ISSI Newsletter

QUATERLY E-NEWSLETTER OF THE INTERNATIONAL SOCIETY FOR SCIENTOMETRICS AND INFORMETRICS
ISSN 1998-5460

#31 / VOLUME 08 NUMBER 3
SEPTEMBER 2012

CONTENTS

BRAZILIAN SCIENTOMETRICS: FROM LITTLE TO BIG?

JACQUELINE LETA
Universidade Federal do Rio de Janeiro, Brazil
jleta[at]bioqmed.ufrj.br

Brazilian studies on Bibliometrics / Scientometrics started in the 1970s, when the country's oldest graduate school in information science, currently named IBICT, invited some international scientists as teachers for its master's students. The internationally renowned invitees included Tefko Saracevic, Wilfrid Lancaster and Derek de Solla Price, who taught classes and trained dozens of Brazilian students up to the 1980's (IBICT, 2012).

The presence of these scientists has promoted the birth of Scientometrics[1] in the country by improving, training and qualifying their Brazilian students to perform original research in this field. In fact, this fruitful period was anticipated by some previous discussions and investigationsin the Brazilian information science lit-

[1] From this point on I will use only the term Scientometrics whenever I refer to the field.

ISSI e-Newsletter (ISSN 1998-5460) is published by ISSI (http://www.issi-society.info/).
Contributors to the newsletter should contact the editorial board by e-mail.

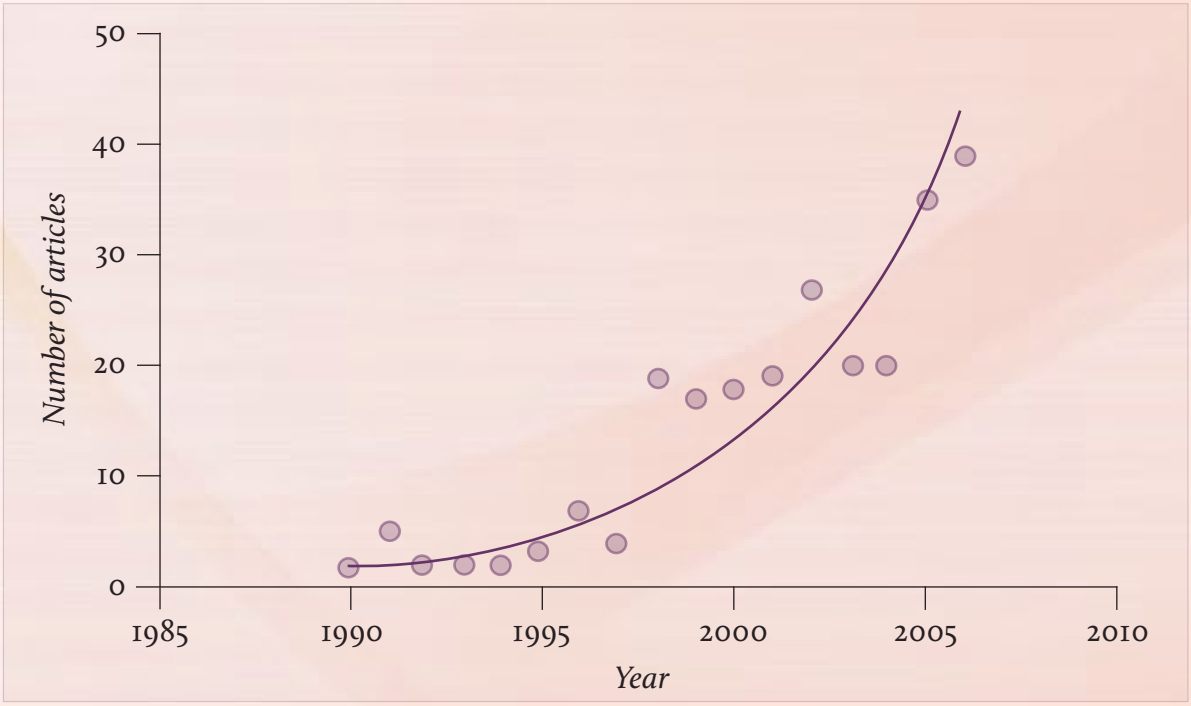

Figure 1. Growth of Brazilian production on scientometrics (from Google Scholar database). Source: Meneghini & Packer, 2010

erature. Urbizagástegui Alvarado (1984), for instance, analyzed different types of publications on the field of information science published in the 1972-1983 period in order to identify the scientometric studies. He found that the majority of these publications were addressed by IBICT, most were published in 1978 and 1979 and had the format of thesis with an international mentor.

During the 1980s and 1990s, researchers from other fields and from other institutes joined the field and started sharing with IBICT the responsibility of training the new generations of Brazilian scientometricians (Pinheiro & Silva, 2008). As a result, the number of studies on Scientometrics has increased remarkably as presented by Meneghini & Packer (Figure 1). Using Google Scholar tools, these authors identified 197 Brazilian publications in the field. For the period 1990-2006, they found that the production increased by 19.5 times. Another feature discussed by Meneghini & Packer was the origin of Brazilian authors: 27% came from Information Science, 26% from Management & Humanities, 19% from Health Sciences, 16% from Biological Sciences and 12% from other experimental fields.

The picture described by Meneghini & Packer reveals not only a growing community but also a wide variety of fields and interest involved in Brazilian scientometric studies. This finding closely parallels with that presented by Glänzel & Schoepflin (1994) when discussing the field within an international context. In fact, the present context of Scientometrics in Brazil encompasses many of the questions raised by these authors to explain a crisis in the field at that time. I would not say we live in a time of crisis but the strengthening and consolidation of the field in Brazil seems to require the overcoming of some old challenges. In the followings, three of them are presented briefly.

Recovering or reconstructing the identity of the field. The diversity of interests and issues revealed by the multiple origins of Brazilian authors points out to a "shift from basic and methodological research to applied bibliometrics". Although this is not an exclusive feature of Brazilian scientometrics, the large and increasing involvement of researchers from other scientific disciplines may bring some consequences to the field identity in Brazil. The integration of more researchers

from information science may be a necessary (but not sufficient) for unifying the field.

Recovering the basic facet of Scientometrics. During the last two decades, scientometric analysis and S&T indicators gained visibility and interest among some Brazilian bodies from public sector, for instance, the Ministry of Science and Technology, the Science Foundation of the State of São Paulo and the Brazilian Academy of Science. They all have used scientometric analyses for evaluating Brazilian researchers, institutes and fields. Such involvement led to a "domination of the interests of science policy" in scientometric analysis which contributed not only to spread the knowledge of some of the field's concepts and analysis but also to consolidate the applicability as its main facet.

