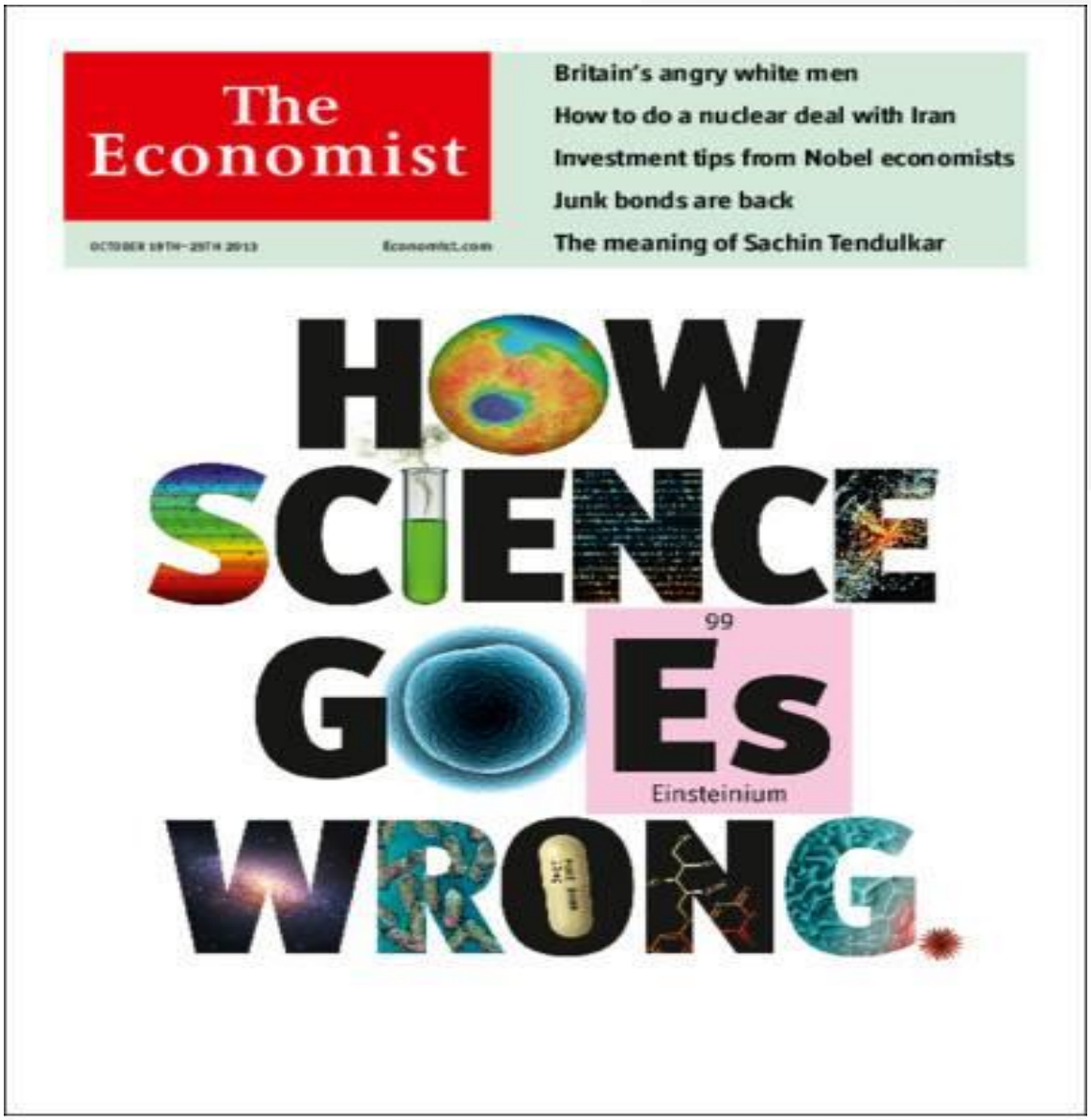
CWTS
Meaningful metrics

Competing for Excellence: Perverse and constructive uses of evaluation machines in academia

