ISSI Newsletter
Quaterly e-Newsletter of the International Society for Scientometrics and Informetrics
ISSN 1998-5460
#55 / Volume 14 Number 3
September 2018

CONTENTS

STI 2018 LEIDEN SCIENCE, TECHNOLOGY AND INNOVATION INDICATORS IN TRANSITION

CONFERENCE REPORT

by **PAUL WOUTERS, ANDREA REYES ELIZONDO, RODRIGO COSTAS, THOMAS FRANSSEN, ALFREDO YEGROS**

The 23rd International STI Conference *Science, Technology & Innovation Indicators in Transition* was held from 12-14 September 2018 in Leiden. The conference was organized in collaboration with the European Network of Indicator Developers (ENID) and hosted by the Centre for Science and Technology Studies (CWTS) at Leiden University.

As reflected by the program, the conference discussed both how these indicators are changing in conceptual and technical terms, and

The opening of the conference *Photo courtesy of © Henri de Winter*

At the posters session
Photo courtesy of © Henri de Winter

The social event at the National Museum of Antiquities
Photo courtesy of © Henri de Winter

how they in their turn shape the process of transformation in science and society. Given the lively debates at the conference in response to the 160 papers and 33 posters in various formats (intense small-scale interactions, large plenary sessions, and lots of face to face thinking together), both topics will only become more important in the coming years.

The program was very diverse, presenting a more interdisciplinary reach of the field of science, technology, and innovation indicators both in terms of where it learns from and what it teaches to.

Attendees to a session *Photo courtesy of © Henri de Winter*

The keynotes were presented by Paula Stephan, renowned economist of science and a path breaking analyst of scientific careers, Cameron Neylon, one of the leading lights and activist in the open science movement, and Paul Wouters, director of the CWTS.

There were also twelve special tracks, ranging from data issues in the study of social media to institutionalist approaches in sociology, assessments as participatory explorations, the policy of responsible research and innovation, and public-private interactions in business.

The conference did not only have breadth, it also had depth. As can be read in the proceedings, there were papers on the technical refinement of new scientometric and computer science tools, the properties of new indicator concepts, the new wealth of data enabled by digital communication networks, the mapping of science and innovation, and the most productive theoretical framework to understand the social life of indicators.

At the conference, 370 participants from 36 different countries based at more than 300 organizations met. The historic center of Leiden with its multitude of cafes and restaurants turned out to be a stimulating context for meeting new colleagues, turning colleagues into friends, and strengthening existing friendships. The reception was held at the National Museum of Antiquities which gave the attendees a superb sample of the rich cultural heritage available in Leiden.

During the 23rd International STI Conference, Clarivate Analytics awarded Dr. Orion Penner (École Polytechnique Fédérale de Lausanne) the 2018 Eugene Garfield Award for Innovation in Citation Analysis for his work on the study of research publication as pivot elements between old and new ideas. There was also a prize for the best three posters presented at the conference, which was sponsored by Digital Science. The winners were:

- 1st place: Dorte Henriksen with "The effect of bibliometric performance systems on Danish economists and political scientists"
- 2nd place: Jonathan Dudek, Timothy D. Bowman, and Rodrigo Costas with "When do tweeters tweet about Science? Exploratory analysis of the Twitter dissemination of scientific publications by weekdays and months"
- 3rd place: Rodrigo Liscovsky Barrera with "Overcoming the divide in SSTI: a mixed method and multi-level analysis of internationalisation in South American biomedical Research"

The next STI conference will be held in Rome (Italy) from 2 to 5 September 2019. It will be hosted by Sapienza University of Rome at the main campus in Piazzale Aldo Moro. The call for papers has already been announced.

'Best poster' awardees (D. Henriksen, J. Dudek, R. L. Barrera) Photo courtesy of © Henri de Winter

Closure of the conference Photo courtesy of © Henri de Winter

ISSI 2019: NOVEL APPROACHES TO THE DEVELOPMENT AND APPLICATION OF INFORMETRIC AND SCIENTOMETRIC TOOLS

WITH A SPECIAL STI INDICATORS CONFERENCE TRACK

17th INTERNATIONAL CONFERENCE ON SCIENTOMETRICS & INFORMETRICS

2–5 SEPTEMBER 2019, ROME ITALY
SAPIENZA UNIVERSITY OF ROME, PIAZZALE ALDO MORO, 5 ROME

CALL FOR PAPERS

The Organizing Committee would like to invite participants to submit a contribution to the 17th International Society of Scientometrics and Informetrics Conference that will be held in Rome, Italy (https://www.issi2019.org/). The ISSI 2019 Conference will provide an international forum for scientists, research managers and administrators, as well as information and communication related professionals to share research and debate the advancements of informetric and scientometric theory and applications. The conference is organized under the auspices of ISSI – the International Society for Informetrics and Scientometrics (http://www.issi-society.org/).