Removing misconceptions. The third and last aspect to be highlighted is the "misuse of bibliometric research results and disregard for scientific standards". Superficial analyses and the lack of a complete understanding of the bibliographic databases lead to improper practice, strongly corroborating misconceptions. These are more frequently observed among non-experts, especially science managers. A quick visit at some websites of Brazilian S&T agencies easily reveals some of the most common misuses: comparison among fields and disciplines; citations as a synonym on quality, impact factor of journals as a quality parameter of individual work, indiscriminate use of indices such as the h-index, without considering the effects of field and number of authors.

Recovering the field's identity and basic facet and removing misconceptions are thus some of the challenges Brazilian scientometricians have to face in the next decade or so. This scenario comes together with the fact that Brazilian scientometric community is too scattered. Thus, in order to overcome the challenges and the difficulties imposed by the isolation and diversity of groups and researchers, the First Brazilian Meeting on Bibliometrics and Scientometrics (abbreviated as EBBC in Portuguese) was planned. In 2008, this meeting took place at the Federal University of Rio de Janeiro (http://www.eventos.bvsalud.org/agendas/ebbc1/?lang=pt). Being the organizer of this first event, my aim was first of all putting together Brazilian specialists on Scientometrics as well as non-specialists and science managers to discuss the field's main issues. Wolfgang Glänzel opened this event and gave the audience a complete panorama of Scientometrics in an international context. The 1st EBBC brought together around 60 participants who had the opportunity of attending the presentations of the most important Brazilian experts on the field. After this first meeting the core researchers established a series of initiatives to promote and stimulate Scientometrics in Brazil. Among these initiatives, a national agenda for EBBC was set up and since then two meetings have already been held.

The 2nd EBBC was organized by Maria Cristina Hayashi, professor at the Federal University of São Carlos (located about 250 km from São Paulo) in 2010 (http://www.ebbc.ufscar.br/). This event was designed just as the field's main conference, ISSI, in terms of program structure. Thus, participants could attend different activities such as four keynote addresses, three round-tables, five oral presentation sessions (25 presentations), a poster session (with 57 posters) and one workshop. Following the practice of the previous meeting, the 2nd EBBC had an international guest, Elias Sanz Casado, who had also opened the meeting with a presentation focused on S&T indicators for university management. Since it was inside the university campus, the 2nd EBBC could count with many undergraduate and graduate students and the total number of participants reached almost 180.

The 3rd EBBC has just finished. As the previous one, the meeting was a success in terms of participants: 170 people. Ida Stumpf, professor at Federal University of Rio Grande do Sul, was the organizer (http://www.ufrgs.br/ebbc2012). The venue of the meeting was in a charming city named Gramado (about 100 km from Porto Alegre, where the main university campus is located). For this time, there were two keynote lectures, given by Ed Noyons and Allan Porter; both addressed the hot topic of Maps of Science. Following the same

Ida Stumpf

Maria Cristina Hayashi

Jacqueline Leta

program structure as the previous meeting, the 3rd EBBC had three round-tables, 9 oral sessions (48 presentations), one poster session (54 posters) and two work-shops.

The success of the three meetings led by three women and their "armies" of students and partners confirmed that gathering of experts and non-experts is the requisite for integrating the diversity as well as for strengthening the whole community to form a more unified group. The meetings confirmed also that integration and information exchange may enhance the quality of the community's work. The results are already appearing. By some means, the meetings themselves as well as the proximity of some key researchers have contributed to stimulate Brazilians to take part in the Scientometrics international arena, as we could witnessed by a record number of Brazilians at Durban's ISSI conference last year.

The closer interaction amongst Brazilian scientometricians promoted by the national meetings brought not only a better knowledge of the whole community but also the knowledge of its potentialities and challenges. New students and experts from other fields, including from computer science, are being incorporated into the community, many students and experts are looking for national and international partners, new research groups are being created, new and more complex research projects are being developed, etc.

This movement is a clear signal that Scientometrics in Brazil is big at the present. In addition, it does have the potential to be consolidated as a scientific field in the next few years. On the other hand, we still have to face some challenges or conditions that can impair this process.

REFERENCES

IBICT, Programa de Pós-Graduação em Ciência da Informação. Apresentação. Available at: http://www.ppgci.ufrj.br/index.php/programa/apresentacao. April, 2012.

URBIZAGÁSTEGUI ALVARADO, R. (1984) A Bibliometria no Brasil. *Ciência da Informação* 13 (2) 1:91-105. Available at: http://revista.ibict.br/ciinf/index.php/ciinf/article/view/1444/1063. September, 2011.

PINHEIRO, LVR; SILVA, GS (2008) Cartografia Histórica e conceitual da Bibliometria / Cientometria no Brasil. Available at: http://repositorio.ibict.br/bitstream/123456789/67/1/PinheiroCIPECC.pdf September, 2011.

MENEGHINI, R.; PACKER, A. (2010) The extent of multidisciplinary authorship of articles on Scientometrics and Bibliometrics in Brazil. *Interciência* 35 (7) 510-514.

GLÄNZEL, W. & SCHOEPFLIN, U. (1994) Little Scientometrics, Big Scientometrics... and beyond? *Scientometrics* 30 (2) 375-384.

MANFRED BONITZ (7.3.1931 – 14.8.2012)

OBITUARY

BY ANDREA SCHARNHORST
(With thanks to his sons Marc and Michael)

Devoted passionately to science and still full of ideas and plans, Manfred Bonitz, long-term editorial board member of the journal "Scientometrics" and member of our society since its beginning, could not defeat illness any longer and passed away.

Scientometrics has been his passion for many years. Many of you might remember Manfred from different bi-annual conferences: engaged presenter, critical and supportive discussant, always in good mood, serving as a "photo-documentalist" of our meetings and becoming a good friend with many colleagues. Less well known might be other details of his long scientific career.

