Towards Indicators for 'Opening Up' Science and Technology Policy

Ismael Rafols[12]   **Tommaso Ciarli**[1]
Paddy van Zwanenberg[1]   Andy Stirling[1]

[1]SPRU – Science and Technology Policy Research, University of Sussex,
[2]INGENIO (CSIC-UPV), Universitat Politècnica de València
n.surname@sussex.ac.uk

The 'problematic' use of conventional S&T indicators

Closes down policy options (as well as technologies, in particular those closely associated with power, e.g. nuclear)

Narrow inputs – e.g. publications, citations, patents

Scalar outputs – e.g. rankings based on averages

  ▶ Aggregated solutions – missing within group variation

Opaque selections of inputs, outputs and classifications (privately owned databases)

Some quantitative assumptions are debatable

  ▶ Impact Factor of journals: only 2 years, ambiguity in document types
  ▶ Average number of citations for data power law distributed – small organisations penalised (Leydesdorff and Bornmann, 2011)

The 'political' use of conventional S&T indicators

Why have S&T indicators been so "narrow"?

S&T Indicators are **simple**: suitable to policy maker

S&T Indicators have a **performative** role: they don't just measure, they signal to stakeholders what is important

For example, **scientometrics** tools

- ▶ Not 'just happen to be used' in science policy (neutral)
- ▶ Part of the incumbent's power (loaded): e.g. evaluation of research
  - ▶ Policy makers, scientific community, job market (firms)

Scientific disciplines/communities and techniques such as statistics are a crucial 'part of the technology of power in a modern state' (Hacking, 1991, p. 181)

Institutions use these techniques to articulate framings, goals and narratives and get people to accept them

Ideas grounded on Foucault: "knowledge and power are inseparable"

Claims of the presentation

Need for more **inputs** (variables) to build indicators: '**broadening out**'

- ▸ Already happening

Need for multiple **outputs** – based on alternative assumptions – allowing for policy evaluation of the diverse options in building the indicator: '**opening up**'

How? Which tools?

Improving the use of tools for measuring S&T

Conventional indicators using narrow inputs

Can openly compare multiple outputs making explicit underlying concepts and enabling heuristic tools to facilitate exploration.

Complexity science tools and new science mapping tools

- ▶ More inputs: pubs, but also news, webs (Altmetrics), etc.
- ▶ Multidimensional outputs: interactive maps
- ▶ Multiple solutions for one indicator – assumptions
  - ▶ Defining disciplinary areas when not comparable
  - ▶ Different levels of aggregation
  - ▶ More inclusive and contrasting classifications
- ▶ Analysis of distributions / variance

Outline of the presentation

1. Intro and motivations
2. Background: policy use of S&T indicators
3. Framework: breadth and openness
4. Examples
   ▶ Opening up using broad inputs
   ▶ Opening using narrow inputs: Academic performance
   ▶ Opening using new tools: Interdisciplinarity
5. Discussion and work in progress

Policy use of S&T indicators: Appraisal

Appraisal

"The ensemble of processes through which knowledges are **gathered** and **produced** in order to inform decision-making and wider institutional commitments" (Leach et al., 2010)

Example: Allocation of resources based on research "(excell)ence"

Breadth – gathering

Extent to which appraisal covers diverse dimensions of knowledge

Narrow: citations/paper
Broad: citations, peer interviews, stakeholders, altmetrics, ...
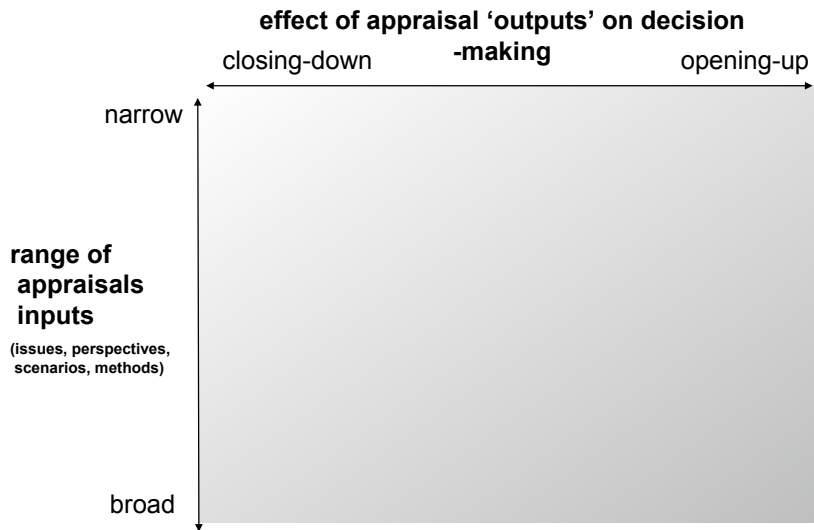
Openness – producing