It will have a Special STI Indicators Conference Track on *Challenges to the Assessment of Science, Technology and Innovation*. This track is organised by the European Network of Indicator Designers (ENID) (http://www.forschungsinfo.de/ENID/). In this way, the current ISSI conference represents a first experiment to bring together the two conferences in a particular year.

CONFERENCE TOPICS

With this scope in mind, major conference topics of interest include, but not limited to:

MAIN TOPICS	SUBTOPICS (NON-EXCLUSIVE LIST)
Informetric, scientometric, bibliometric, webometric, almetrics, datasources	• Data quality, accuracy, completeness, disambiguation • Web of Science, Scopus, Google Scholar • Publication archives and repositories • Research information systems
Electronic scholarly publishing: new developments, access modalities, costs	• Open Access • Bibliotheconomics • Pricing of journals • The role of electronic scholarly books
Full text analyses of scholarly documents	• Computational linguistic techniques • Citation context studies • Novel indicators derived from full texts
Knowledge discovery and data mining	• Big informetric data sets • Novel models and algorithms • Automatic topic clustering • Search engines
Visualisation and Science Mapping: methods and applications	• Novel methodologies and software packages • Emerging topics and research fields • Informetric approaches to S&T forecasting
Usage analysis: methods and applications	• Its potential and novel applications • Patterns in full text downloads • Article recommender systems
S&T indicators	• Mathematical-statistical aspects • Novel indicators • Validation studies • Novel applications
Assessment of higher education institutions	• University rankings • Novel bibliometric and webometric approaches • Indicators of teaching, research and third mission
Assessment of individual researchers and research groups	• Author-level bibliometrics • Indicators for early career scientists • Authorship conventions • Career paths
Scientific-scholarly internationalization, collaboration & mobility	• International and national collaboration • Brain drain phenomena • The effects of internationalization
The science-technology interface	• Authors and inventors • Article-patent citations • The science base of technology
Research integrity	• Research integrity policies • Misconducts in scholarly publishing • Retractions
Open science	• Open data • Reproducibility in science • Open access
The application of informetric methods in other disciplines	• Library and information science • Sociology of science • History of science • Gender studies
Approaches to informetric and related studies borrowed from other disciplines	• Econometric studies of efficiency • Tools from the physics of complex systems • Social network analysis • Higher education studies • Bioinformatics • Computational linguistics

TOPICS OF THE SPECIAL STI INDICATORS-ENID CONFERENCE TRACK: NEW CHALLENGES TO THE ASSESSMENT OF SCIENCE, TECHNOLOGY AND INNOVATION
The implementation and use of indicators in policy practice
Impact of Research and Innovation Policies
Assessing the social value of research
Behavioural effects of indicators
Open Science policies and their impact on scientific knowledge production and exchange
Public-private collaboration
Research careers and mobility
Risk-taking in science, technology and innovation
Assessment of Science-Society interactions
Evolving geographies of ST&I
Social innovation: New concepts and indicators

SCOPE

The goal of ISSI 2019 is to bring together scholars and practitioners in the area of informetrics, bibliometrics, scientometrics, webometrics and altmetrics to discuss new research directions, methods and theories, and to highlight the best research in this area.

In the special STI Conference track, we will focus on the use of indicators in different contexts, ranging from understanding institutional, structural, and developmental processes to their use as analytical tools in knowledge management and STI policy decision-making.

SUBMISSION OF PAPERS, TUTORIALS AND WORKSHOPS PROPOSALS

We ask for researchers worldwide to submit original full research papers, research-in-progress papers or posters, as well as tutorials and workshops, with a special emphasis on the future of this area and on its links with other disciplines.