Very early on, Manfred recognized the possibilities of citation indexing and promoted the work of Eugene Garfield in the socialist countries – despite the cold war. His book "Wissenschaftliche Forschung und wissenschaftliche Infor-

mation" (Scientific research and scientific information) appeared in a series "Contributions to Research Technology" of the publisher Akademie Verlag Berlin in 1981. In short, it is a manifesto of the importance of the information sciences for scientific progress. As recently pointed out by Wolfgang Glänzel, insights and laws[1] formulated by Manfred are still of great importance.[2]

Manfred Bonitz belongs to the pioneers in the information sciences in the East. As many others, he found his way to information and documentation from physics. He loved being at the Nils Bohr Institute in Copenhagen working there two years, but his heart he had lost in the Soviet Union during his doctorate – to Natascha, his partner for life. Back in the German Democratic Republic (GDR), he left his beloved experimental physics in the 1970s to contribute to the need to build information systems for the sciences. The Central Institute for Nuclear Research in Rossendorf near Dresden became his academic home. As many other scientists working for the Academy of Sciences of the GDR he lost his institutional affiliation in the early 1990s. But, at the same time, he conquered new scientific territory with a series of papers about the "Matthew effect in science"[3], contributing empirical evidence to Merton's famous insights, work which is still referred to in the context of scientometrics as well as modeling science.

Science history interested Manfred throughout his life and became his passion again in the last years. He celebrated the importance of Vasili Vassilievich Nalimov[4] for our field, translating the first of his major philosophical works[5] and a biography[6] from Russian into German. He planned to do more investigations into archives about the importance of Klaus Fuchs in the 20th century[7].

Throughout his whole life, Manfred has built bridges: between different countries and cultures, between different scientific areas, and different societal systems – bridges stable enough for travelers in the future. However, making journeys – in geographic and intellectual landscapes – without his inspiring and kind company will be harder and less enjoyable. Let us not forget his scientific and personal heritage!

Berlin—Amsterdam—Dresden
September 2012

1 Bonitz, M. (1991), The impact of behavioral principles on the design of the system of scientific communication, Scientometrics, 20 (1), 107-111.

2 Wolfgang Glänzel (2011), Some thoughts on Manfred Bonitz's two behavioural principles governing research and communication processes. In: The Two Behavioural Principles of Manfred Bonitz - A Festschrift for His 80th Birthday. Ed. by W. Glänzel, Balázs Schlemmer. Special volume of the ISSI e-Newsletter, vol. 25-S March 2011, pp. 11-13. See www.issi-society.info/manfredbonitz/manfredbonitz80.pdf

3 Bonitz, M., Bruckner, E., and A. Scharnhorst (1997) Characteristics and impact of the Matthew effect for countries. Scientometrics 40(3), 407-422.

4 Bonitz, Manfred (Ed.) (2001) Memorial issue of V. V. Nalimov, Scientometrics 52(2), 101-360.

5 Nalimov, V.V. (2009) Spontaneität des Bewusstseins. Trafo Wissenschaftsverlag. Translated from Russian by Manfred Bonitz. Edition V.V. Nalimov (1910-1997) Vol. 1. Ed. by Zh. A. Drogalina-Nalimov and Manfred Bonitz.

6 Zolotukhina-Abolina, E.V. (2010) V.V. Nalimov (1910-1997) Leben und Werk eines russischen Encyclopädisten des XX. Jahrhunderts. Eine Biographie. Trafo Wissenschaftsverlag. Edition V.V. Nalimov (1910-1997) Vol. 0. Ed. by Zh. A. Drogalina-Nalimov and Manfred Bonitz.

7 Bonitz, M. (2005) Klaus Fuchs - the enduring contribution to physics from his British period, Scientometrics 62 (3) 343-350, http://dx.doi.org/10.1007/s11192-005-0026-4

FULL-TIME TENURE-TRACK FACULTY POSITION

THE iSCHOOL AT DREXEL, COLLEGE OF INFORMATION SCIENCE AND TECHNOLOGY

The *iSchool* at Drexel University invites applications for a tenure-track position in bibliometrics/informetrics/scientometrics at the assistant professor level. Applications from senior personnel with a track record of conducting funded research are also encouraged.

We welcome applications from candidates who are eager to join a well-established research program that we have maintained for more than 30 years the successful candidate will become part of a program that involves faculty with interests in both traditional (publication-based) bibliometrics and contemporary webmetrics and altmetrics. Faculty in the program mine, visualize, and analyze large data sets using a wide range of tools—including bibliometrics toolkits and software for social network analysis, information visualization, data mining, trend analysis, and other emerging areas.

The successful candidate will have:

- A completed doctorate in a related field at the time of employment
- Evidence of excellence in teaching and research
- Interest in working with a highly collaborative, interdisciplinary faculty

The *iSchool's* view of information is broad, multidisciplinary, and practical. We currently offer three BS degrees, four Masters degrees, and the PhD. Full-time PhD students are supported either through faculty research grants from a variety of sources or while performing other activities in the College related to research or teaching. The ALA-accredited MS(LIS) degree is highly ranked, both overall and for several of its specializations. Faculty and student interests span a broad spectrum of areas related to information science and technology in contemporary settings. All academic programs emphasize applied research, with a tradition of both quantitative and qualitative approaches that encourages broad collaboration and interdisciplinary activity.

Drexel is a privately endowed technology university founded in 1891. With over 20,000 students, it has one of the largest undergraduate cooperative-education programs in the nation, with formal relationships in place with over 1500 local, national, and multi-national companies. Located in Philadelphia's University City, the campus is at the hub of the academic, cultural, and historical resources of the nation's fifth largest metropolitan region.

To apply for this position, please click here: www.drexeljobs.com/applicants/Central?quickFind=76411, or visit drexeljobs.com and search for position number 5007.

Members of the Search Committee will be available at the ASIST and ALISE annual meetings, iConference, and other venues to discuss open faculty positions.

Drexel University is an Equal Opportunity/Affirmative Action Employer. The *iSchool* at Drexel is especially interested in qualified candidates who can contribute to the diversity and excellence of the academic community.

IFQ – INSTITUTE FOR RESEARCH INFORMATION AND QUALITY ASSURANCE, BERLIN, GERMANY

A PORTRAIT

SYBILLE HINZE
Institute for Research Information and Quality Assurance (iFQ)
Berlin, Germany

In 1999 an international expert commission carrying out an evaluation of the German research system concluded – among other issues – that Germany is lacking a monitoring system that continuously provides information on strength and weaknesses of German research and the policies and instruments put in place to support it. From there it took six years till the foundation of the iFQ, the institution expected to fill this gap. Three objectives lay in front of the novice institution that started off with two and has since grown to approximately 35 staff members, 31 of them with a diverse disciplinary background, even though the social sciences are clearly dominating. These three objectives are: the provision of information on players, programs, projects; the monitoring of research output

and outcome based on well established and newly developed research indicators and the evaluation and assessment of research programs and organizations. The conceptual development of methods, instruments and tools indispensable to fulfill these tasks as well as their technical development and implementation are an integral part of iFQ's activities.

iFQ organizes its activities in five thematic areas:

(i) *Early career researchers and researcher careers* – focusing on the long-term monitoring of researcher careers with specific focus on doctoral candidates and postdoctoral researchers aiming at the identification of factors and conditions impacting researcher careers. For this purpose we developed and implemented a panel survey which is running since 2010 with meanwhile more than 7.100 participants from German universities. This panel approach also allows analyzing whether and how newly introduced measures such as structured doctoral training programs affect researcher careers.