Degree to which outputs provide an array of options for policies

Closed: fixed composite measure of variables → unitary and prescriptive advice
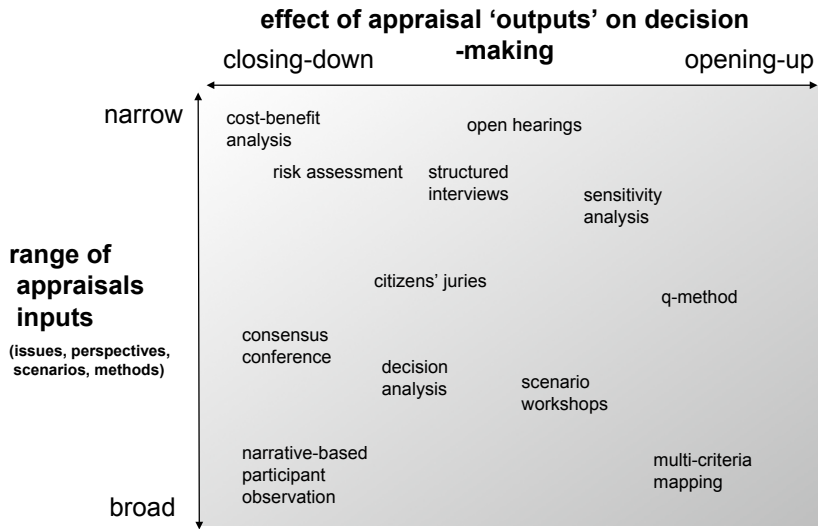**Open**: consideration of various dimensions → plural and conditional advice
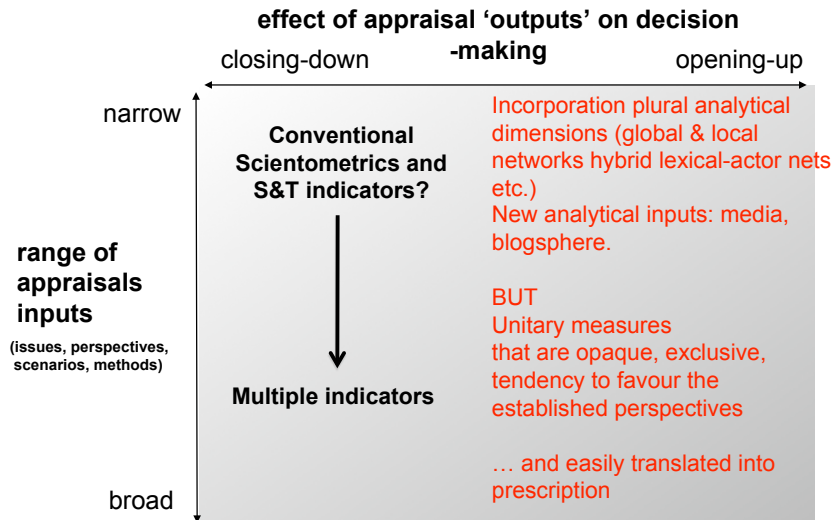
Appraisal methods: broad vs. narrow & close vs. open

Source: Leach et al. (2010)
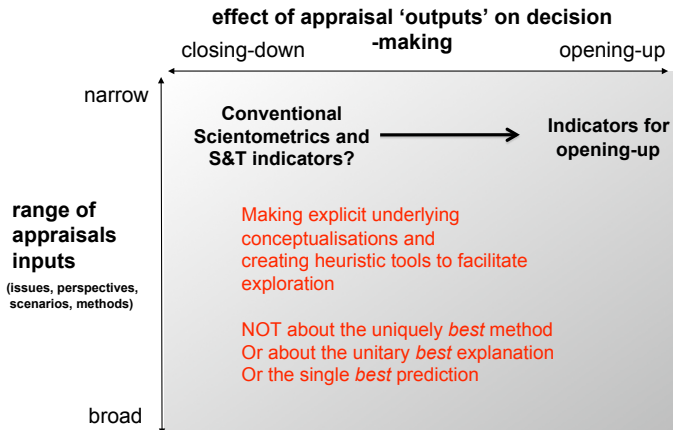
Appraisal methods: broad vs. narrow & close vs. open