Impact and KE Seminar Series
Nuffield College, University of Oxford
3 June 2015

Paul Wouters

Universiteit Leiden

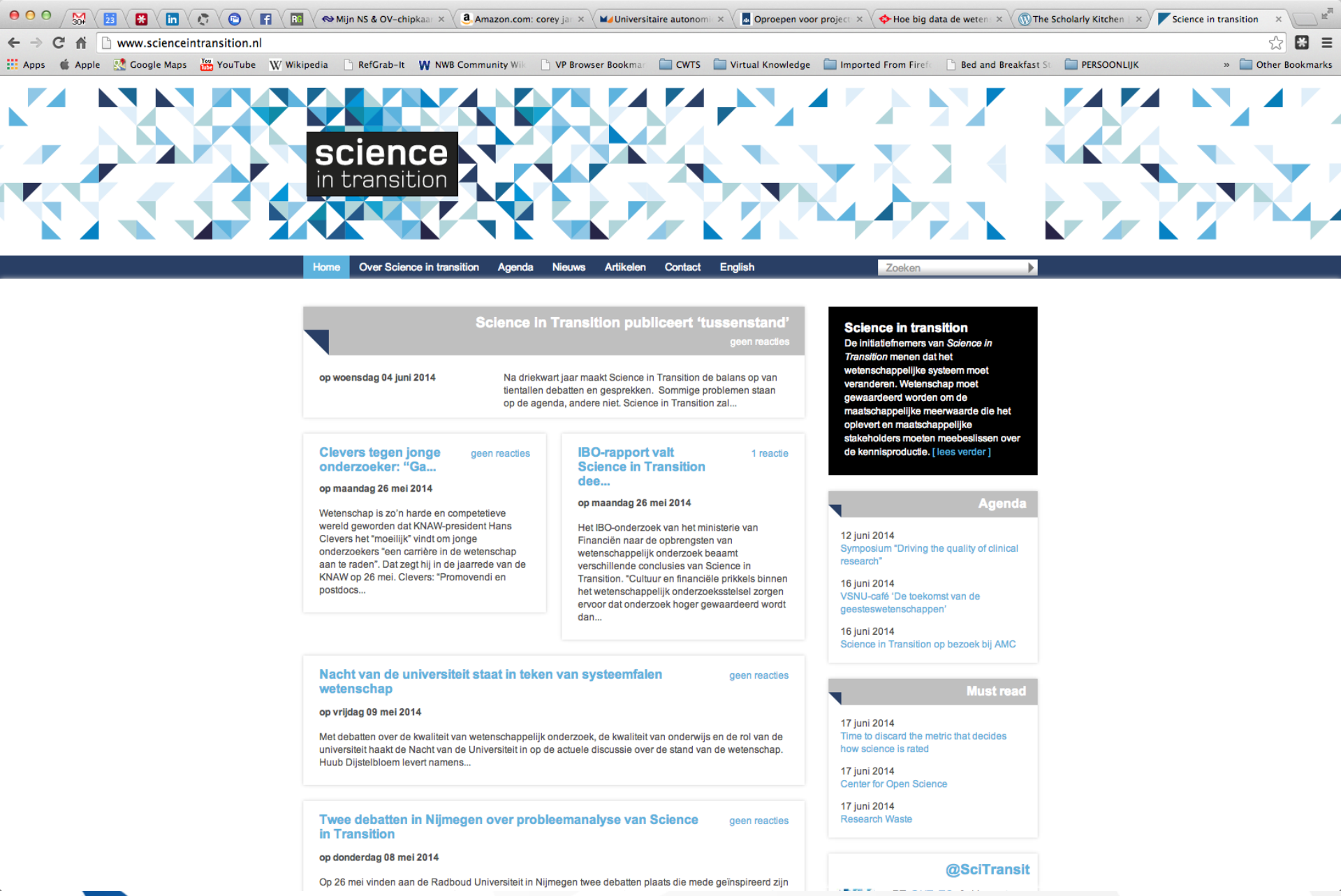


OCTOBER 19TH - 25TH 2013 Worldwide cover

CWTS

A SIMPLE idea underpins science: “trust, but verify”. Results should always be subject to challenge from experiment. That simple but powerful idea has generated a vast body of knowledge. Since its birth in the 17th century, modern science has changed the world beyond recognition, and overwhelmingly for the better. But success can breed complacency. Modern scientists are doing too much trusting and not enough verifying—to the detriment of the whole of science, and of humanity.

Too many of the findings that fill the academic ether are the result of shoddy experiments or poor analysis (see article (http://www.economist.com/news/briefing/21588057-scientists-think-science-self-correcting-alarming-degree-it-not-trouble)). A rule of thumb among biotechnology venture-capitalists is that half of published research cannot be replicated.


www.scienceintransition.nl
science in transition
Home
Over Science in transition
Agenda
Nieuws
Artikelen
Contact
English
Zoeken
Science in Transition publiceert 'tussenstand'
geen reacties
op woensdag 04 juni 2014
Na driekwart jaar maakt Science in Transition de balans op van tientallen debatten en gesprekken. Sommige problemen staan op de agenda, andere niet. Science in Transition zal...
Clevers tegen jonge onderzoeker: "Ga...
geen reacties
op maandag 26 mei 2014
Wetenschap is zo'n harde en competetieve wereld geworden dat KNAW-president Hans Clevers het "moeilijk" vindt om jonge onderzoekers "een carrière in de wetenschap aan te raden". Dat zegt hij in de jaarrede van de KNAW op 26 mei. Clevers: "Promovendi en postdocs...
IBO-rapport valt Science in Transition dee...
1 reactie
op maandag 26 mei 2014
Het IBO-onderzoek van het ministerie van Financiën naar de opbrengsten van wetenschappelijk onderzoek beaamt verschillende conclusies van Science in Transition. "Cultuur en financiële prikkels binnen het wetenschappelijk onderzoeksstelsel zorgen ervoor dat onderzoek hoger gewaardeerd wordt dan...
Nacht van de universiteit staat in teken van systeemfalen wetenschap
geen reacties
op vrijdag 09 mei 2014
Met debatten over de kwaliteit van wetenschappelijk onderzoek, de kwaliteit van onderwijs en de rol van de universiteit haakt de Nacht van de Universiteit in op de actuele discussie over de stand van de wetenschap. Huub Dijstelbloem levert namens...
Twee debatten in Nijmegen over probleemanalyse van Science in Transition
geen reacties
op donderdag 08 mei 2014
Op 26 mei vinden aan de Radboud Universiteit in Nijmegen twee debatten plaats die mede geïnspireerd zijn
Science in transition
De initiatiefnemers van Science in Transition menen dat het wetenschappelijke systeem moet veranderen. Wetenschap moet gewaardeerd worden om de maatschappelijke meerwaarde die het oplevert en maatschappelijke stakeholders moeten meebeslissen over de kennisproductie. [lees verder]
Agenda
12 juni 2014
Symposium "Driving the quality of clinical research"
16 juni 2014
VSNU-café 'De toekomst van de geesteswetenschappen'
16 juni 2014
Science in Transition op bezoek bij AMC
Must read
17 juni 2014
Time to discard the metric that decides how science is rated
17 juni 2014
Center for Open Science
17 juni 2014
Research Waste
@SciTransit

CWTS

Mendeley Desktop

Pan Highlight Note Select Copy Paste Rotate Zoom Fullscreen Sync

varmus

My Library | "Opening Up" and "Closi... | Rescuing US biomedical ...