More information is available at the following webpage: https://www.issi2019.org/

Photo courtesy of © Balázs Schlemmer | schlemmerphoto.com

Photo courtesy of © Balázs Schlemmer | schlemmerphoto.com

Photo courtesy of © Balázs Schlemmer | schlemmerphoto.com

IMPORTANT DATES

Conference dates:	2–5 September 2019
Submission deadline[1]:	15 January 2019
Notification of acceptance:	15 March 2019
Submission deadline for posters[2]:	20 March 2019
Notification of poster submissions:	15 April 2019
Submission deadline for Doctoral Forum:	20 March 2019
Result announcement for Doctoral Forum:	30 April 2019
End of Early Bird registration:	20 May 2019
Submission of final papers/posters[3]:	30 May 2019
Draft program available:	early July 2019

[1] • **full papers:** max 12 pages including all tables, figures and refs according to the template
• **research-in-progress papers:** max 6 pages including all tables, figures and refs according to the template and
• **workshops/tutorials:** max 2 pages length proposal

[2] max 2 pages length, according to the template

[3] at least one author must register

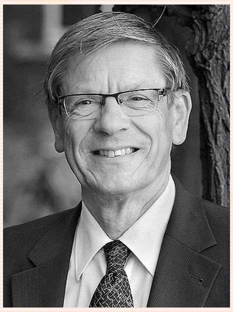
© Góngora / Wikimedia Commons CC-BY-SA-4.0

THE ORGANIZING COMMITTEE OF ISSI 2019

- Prof. Giuseppe Catalano,
- Prof. Cinzia Daraio (E-mail: cinzia.daraio@uniroma1.it),
- Prof. Giancarlo Ruocco, and
- Prof. Henk F. Moed (Program Chair),

Sapienza University of Rome, Italy.

THE AUSTRALIAN AND NEW ZEALAND FIELDS OF RESEARCH (FoR) CODES



RONALD ROUSSEAU
KU Leuven, Facultair Onderzoekscentrum ECOOM
University of Antwerp, Faculty of Social Sciences
ronald.rousseau@kuleuven.be

1. INTRODUCTION

This short note is inspired by two contributions to the latest Science and Technology Indicators Conference, held in Leiden, September 2018. The first one is the article by Haunschild et al. (2018) in which the authors compared three field categorizations for normalized citation scores. The second one is the presentation of *Dimensions*, as a publication (Hook et al., 2018) inserted in the conference bag and as an information booth in the Breezaal.

In the first article Haunschild et al. compared normalized citation scores for the field of chemistry based on journal sets (WoS journals), Chemical Abstracts sections (an intellectual assignment) and on citation relations, as obtained by Waltman and van Eck (2012). They concluded that the agreement between these three methods was lower than what they had expected. They concluded that more investigations – using more classification schemes – are needed. I add that if normalized citation scores turn out to be too dependent on used classification schemes one should thoroughly study these classification schemes and, if necessary, try to come to a universal agreement.

Dimensions is a new scholarly search database (www.dimensions.ai), part of Digital Science, which tries to provide context for the whole scientific research chain, including grants, research, conferences, data sets, publications, tweets and blogs, citations, clinical trials, patents and policy documents. This is not the place to go in detail about all these aspects, but I note that the subject categorization is based among others on the FoR codes (Hood et al., 2018, p.9, Figure 6).

2. FoR CODES

The acronym FoR stands for Fields of Research. It is a classification which is part of the Australian and New Zealand Standard Research Classification (ANZSRC). FoR is a hierarchical classification with three levels, namely Divisions (2 digits), Groups (4 digits) and Fields (6 digits). Each level is identified by a unique number. There are 22 Divisions shown in the appendix. The full classification can be found on the website of the Australian Bureau of Statistics (http://www.abs.gov.au/ausstats). Although this classification is explicitly aimed to be comparable with international classifications such as OECD's Fields of Science 2007 classification, it takes local aspects into account. For example, Group 1802 is Maori law, code 200321 refers to Te Reo Maori (Maori Language) while code 200201 refers to Aboriginal and Torres Strait Islander Cultural Studies.

Activities of members of ISSI will mostly fall under:

08 INFORMATION AND COMPUTING SCIENCES
Group: 0807 Library and Information Studies

and finally in one of these fields:

080704 Information Retrieval and Web Search
080705 Informetrics
080706 Librarianship

But also under:

16 STUDIES IN HUMAN SOCIETY
Group: 1605 Policy and administration
Field: 160511 Research, Science and Technology Policy

Or:

Group: 1608 Sociology
Field: 160808 Sociology and Social Studies of Science and Technology

In the framework of the Excellence in Research for Australia (ERA) framework FoR codes are assigned to journals (between one and three FoR codes, including M for multidisciplinary journals), researchers and research outputs. Assessment panels in Australia too are composed based on FoR codes.