(ii) *Indicators and methods* – the area which for the audience of the ISSI newsletter is probably of most interest. iFQ is coordinating the German centre for bibliometrics (KB), a virtual centre consisting of four partners: iFQ; the Fraunhofer Institute for Systems und Innovation Research (ISI); the Institute for Science and Technology Studies (IWT), University of Bielefeld and the Leibniz Institute for Information Infrastructure (FIZ Karlsruhe). The KB has been building up a quality assured in-house bibliometric data base, with the Web of Science and Scopus as its core components. This tool is now being used for evaluation and monitoring purposes as well as to pursue bibliometric research, e.g. the development of theoretically grounded error calcula, that allow for the assessment of the robustness and interpretative stability of bibliometric indicators. Furthermore, together with the Institute for Science Networking Oldenburg GmbH (ISN), we are developing a natural language processing (NLP) based methodology for field delineation. We provide bibliometric analyses and expertise as integral part of the monitoring of German research funding programs in particular the *excellence initiative*, a funding program for German universities and, together with our partners in the KB also for the *joint initiative for research and innovation*, a funding program implemented to increase the competitiveness of the large German research organizations and. This takes us to the next thematic area

(iii) *Research evaluation* – apart from the already mentioned evaluations of funding programs, the excellence initiative and the joint initiative for research and innovation on behalf of the European Research Council (ERC) we carry out a monitoring of the ERC's starting grants program, one objective being to assess the effect of the program on early researchers' careers and mobility. The second trajectory we are pursuing in this area concerns the re-

flection of existing procedures of quality assurance in science. In particular we are analyzing peer review procedures implemented within the DFG. After focusing on their rather recently introduced system of so-called *Fachkollegiaten* the focus now is on panel peer review and the effects these group reviews have on the review process.

(iv) The analysis of effects and impact of research funding, in particular its systemic effects are the central issue we are addressing in the area of *analyses of the research system*. In 2010 we, for the first time, carried out a large survey among professors at German universities. The major issue addressed was third party funding – the state of play, framework conditions, perceptions, its impact on the researchers and their behavior but also the perceived impact on the research system. It is intended regularly repeat this survey in order to enable trend analysis.

(v) The fifth and last area is *research information*, not literally a thematic area but rather reflecting our task to provide continuous information on the research system and its actors. Here we are attempting to join forces of the four thematic areas just described and make relevant information and data available to the interested public.

iFQ is funded by DFG and collaborates strongly with DFG. But, as already indicated above iFQ carries out research and service projects for a variety of national and international organizations and we do so in close collaboration with a number of partners within and beyond Germany.

Training of researchers is another highly relevant area of iFQ activities. Together with our partners – the University of Leuven, the University of Vienna and Humboldt University we are organizing the European Summer School for Scientometrics – by which we attempt to respond to the lack of a pertinent scientometrics education while at the same time the demand for well trained experts in the field is continuously increasing.

esss 2012 – SCIENTOMETRIC EDUCATION AT KU LEUVEN






CHRISTIAN GUMPENBERGER[1] JUAN GORRAIZ[1] WOLFGANG GLÄNZEL[2] KOENRAAD DEBACKERE[2] STEFAN HORNBOSTEL[3,4] SYBILLE HINZE[3]

[1] University of Vienna, Vienna, Austria
[2] Katholieke Universiteit Leuven, Leuven, Belgium
[3] Institute for Research Information and Quality Assurance (iFQ), Berlin, Germany
[4] Humboldt University of Berlin, Germany

According to the annual rotation idea the third esss (European Summer School for Scientometrics) was hosted by the Katholieke Universiteit (KU) Leuven from 1 -7 July 2012 (1,2). It was again jointly organized by the University of Vienna (Austria), the Humboldt University (Germany), KU Leuven (Belgium) and iFQ (Institute for Research Information and Quality Assurance, Germany).

Like in precedent years, the whole week event was fully booked (50 seats) within only few days after registration opened. The esss conference part was attended by 80 persons (excluding organizers, speakers and several representatives from industry).

While last year's event (3) was six days including a free pre-programme, a conference, seminars, hands-on sessions and a workshop, this year's programme ran seven days. The tutorial day as well as the conference took place in the J. van der Eecken Auditorium at the Faculty of Business and Economics, one of the many buildings KU Leuven. esss seminars and hands-on sessions were then all held at the IT Seminar Rooms of the Van Den Heuvelinstituut.

esss 2012 started officially with a tutorial day on 1 July, which was a bibliometric crash course for esss participants short on experience in the field. The attendees were familiarized with the main terms and concepts of bibliometrics by Wolfgang Glänzel (KU Leuven, Belgium).

His presentation was followed by Thomson Reuters and Elsevier database demonstrations.

Malgorzata Krasowska (Business Development Manager, Thomson Reuters) gave a lively and user-oriented introduction to "Naviga-

tion, search and analysis features of the Web of Knowledge Platform". Then Jan-Albert Majoor (Account Development Manager Elsevier S&T) provided his "SciVerse Scopus: introduction and walk trough demonstration".

esss 2012 was officially opened on 2 July by Koenraad Debackere, who emphasized the importance of bibliometrics and gave a warm welcome to the attendees of the two days esss conference.

Conference day 1 was dedicated to "Theoretical and Practical Aspects" of scientometrics. Nicola De Bellis as the first speaker talked about "The Philosophical Foundations of (Biblio/Sciento/Infor) Metrics" and clearly demonstrated that bibliometrics is certainly not strictly limited to evaluation purposes (Photo 1). Wolfgang Glänzel (KU Leuven, Belgium) and Stefan Hornbostel (iFQ, HU Berlin, Germany) followed providing a concise overview of the "History and Institutionalization of Scientometrics". Sybille Hinze gave a brief introduction to "Scientometric indicators in use: an overview". For the first time in the esss Wolfgang Glänzel talked about the "Mathematical Foundation of Scientometrics", followed by the presentation of Bart Thijs about "The application of Network analysis in science studies: Common theoretical background for broad applications".