Showing match 1 of 1 Previous Next

PERSPECTIVE

Rescuing US biomedical research from its systemic flaws

Bruce Alberts[a], Marc W. Kirschner[b], Shirley Tilghman[c,1], and Harold Varmus[d]

[a]Department of Biophysics and Biochemistry, University of California, San Francisco, CA 94158; [b]Department of Systems Biology, Harvard Medical School, Boston, MA 02115; [c]Department of Molecular Biology, Princeton University, Princeton, NJ 08540; and [d]National Cancer Institute, Bethesda, MD 20892

Edited by Inder M. Verma, The Salk Institute for Biological Studies, La Jolla, CA, and approved March 18, 2014 (received for review March 7, 2014)

The long-held but erroneous assumption of never-ending rapid growth in biomedical science has created an unsustainable hypercompetitive system that is discouraging even the most outstanding prospective students from entering our profession—and making it difficult for seasoned investigators to produce their best work. This is a recipe for long-term decline, and the problems cannot be solved with simplistic approaches. Instead, it is time to confront the dangers at hand and rethink some fundamental features of the US biomedical research ecosystem.

graduate education | postdoctoral education | federal funding | peer review

By many measures, the biological and medical sciences are in a golden age. That fact, which we celebrate, makes it all the more difficult to acknowledge that the current system contains systemic flaws that are threatening its future. A central flaw is the long-held assumption that the enterprise will constantly expand. As a result, there is now a severe imbalance between the dollars available for research and the still-growing scientific community in the United States. This imbalance has created a hypercompetitive atmosphere in which scientific productivity is reduced and promising careers

DNA sequencing, sophisticated imaging, structural biology, designer chemistry, and computational biology—has led to impressive advances in medicine and fueled a vibrant pharmaceutical and biotechnology sector.

In the context of such progress, it is remarkable that even the most successful scientists and most promising trainees are increasingly pessimistic about the future of their chosen career. Based on extensive observations and discussions, we believe that these concerns are justified and that the biomedical research enterprise in the United States is on an unsustainable

doubling of the NIH budget ended, the demands for research dollars grew much faster than the supply. The demands were fueled in large part by incentives for institutional expansion, by the rapid growth of the scientific workforce, and by rising costs of research. Further slowdowns in federal funding, caused by the Great Recession of 2008 and by the budget sequestration that followed in 2013, have significantly exacerbated the problem. (Today, the resources available to the NIH are estimated to be at least 25% less in constant dollars than they were in 2003.) The consequences of this im-

Page 1 of 5

Details Notes Contents Enrichments

Authors: B. Alberts, M. Kirschner, S. Tilghman et al.

View research catalog entry for this paper

Journal: *Proceedings of the National Academy of Sciences of the United States of America*

Year: 2014

Volume: 111

Issue: 16

Pages: 5773-7

Abstract:

The long-held but erroneous assumption of never-ending rapid growth in biomedical science has created an unsustainable hypercompetitive system that is discouraging even the most outstanding prospective students from entering our profession--and making it difficult for seasoned investigators to produce their best work. This is a recipe for long-term decline, and the problems cannot be solved with simplistic approaches. Instead, it is time to confront the dangers at hand and rethink some fundamental features of the US biomedical research ecosystem.

Tags:

Author Keywords:

Biomedical Research; Biomedical Research: economics; Biomedical Research: manpower; Biomedical Research: organization & administration; Health Planning Guidelines; Humans; Research Personnel; Research Personnel: supply & distribution; Research Support as T...

URL:

http://www.ncbi.nlm.nih.gov/pubmed/24733905

Add URL...

Catalog IDs

ArXiv ID:

DOI: 10.1073/pnas.1404402111

ISBN: 1091-6490 (Electronic)\r0027-8424 (Linking)

ISSN: 1091-6490

PMID: 24733905

Files:

Alberts et al. - 2014 - Rescuing US biomedical rese...

Add File...

Other Settings

Unpublished work - exclude from Mendeley Web catalog

CWTS

- A severe imbalance between the dollars available for research and the still-growing scientific community in the United States.
- The training pipe-line produces more scientists than relevant positions in academia, government, and the private sector are capable of absorbing
- Hyper-competition for the resources and positions that are required to conduct science suppresses the creativity, cooperation, risk-taking, and original thinking required to make fundamental discoveries.
- Overvaluing translational research is detracting from an equivalent appreciation of fundamental research of broad applicability
- As competition for jobs and promotions increases, the inflated value given to publishing in a small number of so-called “high impact” journals has put pressure on authors to rush into print, cut corners, exaggerate their findings, and overstate the significance of their work.
- Today, time for reflection is a disappearing luxury for the scientific community.
- The quality of evaluation has declined

Research leaders face key questions

- How should we monitor our research?
- How can we profile ourselves to attract the right students and staff?
- How should we divide funds?
- What is our scientific and societal impact?
- What is actually our area of expertise?
- How is our research trans-disciplinary connected?

CWTS

Research leaders need more, not less, strategic intelligence

- Increasing demand for information about research:
  - hyper competition for funding
  - globalization
  - industry – academic partnerships
  - interdisciplinary research challenges
  - institutional demands on research & university management
- Increased supply of data about research:
  - web based research
  - deluge of data producing machines and sensors
  - increased social scale of research: international teams
  - large scale databases of publications, data, and applications

CWTS

Four main problems in current academic research

- The funding system
- The career structure
- The publication system
- The evaluation system

Funding system

- level of funding
- balance between project and infrastructure funding
- balance between blue sky and focused funding
- relationship research and teaching
- 1 size 4 all?