3. FoR CODES IN THE SCIENTIFIC LITERATURE

FoR codes have almost exclusively been studied by Gaby Haddow (Haddow & Noyons, 2013; Haddow, 2015), more specifically in the context of their use in the ERA framework. When institutions submit their ERA data they can request that a different code or codes are assigned to a specific article if they can argue that 66% or more of the content is in another code. For MD assigned journals article codes are selected by the institution and/or the individual. In (Haddow & Noyons, 2013) the authors conclude that the assignment of FoR codes to journals is inadequate for the purpose of valuing contributions at the micro level. Indeed, in the ERA exercise a journal set classification is used as a surrogate for a publication's "true" field classification.

In (Haddow, 2015) the author explores the impact of the FoR research classification on the visibility of research contributions in *Education* and in *Language, communication and culture.* In this study the author concludes that for a sizeable proportion of the articles written by authors from these two fields the classification by FoR codes does not correspond with the authors' organizational unit name. This has the potential to disadvantage researchers and their organizational units in the sense that research results may be assessed as belonging to a field outside the authors' affiliation.

In the only article I know that uses FoR codes, in contrast with those that investigate its use, and published in an international mainstream journal, Bromham et al. (2016) found that whatever the general perception, interdisciplinary research, as measured using FoR codes, was less funded than non-interdisciplinary research.

CONCLUSION

In view of its use in *Dimensions* and with the results of Haunschild et al. (2018) in mind I expect an increased use of FoR codes in the scientific literature. I myself have already been asked by a reviewer to include *Dimensions* as a database in my investigations (Hu & Rousseau, 2018). I think that a thorough investigation of the pros and cons of the Australian and New Zealand Fields of Research (FoR) codes is in order.

ACKNOWLEDGMENT

I thank Gaby Haddow (Curtin University, Australia) for helpful information about the Australian and New Zealand Fields of Research (FoR) codes.

REFERENCES

Bromham, L., Dinnage, R., & Hua, X. (2016). Interdisciplinary research has consistently lower funding success. *Nature,* 534(7609), 684-687.

Haddow, G. (2015). Research classification and the social sciences and humanities in Australia: (Mis)Matching organizational unit contribution and the impact of collaboration. *Research Evaluation*, 24(3), 325-339.

Haddow, G., & Noyons, E. (2013). Misfits? Research classification in research evaluation: Visualizing journal content with fields of research codes. In: (Gorraiz, J., Schiebel, E., Gumpenberger, C., Hörlesberger, M., & Moed, H., Eds.) *Proceedings of ISSI 2013* (pp. 1210-1224). Vienna: AIT.

Haunschild, R., Marx, W., French, B., & Bornmann, L. (2018). Relationship between field-normalized indicators calculated with different approaches of field-categorization. In: (Costas, R., Franssen, T. and Yegros-Yegros, A., Eds.) *Proceedings of the 23rd International Conference on Science and Technology Indicators* (pp. 11-18).

Hook, D.W., Porter, S.J., & Herzog, C. (2018). Dimensions: Building context for search and evaluation. *Frontiers in Research Metrics and Analytics*, 3, article 23. https://doi.org/10.3389/frma.2018.00023

Hu, XJ., & Rousseau, R. (2018). A new approach to explore the knowledge transition path in the evolution of Science & Technology: From the biology of restriction enzymes to their application in biotechnology. *Journal of Informetrics*, 12(3), 842-857.

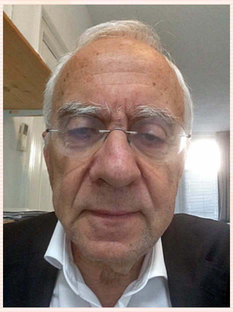
Waltman, L., & van Eck, N.-J. (2012). A new methodology for constructing a publication-level classification system of science. *Journal of the American Society for Information Science and Technology*, 63(12), 2378-2392.