Another novelty was "The Bibliometric Agora: Open discussion of controversial issues" (Photo 2), featuring Marc Luwel (Hercules Foundation, Belgium) and Wolfgang Glänzel in a lively, most interesting dispute, which was brilliantly moderated by Stefan Hornbostel. Valérie Thiel (SciVal Consultant, Elsevier) concluded conference day 1 with her presentation "The SciVal Suite - Helping Institutions to Establish, Evaluate and Execute Research Strategies".

The overall topic of conference day 2 (3 July) was "Procedures and Indicators: State-of-the-Art". Donald de B. Beaver (Williams College, USA) set the stage and introduced "The many faces of collaboration and teamwork in scientific research".

Then Anthony van Raan (Leiden University, the Netherlands) talked about "Advanced bibliometric methods for evaluation of research groups, ranking and benchmarking of universities, and mapping of research related to socio-economic issue" in his usual entertaining way. Éric Archambault (Science-Metrix, Canada) presented "Journal-level classifications - current state of the art" followed by Michel Zitt (INRA, LERECO Lab, Research Department SAE2, France) who introduced "Meso-level retrieval: field delineation and hybrid methods".

After the lunch break Henk Moed (Senior Scientific Advisor, Elsevier) spoke about "New developments in bibliometrics and research assessment" (Photo 3). Patricia Bren-

nan (Director Product Strategy, Thomson Reuters) presented the new "Book Citation Index" and discussed its limitations.

After the coffee break Koenraad Debackere (Katholieke Universiteit (K.U.) Leuven) grabbed the audience's attention by talking about "Scientometrics, Technometrics, Innovation: Policy Aspects". The final presentation of conference day 2 was dedicated to the visualization trend in the discipline. Katy Börner (Royal Netherlands Academy of Arts and Sciences (KNAW), The Netherlands / Cyberinfrastructure for Network Science Center, Director, School of Library and Information Science, Indiana University, Bloomington, IN) gave a detailed description of the science mapping and visualization tool SCI2 and informed about its wide range of functionalities, many of them just recently added.

Day four to seven were all about the true nature of a summer school: seminars, hands-on sessions, group work and presentations all taking place in one big computer room. Like last year a maximum of 50 people could participate (Photo 4).

4 July was dedicated to "Data sources, Cleaning & Impact Measures". The day started with Matthias Winterhager (Bielefeld University,

Institute of Science and Technology Studies (IWT), Germany) who discussed the highly relevant topic of "Data cleaning and Processing". After that Wolfgang Glänzel and Juan Gorraiz introduced alternative and complementary data sources for bibliometric purposes and discussed the various journal impact measures. The imparted theory was then consolidated in practical hands-on exercises in the afternoon.

The focus topics of 5 July were "Mapping Science" on the one hand and "Network Analysis" on the other. De Solla Price Medal winner Olle Persson (Sociology Department, Umeå universitet, Sweden) started the day with "Mapping Science (on the basis of Bibexcel Software)" . In the afternoon Bülent Özel (Department of Computer Science, Istanbul Bilgi University (IBU), Turkey) continued with presenting "Network Analysis with R".

The workshop "Research Evaluation in Practice" was scheduled for 6 July. After an introduction for all by Wolfgang Glänzel, the participants were asked to work on specific tasks in small groups. Two topics were given (1. Bibliometric Research and Application, 2. Adolescent and maternal smoking). Like in the year before the participants had to determine the most important players, the most frequent journals and subject categories in the field and the topics within the field and their distribution. Furthermore they should determine basic indicators of citation impact and analyse the h-score. Then an international collaboration analysis should be performed and the most productive countries as well as the main partners identified. Finally publication activity and citation impact of the most productive countries in the field should be determined. The students were encouraged to also visualize the results using the mapping tools introduced in the preceding days.

Group work was individually supported by Wolfgang Glänzel, Sybille Hinze, Juan Gorraiz and Bart Thijs.

7 July was the final day of the esss course and intended as a Q & A session. The organizers were happy to answer many questions of the participants and officially closed the event shortly before lunchtime.

esss 2012 was again very successful with participants from 24 countries (Austria, Belgium, Brazil, Colombia, Czech Republic, Estonia, France, Germany, Greece, Iran, Japan, Latvia, Lithuania, Netherlands, Norway, Poland, Russia, South Korea, Spain, Sweden, Switzerland, Turkey, United Kingdom, United States). The overall positive feedback is a clear signal to continue this international initiative, and the organizers are looking forward to next year's event held at the Humboldt University Berlin in Germany in September. Further announcements will be made via esss website (www.scientometrics-school.eu) and via the esss mailing list (to register please send an informal email to office@scientometrics-school.eu).

REFERENCES

Gorraiz J., Gumpenberger C., Hornbostel S., Hinze, S., Glänzel, W. & Debackere, K. (2010). European Summer School for Scientometrics (esss) to be launched. Scientometrics 83(2), 601-602. DOI: 10.1007/s11192-010-0206-8

Gorraiz J., Gumpenberger C., Glänzel, W., Debackere, K., Hornbostel S., Hinze, S. (2011). esss 2010: A review of the inaugurational European Summer School for Scientometrics in Berlin. Scientometrics 86 (1), 235-236. DOI: 10.1007/s11192-010-0279-4

Gumpenberger C., Gorraiz J., Glänzel, W., Debackere, K., Hornbostel S., Hinze, S. (2012). Event report: esss 2011-Scientometric education in Indian summer at the University of Vienna. Scientometrics 91 (1), 311-313. DOI: 10.1007/s11192-011-0578-4

VI. INTERNATIONAL SEMINAR ON THE QUANTITATIVE AND QUALITATIVE STUDY OF SCIENCE AND TECHNOLOGY "PROF. GILBERTO SOTOLONGO AGUILAR"

MEETING REPORT

18-20 APRIL, 2012 – HAVANA, CUBA

DR. MARÍA VICTORIA GUZMÁN
mvguzman [at] finlay.edu.cu

DR. CÉSAR A. MACÍAS CHAPULA
chapula [at] data.net.mx

MID ISIDRO AGUILLO HonPhD
isidro.aguillo [at] cchs.csic.es

DR. JANE M. RUSSELL
jrussell [at] unam.mx

INTRODUCTION

One of the main objectives of the Seminar which is a permanent feature of the biennial Cuban Information Congress (INFO) is to provide a forum to debate, exchange and discuss issues related to metric studies (Bibliometrics, Scientometrics, Informetrics, Webmetrics, etc.). An overriding purpose has been the widespread adoption of metric studies by the region's young specialists and the diversification of applications within the different spheres of scientific and technological development in

Latin America. The speakers and delegates were a mix of researchers from information science and management from both industry and academia.