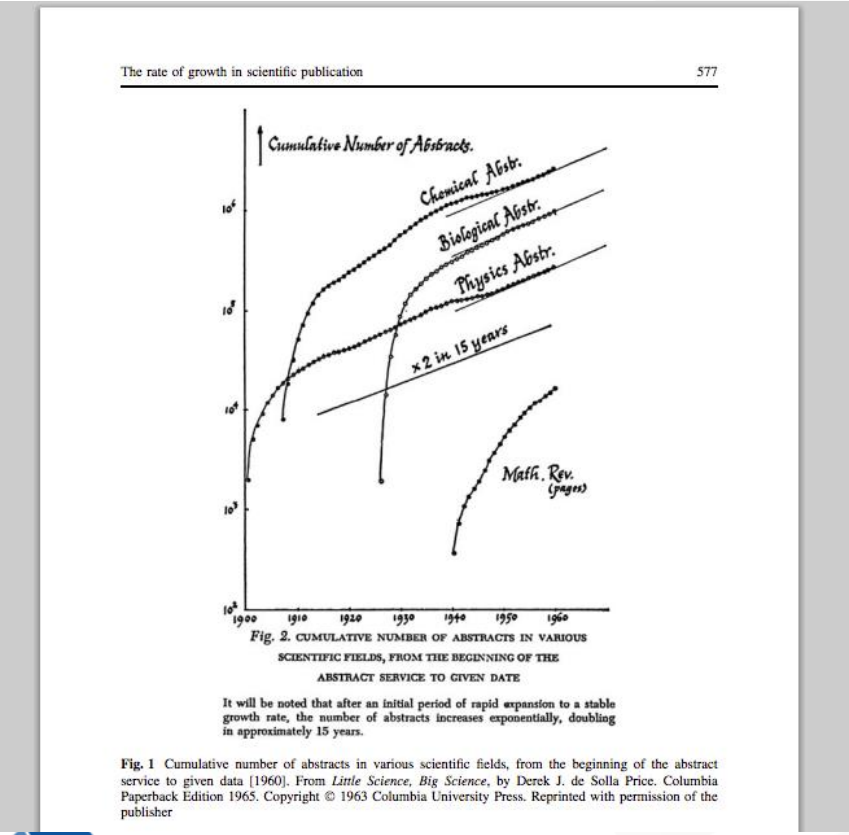
Career structure

- PhDs and postdocs as cheap labour
- hyper-competition
- mismatch training and job opportunities
- lack of dual careers
- emerging separation between researchers and teachers
- increasing inequalities
- lack of diversity in workforce (this may be improving)

CWTS

Growth of scientific literature

Price (1963)



The rate of growth in scientific publication 577

Cumulative Number of Abstracts.

Chemical Abstr.

Biological Abstr.

Physics Abstr.

×2 in 15 years

Math. Rev. (pages)

10^6 10^5 10^4 10^3 10^2

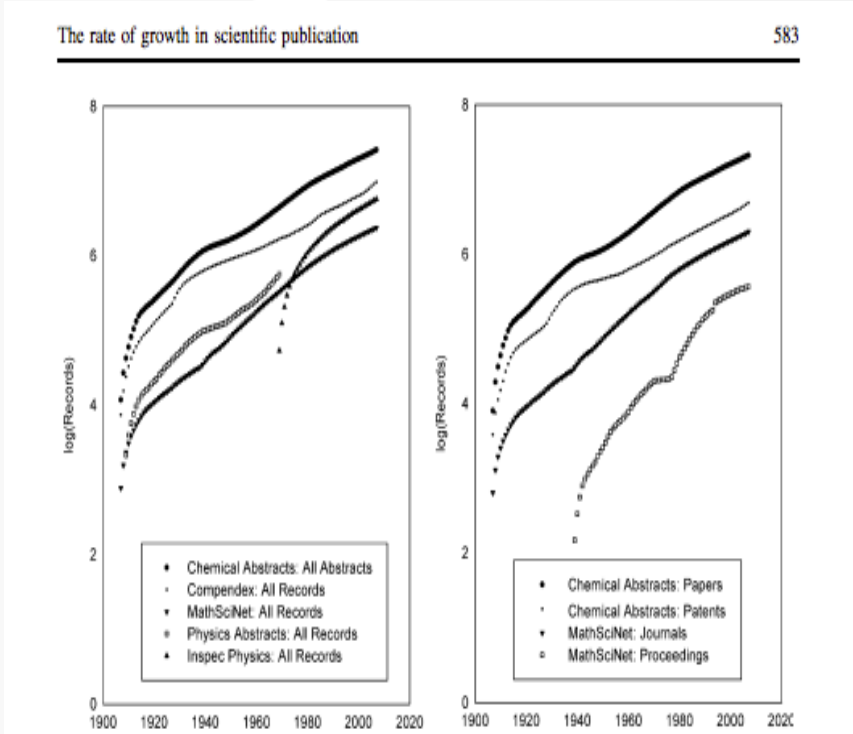
1900 1910 1920 1930 1940 1950 1960

Fig. 2. CUMULATIVE NUMBER OF ABSTRACTS IN VARIOUS SCIENTIFIC FIELDS, FROM THE BEGINNING OF THE ABSTRACT SERVICE TO GIVEN DATE

It will be noted that after an initial period of rapid expansion to a stable growth rate, the number of abstracts increases exponentially, doubling in approximately 15 years.

Fig. 1 Cumulative number of abstracts in various scientific fields, from the beginning of the abstract service to given data [1960]. From *Little Science, Big Science*, by Derek J. de Solla Price. Columbia Paperback Edition 1965. Copyright © 1963 Columbia University Press. Reprinted with permission of the publisher

Larsen & Von Ins (2010)



The rate of growth in scientific publication 583

log(Records)

0 2 4 6 8

1900 1920 1940 1960 1980 2000 2020

Chemical Abstracts: All Abstracts
Compendex: All Records
MathSciNet: All Records
Physics Abstracts: All Records
Inspec Physics: All Records

Chemical Abstracts: Papers
Chemical Abstracts: Patents
MathSciNet: Journals
MathSciNet: Proceedings

Fig. 2 Cumulative number of records for nine databases 1907–2007 (semi logarithmic scale)

CWTS

Codification by publication

- The publication system is the basis for communication, teaching and codification
- Hence, all evaluation systems in science and scholarship are in the end based on publications
- Commercial interests have been able to use the publication system as source of vast profits
- Publishing for the smallest audience possible?
- Evaluation systems have developed on the basis of information (systems) of these publications:
  - peer review in various formats
  - scientometrics and bibliometrics

CWTS

Evaluation Gap

- discrepancy between evaluation criteria and the social and economic functions of science
- evaluation methods (esp. qualitative) have not adapted to increased scale of research
- available quantitative measures are often not applicable at the individual level
- lack of recognition for new types of work that researchers need to perform

CWTS

Diagnosis

Research

The research at CWTS is organized into three chairs for full professors and has five working groups on key research themes.

Chairs

Scientometrics

Science & innovation studies

Science policy studies

Working groups

Advanced bibliometric methods

Evaluation practices in context

Societal impact of research

Career policy & paths in science

Social sciences & humanities

Questions

- How does the evaluation of scientific or scholarly quality affect the creation of knowledge?
  - Which concept of “quality” can be used to understand this interaction?
  - Which concept of “science” or “knowledge” should we use?