APPENDIX

The Australian and New Zealand Standard Research Classification (ANZSRC) Divisions

01 MATHEMATICAL SCIENCES
02 PHYSICAL SCIENCES
03 CHEMICAL SCIENCES
04 EARTH SCIENCES
05 ENVIRONMENTAL SCIENCES
06 BIOLOGICAL SCIENC
07 AGRICULTURAL AND VETERINARY SCIENCES
08 INFORMATION AND COMPUTING SCIENCES
09 ENGINEERING
10 TECHNOLOGY
11 MEDICAL AND HEALTH SCIENCES
12 BUILT ENVIRONMENT AND DESIGN
13 EDUCATION
14 ECONOMICS
15 COMMERCE, MANAGEMENT, TOURISM AND SERVICES
16 STUDIES IN HUMAN SOCIETY
17 PSYCHOLOGY AND COGNITIVE SCIENCES
18 LAW AND LEGAL STUDIES
19 STUDIES IN CREATIVE ARTS AND WRITING
20 LANGUAGE, COMMUNICATION AND CULTURE
21 HISTORY AND ARCHAEOLOGY
22 PHILOSOPHY AND RELIGIOUS STUDIES

REVISITING RELATIVE INDICATORS AND PROVISIONAL TRUTHS



LOET LEYDESDORFF
Amsterdam School of Communication Research (ASCoR), University of Amsterdam, The Netherlands
loet@leydesdorff.net

TOBIAS OPTHOF
Experimental Cardiology Group, Heart Failure Research Center, Academic Medical Center, The Netherlands
tobias.opthof@gmail.com

Abstract: Following discussions in 2010 and 2011, scientometric evaluators have increasingly abandoned relative indicators in favor of comparing observed with expected citation ratios. The latter method provides parameters with error values allowing for the statistical testing of differences in citation scores. A further step would be to proceed to non-parametric statistics (e.g., the top-10%) given the extreme skewness (non-normality) of the citation distributions. In response to a plea for returning to relative indicators in the previous issue of this newsletter, we argue in favor of further progress in the development of citation impact indicators.

Keywords: indicators, statistics, citation, percentiles

INTRODUCTION

In the *ISSI Newsletter 14*(2), Glänzel & Schubert (2018) argue for using "relative indicators" – e.g., the Mean Observed Citation Rate relative to the Mean Expected Citation Rate *MOCR/MECR* (Schubert & Braun, 1986; cf. Vinkler, 1986) – instead of testing citation scores against their expected values using the Mean Normalized Citation Score *MNCS* (Waltman, Van Eck, Van Leeuwen, Visser, & Van Raan, 2011a and b). The authors note our "concern" about using these relative indicators (Opthof & Leydesdorff, 2010; cf. Lundberg, 2007). However, Glänzel & Schubert (2018) state (at p. 47) that they do not wish to "resume the debate but attempt to shed some light on the premises and the context of indicator design in the mirror of the rules of mathematical statistics."

In their discussion of the indicators, Glänzel & Schubert (2018) pay insufficient attention to the differences in terms of the results of a scientometric evaluation. Are the indicators valid and reliable (Lehman et al., 2006)? Our "concern" was never about the relative indicators as mathematical statistics, but about their use in evaluations. From this latter perspective, the division between two averages instead of first normalizing against expected values can be considered as a transgression of the order of mathematical operations by which division precedes addition.

In the case of *MOCR/MECR*, one first sums in both the numerator and denominator and then divides, as follows:

$$\frac{MOCR}{MECR} = \frac{\sum_{i=1}^{n} c_i / n}{\sum_{i=1}^{n} e_i / n} = \frac{\sum_{i=1}^{n} c_i}{\sum_{i=1}^{n} e_i} \quad (1)$$

In the case of *MNCS*, one first divides and sums thereafter:

$$MNCS = \frac{1}{n} \sum_{i=1}^{n} \frac{c_i}{e_i} \quad (2)$$

Eq. 1 has also been called the "Rate of Averages" (RoA) versus the "Average of Rates" (AoR) in the case of Eq. 2 (Gingras & Larivière, 2011).

THE OLD "CROWN INDICATOR"

The "relative indicators" of Eq. 1 were introduced by the Budapest team in the mid-1980s (Schubert & Braun, 1986; Vinkler, 1986). One of these relative indicators – using the field of science as the reference set – has been used increasingly since approximately 1995 as the so-called "crown indicator" (*CPP/FCSm*)[1] by the Leiden unit *CWTS* (Moed, De Bruin, & Van Leeuwen, 1995). These "relative indicators" are still in use for research evaluations by the ECOOM unit in Louvain headed by Glänzel.

1 *CPP/FCSm* is the total "citations per publication" for a unit under evaluation divided by the mean of the citations in the respective field.

