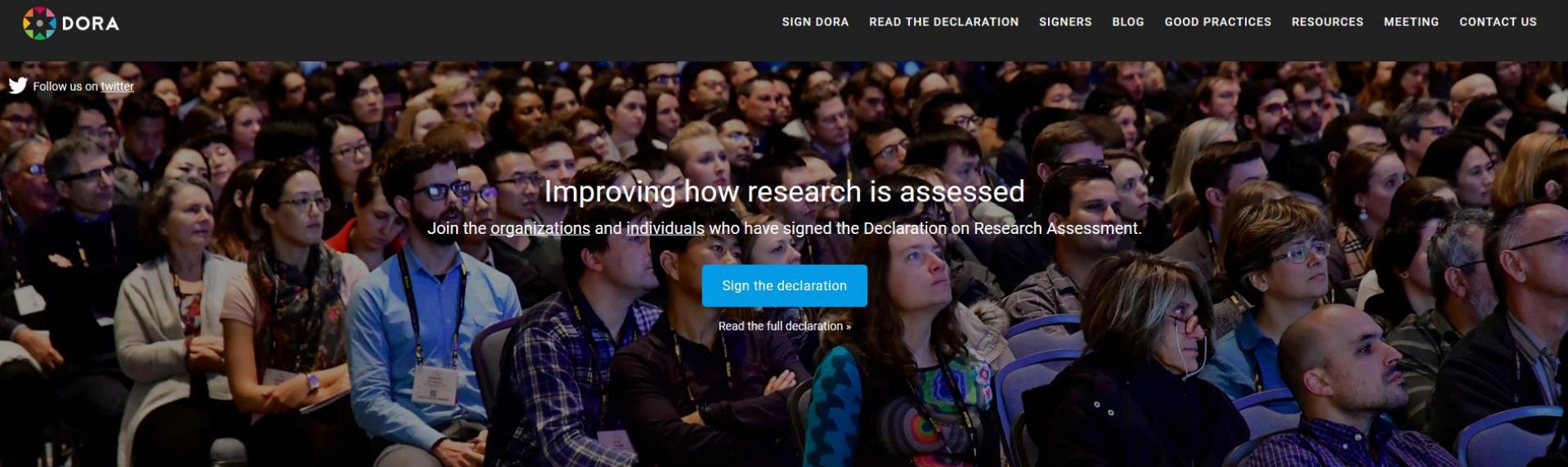
D. Hicks et al., *Nature* 520, Comment, April 2015

Best practice in metrics-based research assessment: 10 principles

<http://www.leidenmanifesto.org/>

1. Quantitative evaluation should support qualitative, expert assessment.
2. Measure performance against the research missions of the institution, group or researcher.
3. Protect excellence in locally relevant research.
4. Keep data collection and analytical processes open, transparent and simple.
5. Allow those evaluated to verify data and analysis.
6. Account for variation by field in publication and citation practices.
7. Base assessment of individual researchers on a qualitative judgement of their portfolio.
8. Avoid misplaced concreteness and false precision.
9. Recognize the systemic effects of assessment and indicators.
10. Scrutinize indicators regularly and update them.

San Francisco Declaration on Research Assessment (DORA)



The image shows a large, diverse audience of people seated in rows, facing forward in what appears to be a conference or meeting hall. The lighting is focused on the audience, creating a sense of depth and engagement. In the foreground, there is a graphic overlay with text and buttons.

Follow us on twitter

Improving how research is assessed

Join the organizations and individuals who have signed the Declaration on Research Assessment.

[Sign the declaration](#)

[Read the full declaration »](#)

<https://sfdora.org/>

Annual Meeting of The American Society for Cell Biology (ASCB)
San Francisco, CA, on December 16, 2012

San Francisco Declaration on Research Assessment (DORA)

Oberstes Ziel

Förderung eines echten Wandels in der Forschungsbewertung

Empfehlungen

für Förderorganisationen, Institutionen, Verlage,
Organisationen, die Metriken liefern, Wissenschaftler*innen

Unterzeichner (Organisationen / Einzelpersonen)

Mai 2018 473 / 12.023

März 2019 1.233 / 13.869

Nov. 2019 1.557 / 15.006

Durchgängige Themenschwerpunkte beziehen sich auf die Notwendigkeit,

- auf die Verwendung von zitations-basierten Metriken (wie den IF) bei Finanzierung, Berufung und Beförderung zu verzichten
- Forschung nach ihren eigenen Werten / Verdiensten / Leistungen, und nicht auf der Grundlage des Journals, in dem die Forschung veröffentlicht wird, zu bewerten
- die Möglichkeiten der Online-Publikation zu nutzen (z.B. die unnötige Begrenzung der Anzahl der Wörter, Zahlen und Verweise in Artikeln zu lockern)
- neue Indikatoren für Bedeutung und Wirkung einer Publikation zu erforschen.