Knowledge as infrastructure

- Infrastructures are not constructed but evolve
- Transparent structures taken for granted
- Supported by invisible work
- They embody technical and social standards

(Edwards, A Vast Machine, 2010)

CWTS

Mushroom growth of evaluation

- Relatively recent phenomenon (since mid 1970s)
- Formal evaluation protocols: performance indicators all over the place *but* citation indicators hardly visible
- Science policy studies tend to underestimate the proliferation and impact of indicator based evaluations
- Recent studies focus on performance based funding
- “Anecdotal evidence” shows the proliferation of especially the Hirsch Index and the JIF

CWTS

New trends in assessment

- Increased bibliometric services at university level available through databases
- Increased self-assessment via “gratis bibliometrics” on the web (h-index; publish or perish; etc.)
- Emergence of altmetrics
- Increased demand for bibliometrics at the level of the individual researcher
- Societal impact measurements required
- Career advice – where to publish?

CWTS

Peter Dahler-Larsen *The Evaluation Society*

- *"Evaluations are not something that the individual can reject"*

- Evaluation as disembedded reflexive social practice

- Evaluation consists of:
  - Evaluand
  - Criteria
  - Systematic methodology
  - Purpose

CWTS

Evaluations are liminal

> One often has the feeling that there should have been a clear-cut plan for the purpose and process of an evaluation, but this is often not the case. (…) people realize too late that they had very different notions of plans for evaluation (…) The purpose of the evaluation constitutes an ongoing controversy rather than a common logical starting point.

(p. 15)

CWTS

A good question

If we have only imprecise or invalid indicators available, how can one be so sure that it is better to evaluate than not to evaluate?

CWTS

Evaluation in organization theory

- The rational organization
- The learning organization
- The institutionalized organization

- Clear goals
- Predictable
- Learning cycle
- Positive feedback
- Ritualistic
- Legitimacy
- Power struggle

CWTS

Evaluation Machines

- Primary function: make stuff auditable
- Mechanization of control – degradation of work and trust? (performance paradox)
- Risks for evaluand and defensive responses
- What are their costs, direct and indirect?
- Microquality versus macroquality – lock-in
- Goal displacement & strategic behaviour

CWTS

Constitutive effects

- Limitations of conventional critiques (eg ‘perverse or unintended effects’)
- Effects:
    - Interpretative frames
    - Content & priorities
    - Social identities & relations (labelling)
    - Spread over time and levels
- Not a deterministic process
- Democratic role of evaluations

CWTS

Effects of indicators

- Intended effect: behavioural change
- Unintended effects:
  - Goal displacement
  - Structural changes
- The big unknown: effects on knowledge?
- Institutional rearrangements
- Does quality go up or down?

CWTS

Responses scientific community

- Strategic behaviour
- Ambivalence
- Sophisticated understanding of indicators and citation numbers
- Responses vary by discipline, style, position (Hargens and Schuman 1990)
- “Self-interest” not a valid explanation

CWTS

Citation theories

Theories about citing behaviour

- How do researchers decide what to cite?

- What can be inferred from patterns of citation?