In a vivid debate, Van Raan et al. (2010) first argued that the distinction between RoA and AoR was small and therefore statistically irrelevant. However, both Opthof & Leydesdorff (2010) and Gingras & Larivière (2011) provided examples showing significant differences between the two procedures. Using AoR, one is able to test for the statistical significance of differences in citations among sets of documents. Unlike AoR, RoA comes as a pure number (without error); using this indicator at the time, CWTS and ECOOM invented "rules of thumb" to indicate significance in the deviation from the world standard as 0.5 (Van Raan, 2005) or 0.2 (CWTS, 2008, at p. 7; cf. Schubert & Glänzel, 1983; Glänzel, 1992 and 2010). Even if one tries to circumvent the violation of basic mathematical rules by adding brackets to the equations, these conceptual issues remain.

AoR AND RoA IN THE BANKING WORLD

Glänzel & Schubert (2018) refer to a paper published in the arXiv by Matteo Formenti (2014) from the Group Risk Management of the UniCredit Group. In this risk assessment, the author compares default rates of mortgages issued in the years 2008-2011 during the subsequent five years as risks for the bank. The time of default applies to any mortgage that ends before the scheduled date planned by the bank, either because the individual fails to pay or because the mortgage is paid off before the planned date, which also implies less income for a portfolio holder such as a bank.

The problem formulation is different from that of research evaluation using citations:

1. For a bank it does not matter which customers fail to pay the mortgage in the future, as long as the sumtotal of individual positions of customers does not provide a risk for the bank. The sumtotal provides the reference in RoA;

2. Formenti (2014) missed an important issue: in his test portfolio there are 12 risk groups from 'M1' to 'M12', with the highest risk residing in 'M12'. Neither RoA nor AoR are able to estimate the risk in the highest risk group or the risk groups with a lower but still substantial risk profile; both indicators underestimate the risk by an order of magnitude. Analogously, the risks in the lowest risk group ('M1') are grossly overestimated, regardless of whether RoA or AoR is used. (Because both estimates thus fail, holders of home mortgages pay an interest rate on loans much higher than the current one on the market.)

We do not understand the relation between this example and research evaluations. Are funding agencies distributing money over the scientific community with the aim of avoiding their own bankrupcy?

THE NEW "CROWN INDICATOR"

In the weeks after voicing our critique (in 2010), the Leiden unit turned up another "crown indicator:" *MNCS* or the "mean normalized citation score" (Eq. 2; Waltman, van Eck, van Leeuwen, Visser, & van Raan, 2011 a and b). In our response, we expressed our concern about moving too fast – without sufficient debate – to this alternative (Leydesdorff & Opthof, 2011). Following up on Bornmann & Mutz (2011), we then proposed "to turn the tables one more time" by first specifying criteria for comparing sets of documents in terms of performance indicators independently from specific evaluation contexts and existing infrastructures (Leydesdorff, Bornmann, Mutz, & Opthof, 2011). We formulated these criteria (at pp. 1371f.), as follows:

1. A citation-based indicator must be defined so that the choice of the reference set(s) (e.g, journals, fields) can be varied by the analyst independently of the question of the evaluation scheme. In other words, these two dimensions of the problem (the normative and the analytical ones) have to be kept separate.

2. The citation indicator should accommodate various evaluation schemes, for example, by funding agencies. Some agencies may be interested in the top-1% (e.g., National Science Board, 2010) while others may be interested in whether papers based on research funded by a given agency perform significantly better than comparable non-funded ones (e.g., Bornmann et al., 2010).

3. The indicator should allow for productivity to be taken into account. One should, for example, be able to compare two papers in the 39th percentile with a single paper in the 78th percentile (with or without weighting the differences in rank in an evaluation scheme as specified under 2.).

4. The indicator should provide the user, among other things, with a relatively straightforward criterion for the ranking (for example, a percentage of a maximum) that can then be tested for its statistical significance in relation to comparable (sets of) papers.

5. It should be possible to calculate the statistical errors of the measurement.

Using the publications of seven principal investigators at the Amsterdam Medical Center (AMC), we showed in detail how one can use percentiles and test the non-parametric differences (e.g., in SPSS) using Bonferroni corrections. In our opinion, this should have become the basis for a new "crown indicator", but we are not in the business of using indicators in evaluation practices.