Good practices (eine Auswahl)

Geldgeber / Kostenträger

- EMBO Long-term fellowships
- European Commission
- Higher Education Funding Council for England
- U.S. National Institutes of Health
- U.S. National Science Foundation
- Wellcome

Fachgesellschaften

- Académie des Sciences, Leopoldina and Royal Society
- American Society for Cell Biology

Forschungsinstitute / Universitäten

- University of California, Berkeley, US (Helen Wills Neuroscience Institute)
- University College of London, UK
- University of Colorado School of Medicine, US
- University of Glasgow, UK

Recent developments

Support The Guardian Subscribe Find a job Sign in Search ▾

News Opinion Sport Culture Lifestyle More ▾

The Guardian view Columnists Cartoons Opinion videos Letters

Peer review and scientific publishing

Opinion

How journals like Nature, Cell and Science are damaging science

Randy Schekman

Mon 9 Dec 2013 19.30 GMT

The incentives offered by top journals distort science, just as big bonuses distort banking



WORLD VIEW

A personal take on events



How will you judge me if not by impact factor?

Stop saying that publication metrics don't matter, and tell early-career researchers what does, says John Tregoning.

„The journal impact factor is wrong in so many ways, but it is so easy.“ (John Tregoning, *Nature* 558, World View, June 2018)

International edition ▾

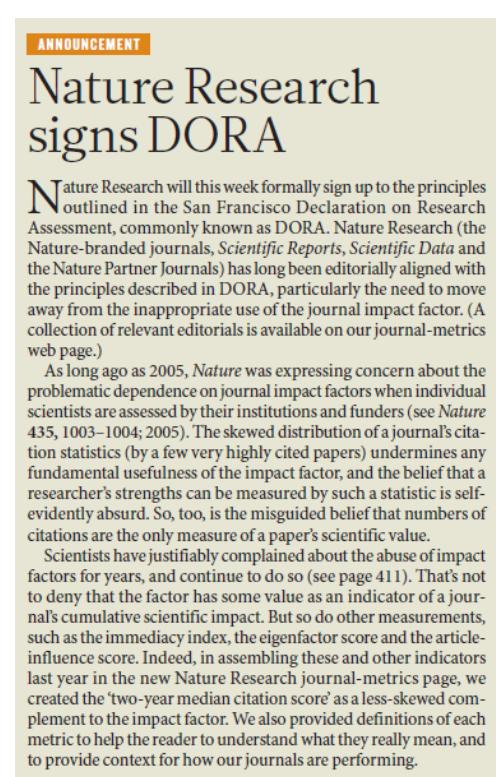
ANNOUNCEMENT

Nature Research signs DORA

Nature Research will this week formally sign up to the principles outlined in the San Francisco Declaration on Research Assessment, commonly known as DORA. Nature Research (the Nature-branded journals, *Scientific Reports*, *Scientific Data* and the Nature Partner Journals) has long been editorially aligned with the principles described in DORA, particularly the need to move away from the inappropriate use of the journal impact factor. (A collection of relevant editorials is available on our journal-metrics web page.)

As long ago as 2005, *Nature* was expressing concern about the problematic dependence on journal impact factors when individual scientists are assessed by their institutions and funders (see *Nature* 435, 1003–1004; 2005). The skewed distribution of a journal's citation statistics (by a few very highly cited papers) undermines any fundamental usefulness of the impact factor, and the belief that a researcher's strengths can be measured by such a statistic is self-evidently absurd. So, too, is the misguided belief that numbers of citations are the only measure of a paper's scientific value.

Scientists have justifiably complained about the abuse of impact factors for years, and continue to do so (see page 411). That's not to deny that the factor has some value as an indicator of a journal's cumulative scientific impact. But so do other measurements, such as the immediacy index, the eigenfactor score and the article-influence score. Indeed, in assembling these and other indicators last year in the new Nature Research journal-metrics page, we created the 'two-year median citation score' as a less-skewed complement to the impact factor. We also provided definitions of each metric to help the reader to understand what they really mean, and to provide context for how our journals are performing.



Nature 544, Editorial, April 2017

Recent developments

DAVE GUTTRIDGE



Words were a good start – now it is time for action

*Five years ago, the Declaration on Research Assessment was a rallying point. It must now become a tool for fair evaluation, urges **Stephen Curry**.*

Nature 554, World View, February 2018

Multiple Assessment

(relative Bedeutung variiert je nach Disziplin)

Peer Review (aber: Bias möglich)

Citation statistics (aber: Missbrauch möglich)

Fördermittel

Einladungen (zu Vorträgen, Gastaufenthalten, etc.)

Mitgliedschaften in Editorial Boards

Auszeichnungen

Mitarbeit in Komitees

Öffentliches Engagement

Manuskriptbegutachtungen

Bedeutung eines “Teamplayers”

...

„In ihrer Denkschrift mahnt die DFG eine Überprüfung der Praxis an, quantitative Kriterien als dominierenden Bewertungsmaßstab zu nutzen. Dort wird explizit auch die ergänzende Orientierung am Impact-Faktor von Zeitschriften kritisiert, der durch den Bezug auf die Zitationshäufigkeit ebenfalls nur einen quantitativen, bedingt interpretierbaren Faktor darstelle.“

Empfehlungen zu wissenschaftlicher Integrität, Positionspapier des Wissenschaftsrates 2015

“Not everything that counts can be counted, and
not everything that can be counted counts.”

William Bruce Cameron
(häufig Albert Einstein zugeschrieben)