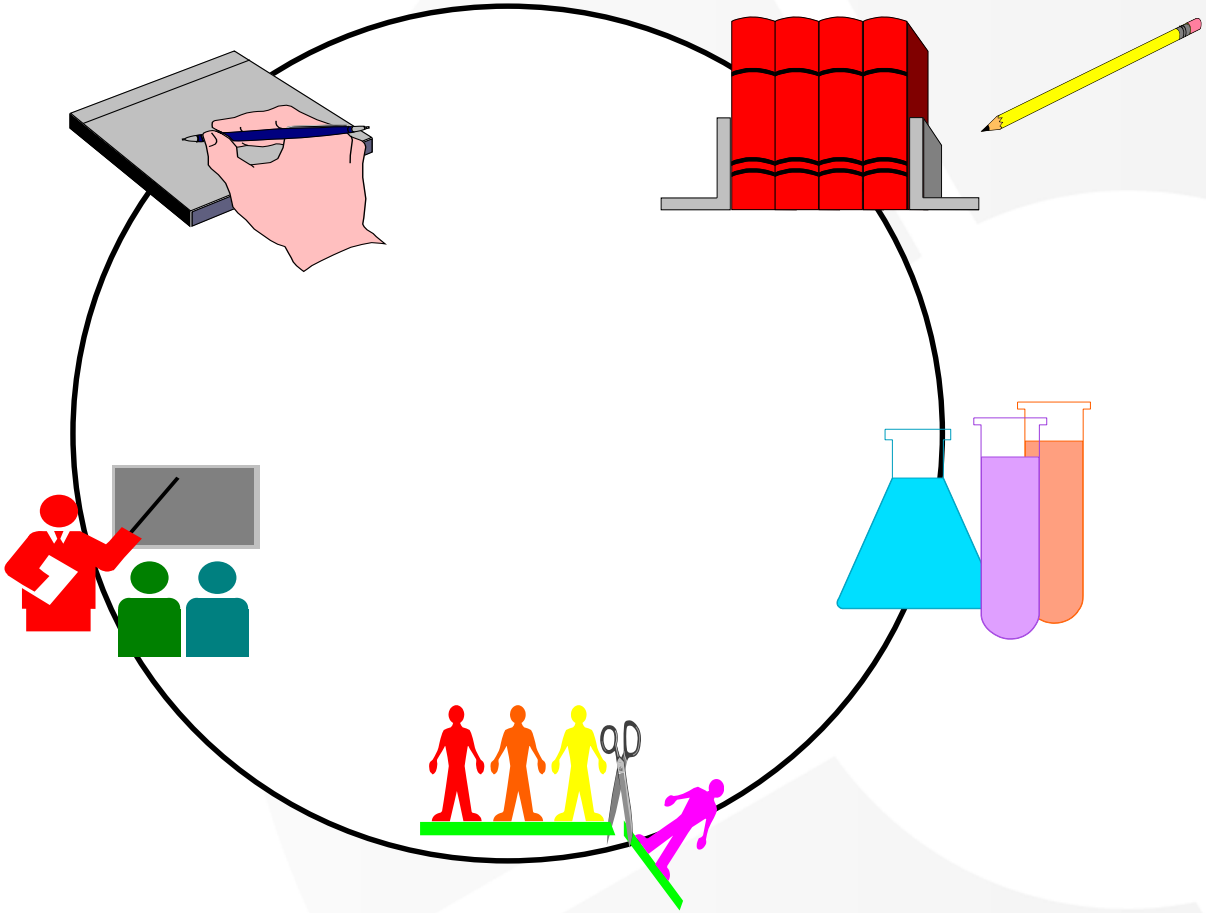
CWTS

Indicator theories

- Van der Veer Martens (2001, online): 'holy grail' in scientometrics" is the development of indicator theories rather than the development of theories of citing behavior
- Two different approaches:
  - Semantic studies: Small (1978): co-citations as concept markers
  - Semiotic studies:
    - Wouters (1998; 1999): analyze citations as sign systems
    - Further extended by Cronin (2000)
    - Wouters (2014, 2015) develops an argument for material semiotics as framework