The sixth edition comprised 56 studies from a diverse range of countries (Austria, Belgium/Hungary, Chile, Cuba, France, Mexico, Spain and Uruguay), 29 were oral presentations and 27 were posters. Several sessions focused on the evaluation of the performance of researchers, institutions, countries and subject areas, while others specialised in topics related to Latin America and the Caribbean, patentometrics, webmetrics and health issues due to a number of contributions on the application of metric studies to the biomedical field

OPENING SESSION

The Opening Lecture on the first day of the Seminar was an overview of bibliometric methods for detecting and analysing emerging research topics given by Professor Wolfgang Glänzel of the Katholieke Universiteit Leuven (Belgium) and the Hungarian Academy of Sciences, Institute for Research Policy Studies (Budapest, Hungary) which Jane Russell (Mexico) chaired. A combined text- and citation-based (clustering) hybrid approach was proposed to analyse the underlying structures within which emerging topics are detected. According to the speaker the key phases of the bibliometric analysis of emerging topics require both structural and dynamic analyses of the discipline under investigation. The study presented was carried out in the field of nanopollution and results clearly indicated that emerging topics in this field are no longer the prerogative of research from the industrialised world.

Following the questions and comments of the audience, Dr. Glänzel concluded by saying that no one particular method exists (when referring to methodological considerations) nor is there an ideal tool (when referring to software) for carrying out metric studies. Their use will depend upon the particular requirements of the study at hand.

A second keynote address was given by Anna María Prat from Chile who for 25 years was the head of the Information Department of the Chilean National Commission for Scientific and Technological Research (CONICYT). The speaker reviewed the experiences acquired from bibliometric studies from the perspective of a public institution responsible for science policy such as CONICYT. In this setting the development and growth of bibliometrics was closely associated with the necessity and demands of science policy decision making in the public research sector as well as those of government projects on scientific research and technological development. The Chilean experience began with the construction of a single database which brought together the entire scientific production of the country which was normalized and updated weekly.

During the discussion which was chaired by Maria Victoria Guzmán (Cuba), attention was directed towards the need of countries from Latin America and the Caribbean to develop their own databases to provide reliable information for carrying out metric studies which present an objective picture of regional science to complement and contextualize analyses from international databases. Questions from the audience related to the use and general application of manuals and methodological applications (Manual of Bogota and Manual of Santiago) developed by the Ibero-American and Inter-American Network for Science and Technology Indicators (RICYT) of which the speaker is a member. Given these and other concerns, the conclusion reached was the need to create and make extensive suitable tools for the analysis of problems in science and technology in the region.

SCIENTIFIC SESSIONS

In the sessions that followed the keynote speakers, two new indicators were proposed to evaluate individual scientific performance. The first presented by Yaniris Rodríguez (Cuba) and termed IBAC (Bibliometric Index of Scientific Activity) is a multifactorial bibliometric indicator which takes into consideration the following factors: production in different scientific journals, citations received and variation in the h2 index. According to the authors, quantitative comparisons can be established between researchers working in the same disciplinary area and qualitative comparisons between researchers in different disciplinary areas.

The second indicator put forward by a team of researchers from Mexico and Cuba and presented by Magdalena Sierra, proposed a bibliometric measurement for the autoevaluation of scientific articles designed to answer the following questions: How can a researcher self-evaluate his or her scientific production? How can evaluators and decision makers assess the impact of individual researchers within a given research field? Based on the evaluation of citations given to an article by making a comparison with those received by the articles included in the reference list, the method positions the article under evaluation within the citation distribution of the field to which the research reported belongs. The indicator was tested in three different astrophysical fields.

A new strategy to extract negative results from secondary information sources was presented by Ivana Roche (Austria-France) involving semantic text analysis essential for arriving at a fair interpretation of the significance of the negative assertions present in the article. On the same lines, Ricardo Arenciba (Cuba) described a structural-analytical technique to identify macrostructures in biomedical knowledge domains based on the analysis of conglomerates.

Maria Victoria Guzmán (Cuba) applied ViBlioSOM (Bibliometric Visualisation using

the SOM algorithm) methodology to study the use of diverse metric indicators for the analysis of translational medicine. Elio Villaseñor (Mexico - Cuba) presented a dynamic analysis of the relevance of MeSH terms and their application in the visualisation of translations in biomedical knowledge domains. The methodology combines temporal clustering recognition with the analysis of the relevance of nodes in complex networks. By analysing distinctive topics, for example using MeSH terms, the goal is the automatic recognition of pattern changes within a particular domain which represent a shift in the interests of the international community. The usefulness of this method was illustrated by means of its application to the evolution of scientific research into tuberculosis vaccines between 1950 and 2011. The method is an adjunct to ViBlioSOM methodology which exploits the visualisation capacity of SOM neuronal networks to analyse the evolution of knowledge domains using bibliographic records retrieved from PubMed.

Also presented were applications on the scientific production in Cuban public health by Grisel Zacca (Cuba –Spain) and from Mexico, Elena Luna demonstrated results on the contribution to mainstream science of an international agricultural research centre situated in Mexico (CIMMYT International Maize and Wheat Improvement Centre) while Francisco Collazo analysed the production and flow of Mexican scientific knowledge in the first half of the 20th century. The tendencies in the scientific output of five Latin American countries were presented by Ricardo Araujo (Cuba) as part of the INASLP (International Network for the Availability of Scientific Information) project. A look at the contribution of Latin America particularly in the social sciences and humanities to the recently available Book Citation Index was presented by Shirley Ainsworth (Mexico) comparing the representation in this new index with that in the citation and proceedings indexes also produced by Thomson Reuters. Maylín Frías (Cuba) proposed for the first time indicators to analyse Cuban cinematographic production. A construction of "personal scientific schools" based on research and collaborative relationships was presented by Roelvis Ortiz (Cuba).

Webmetrics was present by way of the application of an informatics tool, GEWEB, whose use was proposed by Yonny Mondelo (Cuba) while Sergio Carbonell (Cuba) described the behaviour with respect to the scientific production of the semantic web from 2005-2009.

