The role of bibliometrics in research management ...
It’s complicated!

Dr. Thed van Leeuwen
Center for Science & Technology Studies (CWTS)

TU Delft seminar on “Are you ready to publish?”, February 5th 2015
Outline

• Bibliometrics and research management context
• Infamous bibliometric indicators
• Scientific integrity and misconduct
• Applicability of bibliometrics
• Take-home messages …
Bibliometrics and the research management context
What is bibliometrics?

- Bibliometrics can be defined as the quantitative analysis of science and technology (development), and the study of cognitive and organizational structures in science and technology.

- Basic for these analyses is the scientific communication between scientists through (mainly) journal publications.

- Key concepts in bibliometrics are output and impact, as measured through publications and citations.

- Important starting point in bibliometrics: scientists express, through citations in their scientific publications, a certain degree of influence of others on their own work.

- By large scale quantification, citations indicate (inter)national influence or (inter)national visibility of scientific activity, but should not be interpreted as synonym for ‘quality’.
‘Classical’ image of the Credibility cycle

Credibility cycle (adapted from Latour and Woolgar (1979) & Rip (1990))
Rise of performance indicators & bibliometrics

Need for formalised measures

- ‘Push’ from science policy (from 1970s onwards)
- Independent of peer review
- New Public Management / Neo-liberalism (from 1980s onwards)

Matrix-structure science (Whitley)

- Researchers part of international community (Peer review)
- But also part of local institutions (Specific management practices, e.g. appraisals, external evaluations)
- Institute managers not always part of international expert community
- Tighter forms of management (from the 1990s onwards)

→ Distance
Extended credibility cycle
Infamous bibliometric indicators: JIF & H-index
Definitions of Journal Impact Factor & Hirsch Index

• **Definition of JIF:**
  – The mean citation score of a journal, determined by dividing all citations in year T by all citable documents in years T-1 and T-2.

• **Definition of h-index:**
  – The ‘impact’ of a researcher, determined by the number of received citations of an oeuvre, sorted by descending order, where the number of received citations on that single paper equals the rank position.
Problems with JIF

• Methodological issues
  – Was/is calculated erroneously  (Moed & van Leeuwen, 1996)
  – Not field normalized
  – Not document type normalized
  – Underlying citation distributions are highly skewed  (Seglen, 1994)

• Conceptual/general issues
  – Inflation  (van Leeuwen & Moed, 2002)
  – Availability promotes journal publishing
  – Is based on expected values only
  – Stimulates one-indicator thinking
  – Ignores other scholarly virtues
Deconstructing the myth of the JIF…

- Take the Dutch output
- Similar journal impact classes
- Focus on publications that belong to the top 10% of their field
Problems with H-index

• **Bibliometric-mathematical issues**
  – mathematically inconsistent  (Waltman & van Eck, 2012)
  – conservative
  – Not field normalized  (van Leeuwen, 2008)

• **Bibliometric-methodological issues**
  – How to define an author?
  – In which bibliographic/metric environment?

• **Conceptual/general issues**
  – Favors age, experience, and high productivity  (Costas & Bordons, 2006)
  – No relationship with research quality
  – Ignores other elements of scholarly activity
  – Promotes one-indicator thinking
The problem of fields and h-index ...

- Spinoza candidates across all domains ...
- Use output, normalized impact, and h-index
In what database context …

Selected my own publications in WoS and Scopus, Google Scholar has a pre-set profile.

<table>
<thead>
<tr>
<th>Database</th>
<th>H-index</th>
<th>Based upon …</th>
</tr>
</thead>
<tbody>
<tr>
<td>Web of Science</td>
<td>14</td>
<td>Articles in journals</td>
</tr>
<tr>
<td>Scopus</td>
<td>25</td>
<td>Articles, book (chapters), and conference proceedings papers</td>
</tr>
<tr>
<td>Google Scholar</td>
<td>32</td>
<td>All types, incl. Reports</td>
</tr>
</tbody>
</table>
Scientific integrity
and misconduct
Scientific integrity and misconduct

• An often mentioned cause for any kind of misconduct is the “Publish or Perish” culture.

• As such, that is supposed to force individuals to produce more and more output, ....

• … as these are considered the building blocks of an academic career !

• However, pressure is on all of us !
Effects of the “Publish or perish” culture: the case of J.H. Schön
Effects of the “Publish or perish” culture: the case of D. Stapel

D. Stapel caused the Netherlands to be the country ranked first in the field of Social psychology, when it comes to retractions!
Applicability of bibliometrics
Output of a university ... Output, and the way it is represented in Web of Science!

Output, and the wide variety not covered in Web of Science!
Take home messages
Take-home messages on journals

• Journals tend to publish positive/confirming results.
• Editorial boards are driven by market shares as well!
• Therefore, selection is harsh, and rejection rates are high

Take-home messages on data

• Data are not frequently published.
• Therefore, they do not give any credits for the producers!
• This keeps most scientific work non transparent
• Databases with negative results are necessary
Take-home messages on bibliometrics

- Ask yourself the question “What do I want to measure?”
- And also “Can that be measured?”
- Take care of proper data collection procedures.
- Then, always use actual and expected citation data.
- Apply various normalization procedures (field, document, age)
- Always have a variety of indicators.
- Always include various elements of scholarly activity.
- And perhaps most important, include peer review in your assessment procedures!!!
Thank you for your attention!

Any questions?
Ask me, or mail me
leeuwen@cwts.nl
Development of authorship across all domains of scholarly activity
Coverage issues
<table>
<thead>
<tr>
<th>AU</th>
<th>Moed, HF; Garfield, E.</th>
<th>in WOS</th>
</tr>
</thead>
<tbody>
<tr>
<td>TI</td>
<td>In basic science the percentage of 'authoritative' references decreases as bibliographies become shorter</td>
<td></td>
</tr>
<tr>
<td>SO</td>
<td><strong>SCIENTOMETRICS</strong> 60 (3): 295-303, 2004</td>
<td>Y</td>
</tr>
<tr>
<td>RF</td>
<td>ABT HA, <em>J AM SOC INF SC T</em>, v 53, p 1106, 2004</td>
<td>Y</td>
</tr>
<tr>
<td></td>
<td>GARFIELD, <em>CITATION INDEXING</em>, 1979 (BOOK!)</td>
<td>N</td>
</tr>
<tr>
<td></td>
<td>GILBERT GN, <em>SOC STUDIES SCI</em>, v 7, p 113, 1977</td>
<td>Y</td>
</tr>
<tr>
<td></td>
<td>ROUSSEAU R, <em>SCIENTOMETRICS</em>, v 43, p 63, 1998</td>
<td>Y</td>
</tr>
<tr>
<td></td>
<td>ZUCKERMAN H, <em>SCIENTOMETRICS</em>, v 12, p 329, 1987</td>
<td>Y</td>
</tr>
</tbody>
</table>

**WoS Coverage** = 5/7 = 71%
WoS Coverage in 2010 across disciplines

- **Black** = Excellent coverage (>80%)
- **Blue** = Good coverage (between 60-80%)
- **Green** = Moderate coverage (but above 50%)
- **Orange** = Moderate coverage (below 50%, but above 40%)
- **Red** = Poor coverage (highly problematic, below 40%)
Some clear ‘perversions’ of the system ... ?

- “You call me, I call you”
- When time is passing by ...
- Salami slicing to boost an academic career
- Multiple authorship (without serious contributing)
- Putting your name on everything your unit produces
- The role of self citations
- Jumping on hypes and fashionable issues