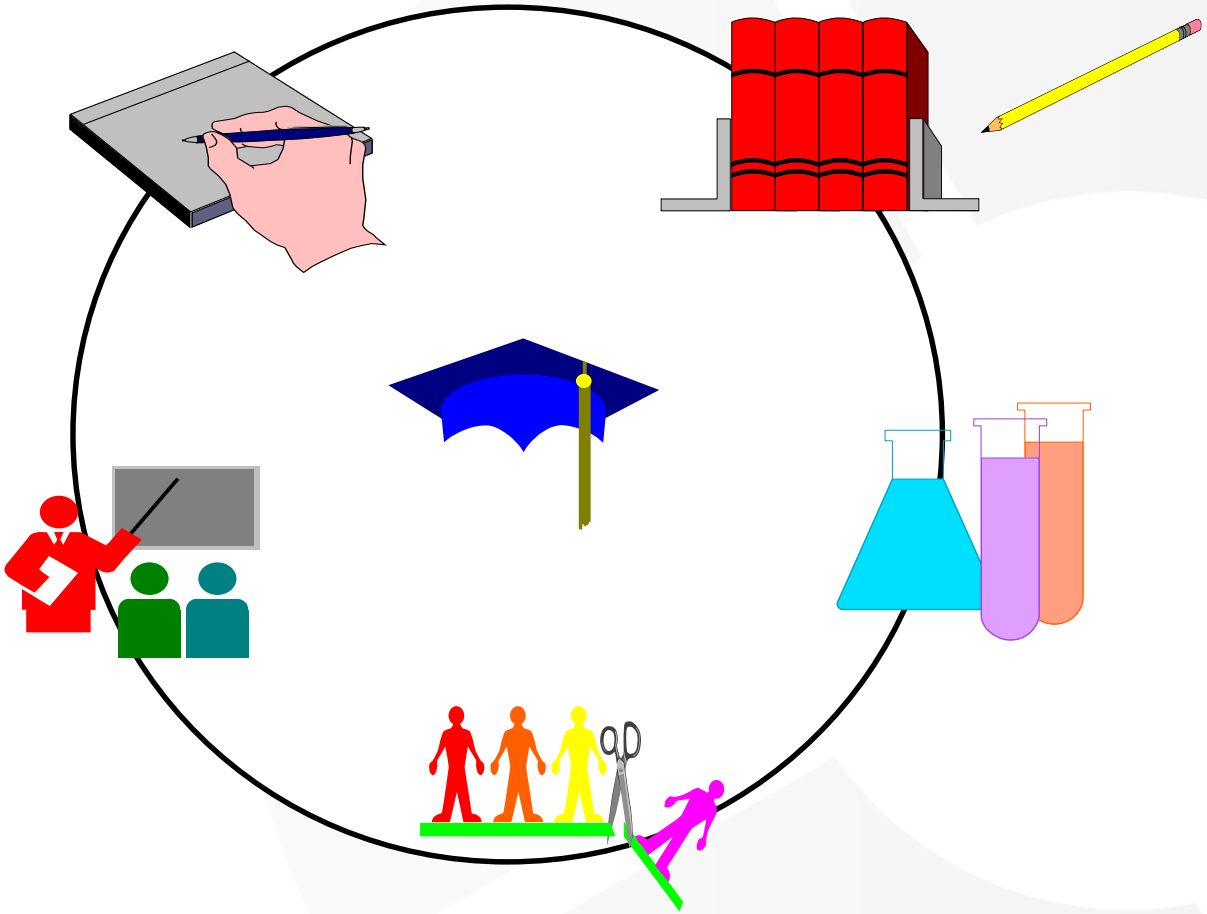
CWTS

The publication cycle



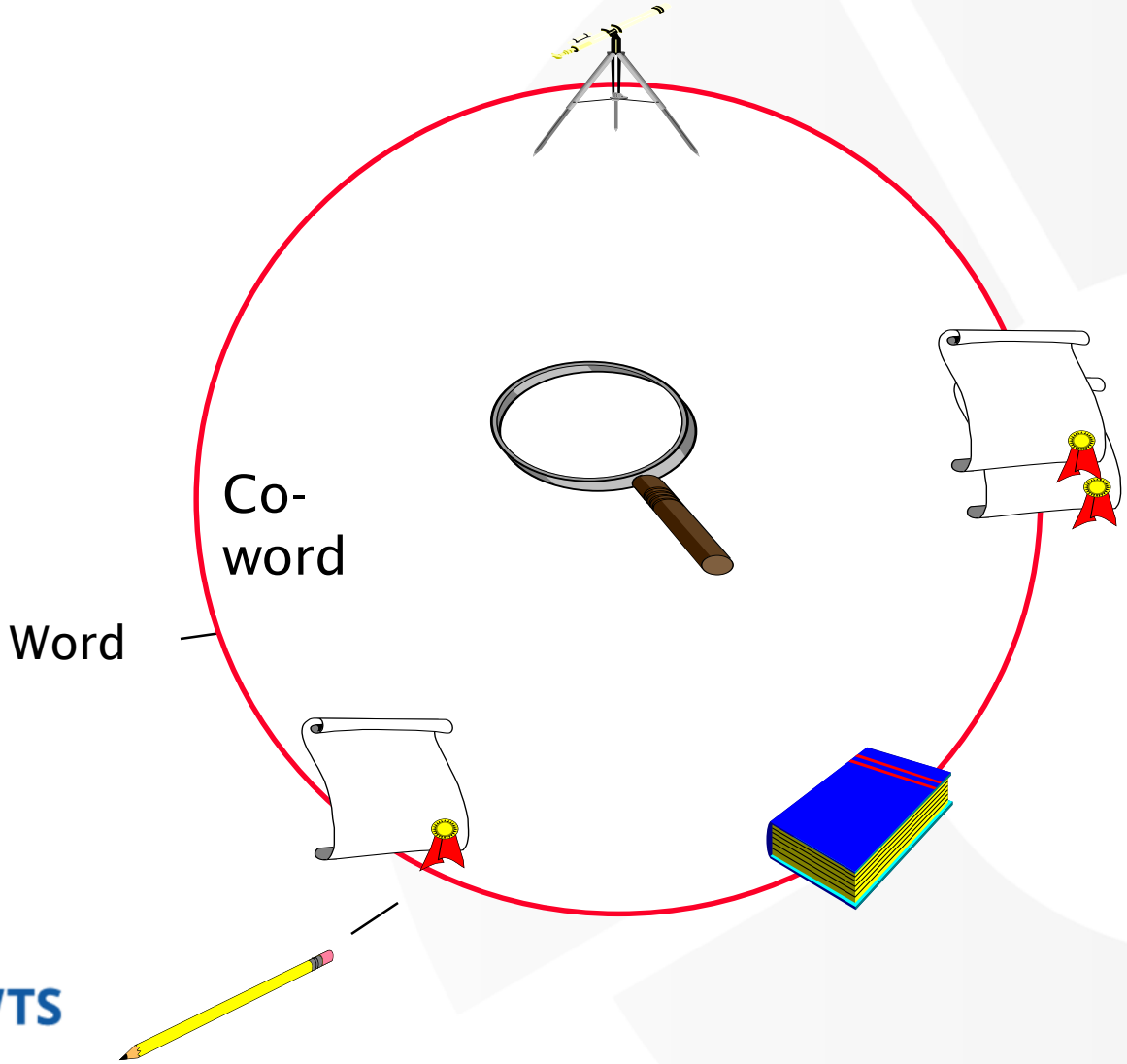
CWTS

The peer review cycle

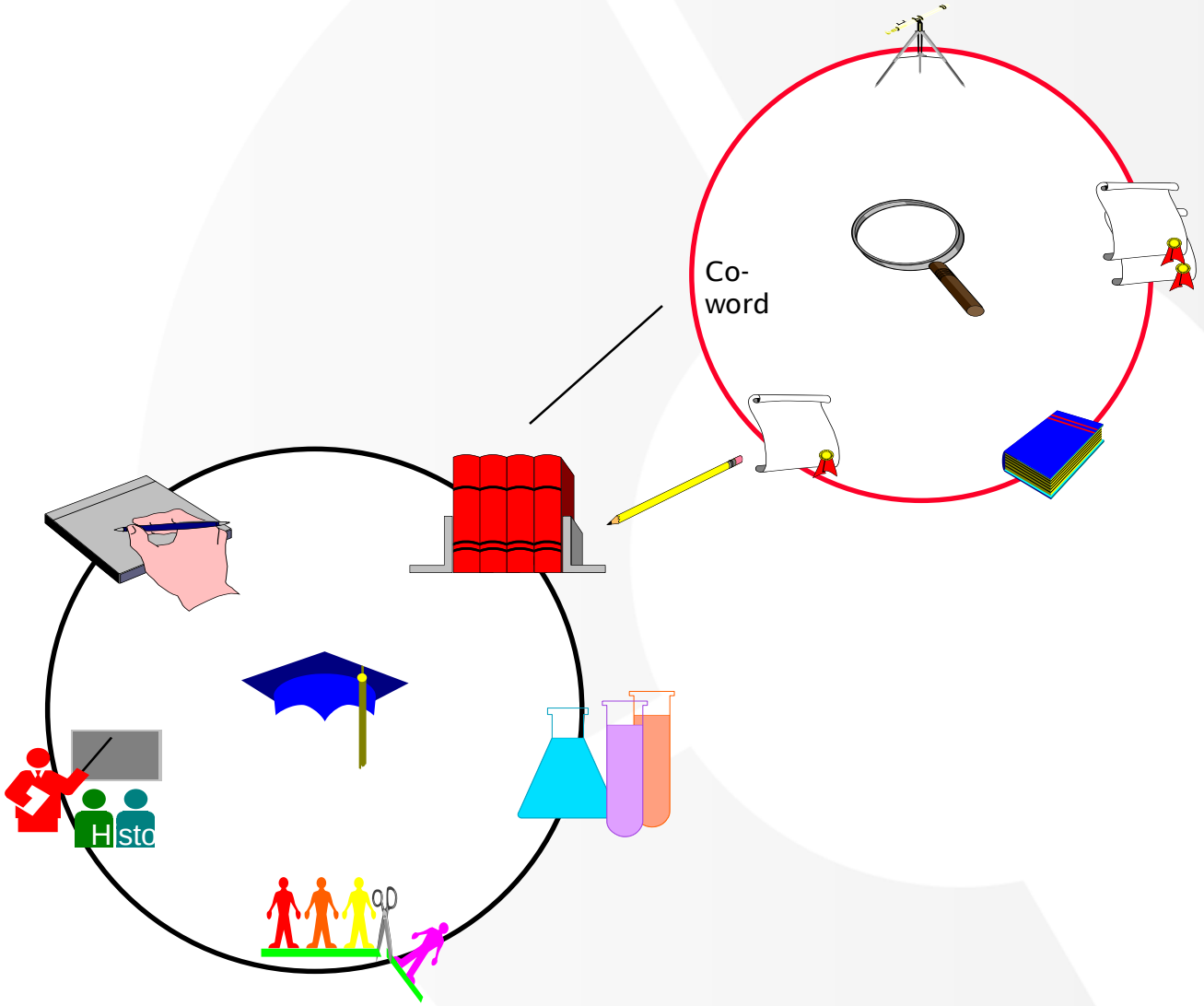


CWTS

The citation cycle



Two interacting cycles



"More specifically, if an individual's, department's or university's ability to amass symbolic capital of this kind were to become the critical determinant of future research funding and career advancement, then it would not be difficult to imagine distortions creeping into the system, as players devised recruitment, publication, collaboration and citation harvesting stratagems to accelerate and maximise the accrual of symbolic capital." (Cronin 2000, p. 450)

