

DER WISSENSCHAFTSRAT BERÄT DIE BUNDESREGIERUNG  
UND DIE REGIERUNGEN DER LÄNDER IN FRAGEN  
DER INHALTLICHEN UND STRUKTURELLEN ENTWICKLUNG DER  
HOCHSCHULEN, DER WISSENSCHAFT UND DER FORSCHUNG.

BERLIN

10 09 2013

# The Role of Indicators in Informed Peer Review. Practical Observations

## Main messages:

1. Effects on objectivity, validity, reliability small.
2. Indicators can help reduce the reviewers' workload.
3. There is a tradeoff between sophistication and usefulness.

## Overview

1. Context: The Wissenschaftsrat. The Research Rating
2. What is informed peer review?
3. Observations on the use of bibliometric indicators
4. Conclusions

# Wissenschaftsrat – German Council of Science and Humanities

The German Council of Science and Humanities (Wissenschaftsrat) provides advice to the German federal government and the state (Länder) governments on the structure and development of higher education and research.

Chairman

Secretary General

Plenary Assembly

Scientific Commission  
32 Votes  
given by  
24 Scientists  
8 Personalities from  
Public Life

Administrative Commission  
32 Votes  
given by  
16 States (*Länder*)  
1 Vote per State  
16 Federal Government

Head Office  
5 units with a staff of  
about 80  
of which about 45  
are scientists

SEND MEMBERS

Committees and Working Groups  
(complemented by external experts)  
Prepare recommendations and reports

# Research rating

## Goals

- \_ Analyse institutional strengths and weaknesses in order to support heads of institutions and governing bodies
- \_ Supply valid information on research performance, improve on customary rankings
- \_ Get complete picture by including non-university research
- \_ Assist research communities in reflecting upon research quality, strengths and weaknesses

# Research rating

## Characteristics

- Multi-dimensional
- Rating, not ranking
- Informed peer review
- Method adjusted to disciplinary practices, by the peers
- Quality assessment on research unit level

# Assessment criteria

E.g. Chemistry

## Dimension

Research

Promotion of young researchers

Knowledge transfer

## Criterion

I. Research Quality

II. Impact/Effectiveness

III. Efficiency

IV. Promotion of young researchers

V. Transfer to other societal agents

VI. Wissensvermittlung und Verbreitung



## Four pilot studies completed 2005 - 2012

<b>Subject</b>	<b>Publication date</b>	<b>No. institutions</b>	<b>No. professors</b>
Chemistry	2008	77	1,038
Sociology	2008	57	376
Electrical Engineering	2011	47	737
English Studies	2012	60	356

## Overview

1. Context: The Wissenschaftsrat. The Research Rating
2. What is informed peer review?
3. Observations on the use of bibliometric indicators
4. Conclusions

## Informed peer review

Evaluation of research performance **on the basis of standardised data reports.**

Goal: Assessment according to pre-defined criteria and pre-defined scale of marks.

Contrasts with: peer review of publications; reputation survey.

## Research rating – the process

1. Definition and operationalisation of criteria
2. Data collection, verification, analysis
3. Reports and Guidelines for reviewers compiled
4. Plenary discussion of exemplary reports
5. Individual assessments (2 – 3 reviewers per institution)
6. Plenary discussion of proposed assessments with focus on (a) divergent opinions, and (b) extreme marks
7. Additional data / additional reviewers in case of doubt
8. Final vote on assessments

## Research rating – the data

5 year window

Questionnaire based data collection combined with re-analysis of various existing databases

Qualitative and quantitative data combined, e.g.:

- \_ Self-evaluation of institutions
- \_ Selected publications
- \_ Lists of publications, dissertation projects etc.
- \_ Personnel, funding etc.
- \_ Bibliometric indicators

## Research rating – the data reports

Data reports for each individual institution include

- \_ All the data collected for that specific institution
- \_ Structured according to the set of assessment criteria
- \_ Percentiles for quantitative data added

Guidelines for reviewers include

- \_ Definitions of criteria
- \_ Description of methods of data collection and analysis
- \_ Statistical information on the whole institutional population under review; examples where appropriate

## Overview

1. Context: The Wissenschaftsrat. The Research Rating
2. What is informed peer review?
3. Observations on the use of bibliometric indicators
4. Conclusions

## Which bibliometrical indicators were used?

	Chemistry	Electrical Engineering
Research Quality	cpp cpp/FCS <sub>m</sub> cpp/JCS <sub>m</sub> JCS <sub>m</sub> /FCS <sub>m</sub> C <sub>max</sub>	cpp cpp/JCS <sub>m</sub> C <sub>max</sub> %Top10
Impact	c C <sub>max</sub> p <sub>non-c</sub> /p	c C <sub>max</sub> %Top10
Efficiency	c/#scientists	-/-
Database	WoS	Scopus incl. Non-Source
Publications	Complete	Selected (5 / Prof)



## First observations on indicator use

Reviewers judged indicators as very helpful

Many comments on absolute citation numbers during the reviewing sessions

Small number of unexpected results that induced changes in preliminary assessments (positive or negative)

\_ Recent achievements vs. reputation earned

\_ Quality of work vs. size of field

Very small number of mistakes in WoS detected

Objectivity & validity increased?

## Correlations with chemistry rating

Criterion	Indicator	$r^2$
Research Quality	cpp/FCS <sub>m</sub>	0.762
	cpp	0.686
	JCS <sub>m</sub> /FCS <sub>m</sub>	0.619
Impact	c	0.866
	p	0.773
Efficiency	c/#scientists	0.821

For chemistry, models can be constructed that predict reviewers' assessment with  $r^2 \approx 0.75$

## Inter-reviewer reliability

Inter-reviewer reliability was rather high: 80 % agreement or near-agreement among reviewers for research quality in chemistry

Lower inter-reviewer reliability for other criteria such as transfer, but still high (59 – 74 %)

Inter-reviewer reliability *re* research quality cannot be traced to use of indicators: equally high in sociology (71 %), English studies (80 %), where no bibliometric indicators were used

Alternative explanation: Discussions on criteria and indicators amount to intense reviewer training

## Reviewer effort

Difficult to measure, based on reviewers' estimates

Subject	Workload per reviewer
Chemistry	4 – 5 weeks
Sociology	8 – 10 weeks
Electrical Engineering	4 – 5 weeks
English Studies	12 – 16 weeks

Partial explanation: subject-specific degree of task-uncertainty?

But: Sociology, English Studies decided to read and review selected publications (90 – 110 per reviewer)

## Overview

1. Context: The Wissenschaftsrat. The Research Rating
2. What is informed peer review?
3. Observations on the use of bibliometric indicators
4. **Conclusions**

1. Effects on objectivity, validity, reliability small.
2. Indicators can help reduce the reviewers' workload.
3. There is a tradeoff between sophistication and usefulness.

## Sophistication vs. usefulness

Positive effects of informed peer review on objectivity and validity require transparency of indicator design.

Reviewers use raw data to:

- \_ Identify publications with  $c_{\max}$ , assess influence on mean;
- \_ Assess effect and adequacy of field normalisation;
- \_ Correct mismatches between publications and research units, document types.

Risk: home-spun bibliometrics.

- \_ Unstandardised use of  $h$  index

For informed peer review,  
keep bibliometrics simple!



Thank you for your attention!

[lange@wissenschaftsrat.de](mailto:lange@wissenschaftsrat.de)