The Seminar programme included various studies oriented towards the analysis of patent indicators such as the one presented by Maidelyn Díaz (Cuba –Spain) on the technological productivity of Iberoamerica from 2003 to 2009. Three studies were from the proGINTEC group from the University of Pinar del Río (Cuba). Raudel Giráldez (Cuba – Mexico) looked at the transference of scientific knowledge to the technology of tuberculosis vaccines by analysing the bibliographic references to scientific articles present in patents. The MeSH terms assigned to these articles were then correlated with the patent topics associated with the international classification of patents. The applications presented by this group used diverse visualisation algorithms employed in

Photo copyright © Zsuzsanna Glänzel

specific softwares. Some of the studies were carried out using network algorithms such as those of Pajek, UCINET and NodeXL. Examples were also shown of visual interfaces obtained via Viscovery SOMine which is based on Kohonen SOM (Self-Organizing-maps). The proINTEC software developed at the University of Pinar specialized in patents was demonstrated. Theoretical aspects of metric visualisation were reviewed by Deborah Torres (Cuba) from the viewpoint of the Information Sciences.

POSTER SESSION

A feature of the poster session of the Seminar is that individual presentations are given a specific time slot and that these are chaired in groups in a way similar to the oral sessions. This year the session was chaired by Ivana Roche (France), Juan Carlos Araujo (Cuba), Shirley Ainsworth and Elio Villaseñor (both of Mexico). Fundamental topics presented were related to the visibility and impact of scientific production. In addition, Celso Musiño (Mexico) proposed a methodology for carrying out information studies in the informal economy sector. Colleagues from the University of Informatic Sciences (Cuba) explained the development of the SEOMaster tool which provides statistical web indicators such as the relevance of certain search criteria "key words or phrases".

DISCUSSION AND PROPOSALS

This was the sixth of the biennial seminars and there is no doubt that these meetings have contributed greatly to the formation of a Latin American community of specialists in metric and other studies of science and technology. The seminars have provided a setting for the free exchange and discussion of experiences and ideas and have led to the conformation of a network of professionals in the field. New collaborative projects have emerged as a result of contacts made through the network as well as publications in coauthorship. An example of a bilateral collaboration between Cuban and Mexican institutions is the project on the development and application of analytical and data visualisation techniques in the biomedical field.

At the end of the seminar participants were invited to express their views and ideas on the content and organization of the Seminar which can be summarised as follows:

1. Dr. Francisco Collazo's name was put forward as a fifth member of the organizing committee of the Seminar. Francisco has given presentations in many if not all, of the previous meetings and has been involved recently in the submission review process.
2. It was suggested that satellite seminars should be held in other parts of the region to achieve greater participation of professionals from all countries in Latin America.
3. The need was expressed to increase awareness and promotion of the Seminar for which the present web site needs to be improved and other information media taken advantage of.

Our immediate goal is to achieve a more balanced representation of countries in the region for INFO 2014.

PREDICTING THE H-INDEX

RONALD ROUSSEAU
1) Faculty of Engineering Technology, KHBO, Oostende, Belgium
2) Department of Mathematics, KU Leuven, Heverlee (Leuven), Belgium
3) IOIW (IBW), University of Antwerp, Antwerp, Belgium
ronald.rousseau[at]khbo.be

XIAOJUN HU
Medical Information Centre, Zhejiang University School of Medicine, Hangzhou, China
xjhu[at]zju.edu.cn

Abstract. We discuss a recent article in Nature about predicting future h-index values.

INTRODUCTION

When Hirsch introduced his indicator (Hirsch, 2005) it immediately led to a flurry of discussions on the h-index, followed by new applications (Braun et al., 2005) and a host of derived h-type indicators, such as the g-index and the R-index (Egghe, 2006b; Jin et al., 2007; Alonso et al., 2009; Egghe, 2010). Also the *ISSI Newsletter* contributed leading to highly cited articles by Egghe (2006a), Kosmulski (2006), Bar-Ilan (2006) and Jin (2007).

Later, Hirsch (2007) showed that the h-index was a better predictor of future scientific achievement than the total number of citations, total number of papers and number of citations per paper, where scientific achievement was defined as cumulative number of citations.

THE NEW ARTICLE

Now some colleagues (from the life sciences) published a method to predict future h-indices (Acuna et al., 2012). As it is published in *Nature* it immediately attracted a lot of attention, leading to a number of comments in *The Scientist* (Richards, 2012). Using data on 3085 neuroscientists, augmented with 57 *Drosophila* researchers and 151 evolutionary scientists they were able to find a formula, using machine-learning techniques, to predict the h-index for next year, 5 year into the future and 10-year into the future, of course with decreasing accuracy.

Although they started with 18 variables it turned out that only 5 were actually significant. Table 1 shows the 18 variables; the

VARIABLES
1. Number of articles in top journals (q)
2. Number of articles in "theoretical" journals
3. Number of years in a postdoctoral position
4. Number of years in graduate school
5. Number of articles with an adviser as a co-author
6. Number of R-type grants
7. Total current yearly cost
8. Current h-index divided by career length (= m-index)
9. Mean h-index divided by career length of PhD supervisor
10. Proportion of articles as last author
11. Proportion of articles as first author
12. Total number of citations
13. Total number of articles (n)
14. Average number of co-authors per article
15. Total number of distinct journals in which articles are published (j)
16. Average number of citations per article
17. Career length defined as number of years since publishing first article (y)
18. Current h-index (h)

Table 1. Variables taken into account for predicting future h-indices

5 significant ones are in bold. For the exact definition of these variables we refer the reader to the original article's supplementary information.

The authors used so-called elastic net regularization (Zou & Hastie, 2005) to take multicollinearity into account and force some coefficients to become zero. This finally led to an equation of the form:

$$h(future) = c_0 + c_1\sqrt{n} + c_2 h - c_3 y + c_4 j + c_5 q$$

where the c_j, j = 0...5 are positive constants depending on the prediction horizon and determined by the authors' algorithm. They showed that their prediction of a future h-index is much better than using only the current h-index (which is no surprise).

DISCUSSION

Many objections can be raised against this approach the most important one being that, as fields differ a lot, the obtained formula is probably only valid for neuroscientists. In a comment Hirsch thinks that the inclusion of the total number of used journals is spurious (Richards, 2012). He claims that publishing in many journals is a sign of shopping around when an article has been rejected by a top journal. We do not agree with this explanation as the number of distinct journals used is also a form of diffusion of knowledge (diffusion by publications), and a sign of the breadth of knowledge of an author (Liu & Rousseau, 2010). Of course, this observation has nothing to do with the possibly that indeed this variable is introduced by a spurious correlation.

Surely, using the number of publications in so-called top journals (*Nature*, *Science*, *Nature Neuroscience*, *PNAS* and *Neuron*) is only valid in a small group of fields (actually just one, because of the inclusion of the journals *Nature Neuroscience* and *Neuron*), and adds to the fetishism surrounding *Science* and *Nature* (giving

authors able to publish in them mythical qualities). Moreover, publications in conference proceedings or in books were not even considered as variables.