CWTS

Citation in two contexts

- Most citation theories based on communication system of science
- This is not identical to the social institution of evaluation in science
- Explaining the social life of citation indicators should be based on the latter
- Example: the black hole in “informed peer review”

CWTS

Implications

- Indicator: embodiment of a specific newly created link between the formal and the paradigmatic in 2 modes:
  - as link
  - as number
- Not one unified but multiple indicator theories (John Law & Annemarie Mol on reality multiple)
- Building indicators is extending the scientific social system with new objects
- Citation theories are performative

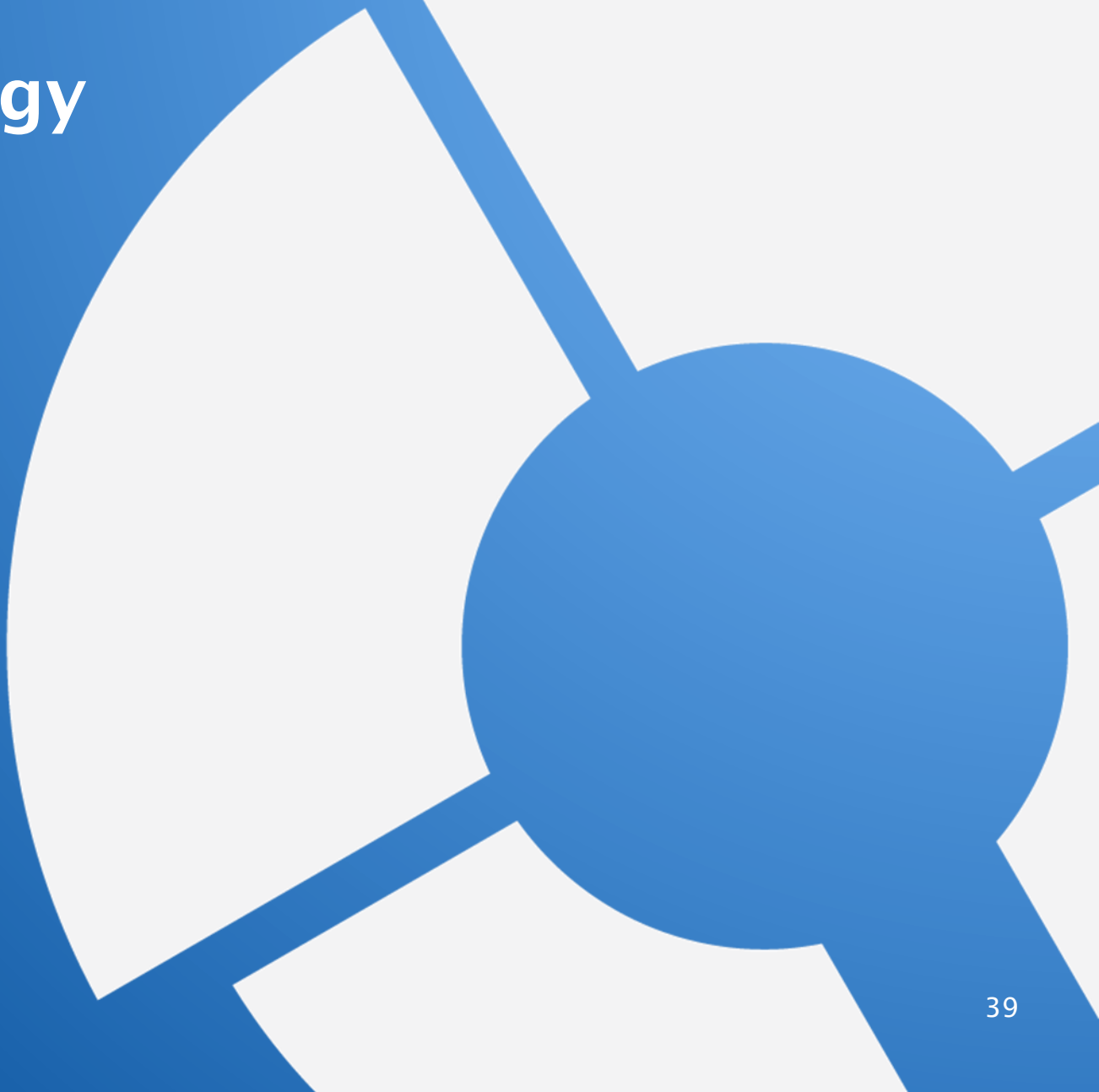
CWTS

Types of interactions

- Indicators directly used in funding decisions
- Indicators may indirectly redefine what *scientific quality* means
- The maps of science may influence priorities
- Scientists may validate indicators or maps
- Scientists may help construct indicators

CWTS

The sociology of quality



Quality

- Substantive (expert based)
- Formalized (procedural – meta method?)
- Ethnographic (actor defined)
- Sociological (power or interest based)
- Semiotic (translation)

- Proposal:

  quality is not an intrinsic property at the level of the individual but an effect of infrastructures

CWTS

Quality – alternative definition

- Quality is the level of “fit” between a particular work and the infrastructure to which it aspires
- Quality is multi-dimensional: more than 1 infrastructure at the same time
- Quality is distinct from the interests of the author
- New infrastructures can emerge from a lack of fit
- Innovativeness can be an aspect of quality but does not have to be required
- Quality can be measured but only partially

CWTS