The proposal by Glänzel & Schubert (2018) to return to the first-generation indicators of the mid-80s and 90s stands athwart this progression. The argument that at the aggregate level, relative indicators provide another and sometimes perhaps richer perspective does not legitimate their use in the practice of research evaluations. In a medical practice, for example, if someone deliberately used a value other than the statistically expected one for making a decision, the doctor would be held responsible for the (potentially lethal) consequences. In the rat-race for university positions and academic status, however, this collateral damage seems to be taken for granted.

In policy-making and managerial contexts, one can work with a flawed or outdated indicator so long as no alternatives are at hand (Leydesdorff, Wouters, & Bornmann, 2016). In other words, the functionality of the indicators is a pragmatic issue, and relatively independent of the validity of the results (Dahler-Larsen, 2014; cf. Hicks et al., 2015). As Lehman, Jackson, & Lantrup (2006) formulated: "There have been few attempts to discover which of the popular citation measures is best and whether any are statistically reliable." Gingras (2016, at p. 76) noted that indicators without a foundation in methodology can only be explained by marketing strategies on the part of the producers.

PERSPECTIVES FOR FURTHER RESEARCH

Two main problems remain when working with *MNCS* as a new crown indicator:

1. Using the mean of the (highly skewed) distribution as the expectation (Seglen, 1992). The Leiden Rankings have proceeded using percentiles (Waltman et al., 2012), but in many other evaluation studies *MNCS* is used based on average citation scores in Web-of-Science Subject Categories.
2. Using the Web-of-Science Subject Categories (WCs) for the delineation of the reference sets. These sets are defined at the journal level. Journals, however, are an amalgam of different subfields and therefor a poor basis for creating reference values (Opthof, 2011). WCs remain at the level of journals because the fields are defined as combinations of journals.

Pudovkin & Garfield (2002) described the method and history of how journals have been assigned Subject Categories in the *JCR*. The authors state that journals are assigned categories by "subjective, heuristic methods" (p. 1113), which the authors clarify in a footnote as follows:

> …This method is "heuristic" in that the categories have been developed by manual methods started over 40 years ago. Once the categories were established, new journals were assigned one at a time. Each decision was based upon a visual examination of all relevant citation data. As categories grew, subdivisions were established. Among other tools used to make individual journal assignments, the Hayne-Coulson algorithm is used. The algorithm has never been published. It treats any designated group of journals as one macrojournal and produces a combined printout of cited and citing journal data. (p. 1113.)

According to the evaluation of these authors, in many fields these categories are sufficient; but they also acknowledge that "in many areas of research these 'classifications' are crude and do not permit the user to quickly learn which journals are most closely related" (p. 1113). These problems have not been diminished but have increased with the more recent expansions of the database (Leydesdorff & Bornmann, 2016).

REFERENCES

Bornmann, L., & Mutz, R. (2011). Further steps towards an ideal method of measuring citation performance: The avoidance of citation (ratio) averages in field-normalization. *Journal of Informetrics, 5*(1), 228-230.

CWTS. (2008). AMC-specifieke CWTS-analyse 1997–2006 (access via AMC intranet; unpublished, confidential). Leiden, The Netherlands: CWTS.

Dahler-Larsen, P. (2014). Constitutive Effects of Performance Indicators: Getting beyond unintended consequences. *Public Management Review, 16*(7), 969-986.

Formenti, M. (2014). Means of ratios or ratio of means: statistical uncertainty applied to estimate Multiperiod Probability of Default. Available at https://arxiv.org/pdf/1409.4896

Gingras, Y. (2016). *Bibliometrics and research evaluation: an overview*. Cambridge MA: MIT Press.

Gingras, Y., & Larivière, V. (2011). There are neither "king" nor "crown" in scientometrics: Comments on a supposed "alternative" method of normalization. *Journal of Informetrics, 5*(1), 226-227.

Glänzel, W. (1992). Publication Dynamics and Citation Impact: A Multi-Dimensional Approach to Scientometric Research Evaluation. In P. Weingart, R. Sehringer & M. Winterhagen (Eds.), *Representations of Science and Technology. Proceedings of the International Conference on Science and Technology Indicators, Bielefeld, 10-12 June 1990* (pp. 209-224). Leiden: DSWO / Leiden University Press.

Glänzel, W. (2010). On reliability and robustness of scientometrics indicators based on stochastic models. An evidence-based opinion paper. *Journal of Informetrics, 4*(3), 313-319.

Glänzel, W., & Schubert, A. (2018). Relative indicators revisited. *ISSI Newsletter, 14*(2), 46-50.