Given the problems surrounding the h-index, e.g. related to its non-consistency, one may wonder why one would want to predict an imperfect indicator. Finally, the definition of the h-index as given by Acuna et al. (2012) is imprecise. They write that a scientist has h-index n if he or she has published n articles receiving at least n citations each. They should have added that it is the largest natural number with this property.

Surely, knowing precisely the variables (research indicators) that really matter in a field would be of great help to peer review committees but we do not think this has been achieved by this publication.

ACKNOWLEDGEMENT

This research is supported by National Natural Science Foundation of China (NSFC Grant Nos. 71173185).

REFERENCES

Acuna, D.E., Allesina, S. & Körding, K.P. (2012). Predicting scientific success. *Nature*, 489, 201-202.

Alonso, S., Cabrerizo, F. J., Herrera-Viedma, E. & Herrera, F. (2009). h-Index: A review focused in its variants, computation and standardization for different scientific fields. *Journal of Informetrics*, 3, 273–289.

Bar-Ilan, J. (2006). H-index for Price medalists revisited. *ISSI Newsletter*, 2, 3-5.

Braun, T., Glänzel, W. & Schubert, S. (2005). A Hirsch-type index for journals. *The Scientist*, 19(22), p. 8-8.

Egghe, L. (2006a). An improvement of the H-index: the G-index. *ISSI Newsletter*, 2, 8-9.

Egghe, L. (2006b). Theory and practise of the G-index. *Scientometrics*, 69, 131-152.

Egghe, L. (2010). The Hirsch index and related impact measures. *Annual Review of Information Science and Technology*, Volume 44 (B. Cronin, ed.), 65-114, Information Today Inc., Medford, (NJ), USA.

Hirsch, J. E. (2005). An index to quantify an individual's scientific research output. *Proceedings of the National Academy of Sciences of the USA*, 102(46), 16569-16572.

Hirsch, J. E. (2007). Does the H index have predictive power? *Proceedings of the National Academy of Sciences of the United States of America*, 104, 19193-19198.

Jin, BH. (2007). The AR-index: complementing the h-index. *ISSI Newsletter*, 3(1), 6-6.

Jin, BH., Liang, LM., Rousseau, R., & Egghe, L. (2007). The R- and AR-indices: complementing the H-index. *Chinese Science Bulletin*, 52, 855-863.

Kosmulski, M. (2006). A new Hirsch-type index saves time and works equally well as the original H-index. *ISSI Newsletter*, 2, 4-6.

Liu, YX., & Rousseau, R. (2010). Knowledge diffusion through publications and citations: a case study using ESI-fields as unit of diffusion. *Journal of the American Society for Information Science and Technology*, 61, 340-351.

Richards, S. (2012). Predicting publishing futures. *The Scientist*, September 12, 2012.

Zou, H. & Hastie, J. (2005). Regularization and variable selection via the elastic net. *Journal of the Royal Statistical Society B*, 67(2), 301-320.

DEAR SIR,

ANDRÁS SCHUBERT
Department of Science Policy and Scientometrics,
Library of the Hungarian Academy of Sciences
Budapest, Hungary

I am observing with some astonishment and growing embarrassment the fervent statements in the RoA vs AoR debate. I might have missed some point, but I feel both the actions and the reactions somewhat exaggerated, all the more, since in my opinion the two supposedly contradictory approaches are, in fact, special cases of the same general procedure.

The problem is as follows: there is a set of items, each characterized by an additive (in physical terms, extensive) measure (in the classical example publications are characterized by their citation count). For each item we have a measured ("actual") value (let us denote it with *X*-s) and also an estimated ("expected") one (the *Y*-s). The quotient of the two, Q = X / Y, is a kind of performance ratio, the relative measure of the actual value as compared to the expected one. So far, so good.

Having an aggregated set of items, one may wonder about the overall performance ratio, *Q,* of the whole set. The debate is whether the ratio of the average actual and expected values are to be formed (ratio of averages, RoA), or the individual ratios should be formed first and their average is to be taken (average of ratios, AoR).

I do not want to recapitulate the arguments and counter-arguments here, and do not want to qualify them, either. I only want to call the readers' attention to a simple and fundamental relation.

In the "real world", outside the realm of pure mathematics, the proper way of averaging is most infrequently the simple arithmetic averaging, since the items to be averaged (e.g. measurement points) are not necessarily equivalent, they may have different weights. Generally speaking, the weights go parallel with the importance or the reliability of the measurements.

The averaging procedure suggested by the AoR advocates is the simple unweighted arithmetic averaging. *Q* of the aggregate is calculated as

$$Q = (\Sigma\, Q_i) / N = (\Sigma\, X_i/Y_i) / N,$$

where lower index *i* denotes the values concerning the individual items and *N* is the number of items in the aggregate.

If weighting is introduced, the formula is modified to

$$Q = (\Sigma\, w_i Q_i) / (\Sigma\, w_i),$$

where w_i denotes the weight of the *i*-th item. Obviously, for the arithmetical average, each $w_i = 1$.

Let us consider now the case when the weights are the Y_i-s themselves. This choice, though first may seem haphazard, is not quite unreasonable, indeed. It attributes greater importance to items of greater expected value (i.e., papers published in higher impact factor journals) and depreciates (in extreme case, nullify) those published in journals of insignificant, almost zero impact factor.

The overall performance measure will now be

$$Q = (\Sigma\, Y_i Q_i) / (\Sigma\, Y_i) =$$

$$= (\Sigma\, X_i) / (\Sigma\, Y_i) =$$

$$= (\Sigma\, X_i / N) / (\Sigma\, Y_i / N),$$

i.e., we have exactly the ratio of averages, the AoR advocates struggle against.

The question is, therefore, not so much about the choice between two alternatives, but about the choice of a proper weighting scheme. Like in all similar cases in the sciences or social sciences, practically never exists an all-purpose optimal weighting scheme; optimal solutions can be searched for only for definite purposes. It is the applier's subtle task to select and justify a proper weighting scheme, whether unit weights (AoR), Y_i-s (RoA) or any other choice is considered.

From the interpretation given here it can be seen that the unweighted scheme overemphasizes items with low expected value (at least as compared with the AoR version); in extreme cases, when some Y_i-s approaches zero, the overall *Q* may diverge. By using Y_i-s as weights this problem is solved, but this is an arbitrary choice, as well; the decision about the suitability – let alone, optimality – of which requires careful consideration. Reciting Woody Allen: *"One path leads to despair and utter hopelessness, the other to total extinction. Let us pray that we have the wisdom to choose correctly."*