Hicks, D., Wouters, P., Waltman, L., de Rijcke, S., & Rafols, I. (2015). The Leiden Manifesto for research metrics. *Nature, 520*, 429-431.

Lehman, S., Jackson, A.D., Lantrup, B.E. (2006). Measures for measures. *Nature, 444*, 1003-1004.

Leydesdorff, L., & Bornmann, L. (2016). The Operationalization of "Fields" as WoS Subject Categories (WCs) in Evaluative Bibliometrics: The cases of "Library and Information Science" and "Science & Technology Studies". *Journal of the Association for Information Science and Technology, 67*(3), 707-714.

Leydesdorff, L., Bornmann, L., Mutz, R., & Opthof, T. (2011). Turning the tables in citation analysis one more time: Principles for comparing sets of documents. *Journal of the American Society for Information Science and Technology, 62*(7), 1370-1381. Leydesdorff, L., & Opthof, T. (2011). Remaining problems with the "New Crown Indicator" (MNCS) of the CWTS. *Journal of Informetrics, 5*(1), 224-225.

Leydesdorff, L., Wouters, P., & Bornmann, L. (2016). Professional and citizen bibliometrics: complementarities and ambivalences in the development and use of indicators – a state-of-the-art report. *Scientometrics, 109*(3), 2129-2150. doi: 10.1007/s11192-016-2150-8

Lundberg, J. (2007). Lifting the crown – citation z-score. *Journal of informetrics, 1*(2), 145-154.

Moed, H. F., De Bruin, R. E., & Van Leeuwen, T. N. (1995). New bibliometric tools for the assessment of national research performance: Database description, overview of indicators and first applications. *Scientometrics, 33*(3), 381-422.

National Science Board (2010). Science and engineering indicators 2010, appendix tables. Arlington, VA, USA: National Science Foundation (National Science Board 10-01).

Opthof, T. (2011). Differences in citation frequency of clinical and basic science papers in cardiovascular disease. *Medical & Biological Engineering & Computing, 49*, 613-621.

Opthof, T., & Leydesdorff, L. (2010). Caveats for the journal and field normalizations in the CWTS ("Leiden") evaluations of research performance. [Article]. *Journal of Informetrics, 4*(3), 423-430. doi: 10.1016/j.joi.2010.02.003 (see also Leydesdorff, L. & Opthof, T. (2010). Normalization, CWTS indicators, and the Leiden Rankings: Differences in citation behavior at the level of fields. Available at http://arxiv.org/abs/1003.3977).

Pudovkin, A. I., & Garfield, E. (2002). Algorithmic procedure for finding semantically related journals. *Journal of the American Society for Information Science and Technology, 53*(13), 1113-1119.

Schubert, A., & Braun, T. (1986). Relative indicators and relational charts for comparative assessment of publication output and citation impact. *Scientometrics, 9*(5), 281-291.

Schubert, A., & Glänzel, W. (1983). Statistical reliability of comparisons based on the citation impact of scientific publications. *Scientometrics,* 5(1), 59-73.

Seglen, P. O. (1992). The Skewness of Science. *Journal of the American Society for Information Science, 43*(9), 628-638.

van Raan, A. F. J. (2005). Measurement of central aspects of scientific research: performance, interdisciplinarity, structure. *Measurement, 3*(1), 1-19.

van Raan, A. F., van Leeuwen, T. N., Visser, M. S., van Eck, N. J., & Waltman, L. (2010). Rivals for the crown: Reply to Opthof and Leydesdorff. *Journal of Informetrics, 4*(3), 431-435.

Vinkler, P. (1986). Evaluation of some methods for the relative assessment of scientific publications. *Scientometrics, 10*(3), 157-177.

Waltman, L., Van Eck, N. J., Van Leeuwen, T. N., Visser, M. S., & Van Raan, A. F. J. (2011a). Towards a New Crown Indicator: Some Theoretical Considerations. *Journal of Informetrics, 5*(1), 37-47.

Waltman, L., van Eck, N. J., van Leeuwen, T. N., Visser, M. S., & van Raan, A. F. J. (2011b). Towards a new crown indicator: An empirical analysis. *Scientometrics, 87*, 467–481.

Waltman, L., Calero-Medina, C., Kosten, J., Noyons, E., Tijssen, R. J., Eck, N. J., . . . Wouters, P. (2012). The Leiden Ranking 2011/2012: Data collection, indicators, and interpretation. *Journal of the Association for Information Science and Technology, 63*(12), 2419-2432.