Source: Leach et al. (2010)
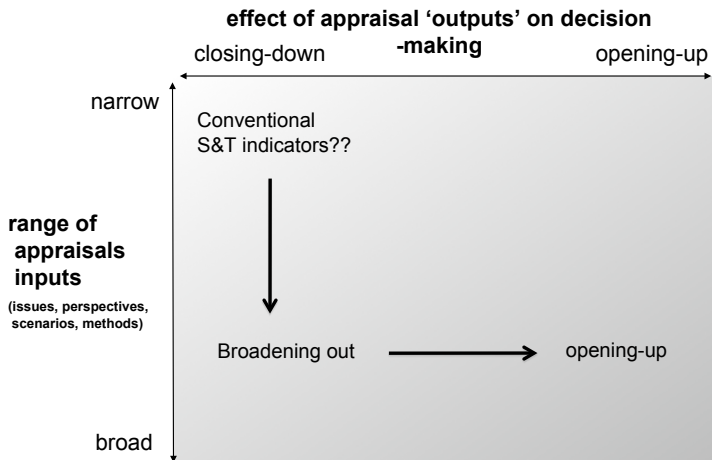
Appraisal methods: broadening out

effect of appraisal 'outputs' on decision -making

closing-down   opening-up

narrow

range of appraisals inputs

(issues, perspectives, scenarios, methods)

broad

Conventional Scientometrics and S&T indicators?

↓

Multiple indicators

Incorporation plural analytical dimensions (global & local networks hybrid lexical-actor nets etc.)
New analytical inputs: media, blogsphere.

BUT
Unitary measures
that are opaque, exclusive, tendency to favour the established perspectives

… and easily translated into prescription

Source: Leach et al. (2010)

Appraisal methods: opening up

effect of appraisal 'outputs' on decision-making

closing-down   opening-up

narrow

Conventional Scientometrics and S&T indicators?   **Indicators for opening-up**

range of appraisals inputs

(issues, perspectives, scenarios, methods)

Making explicit underlying conceptualisations and creating heuristic tools to facilitate exploration

NOT about the uniquely *best* method
Or about the unitary *best* explanation
Or the single *best* prediction

broad

There are different ways of opening up, remaining narrow (i.e. with narrow inputs as scientometrics)

Broadening-out → Opening-up

effect of appraisal 'outputs' on decision -making

closing-down   opening-up

narrow

Conventional
S&T indicators??

**range of
appraisals
inputs**

**(issues, perspectives,
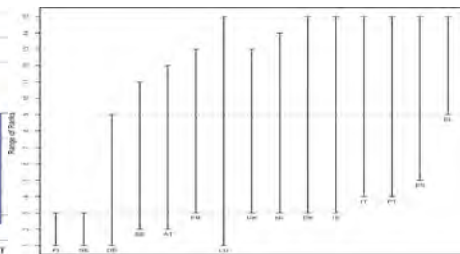scenarios, methods)**

Broadening out   →   opening-up

broad

First broaden, without collapsing the variables in one indicator

EU Innovation Scoreboard: composite indicator
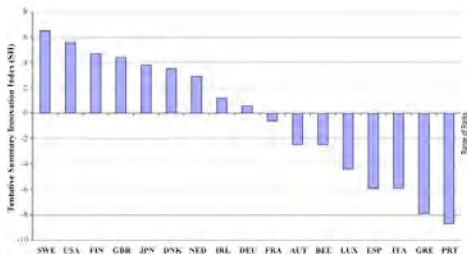
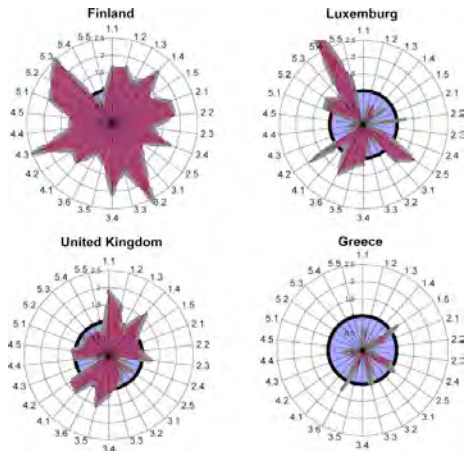(a) Country rankings

(b) Sensitivity analysis

Source: (Grupp and Schubert, 2010)

Broad but narrow S&T indicator

– Ranking (1a) is highly dependent on variables weightings (Grupp and Schubert, 2010)

– Sensitivity (1b): when adopting different weights almost every country could be ranked at any position

EU Innovation Scoreboard: opening the indicator

Source: (Grupp and Schubert, 2010)

Opening

Consider the variables of the indicator contemporaneously but separated

University ranking: opening the indicator

Source: http://www.u-map.eu/finder.shtml

"U-Map offers you tools to enhance transparency"

"A list of higher education institutions (HEIs) that are comparable on the characteristics *you* selected"

Difference in rankings (Innov VS BS) changing normalisation

Review a comparison of performance of six academic organisations using a bibliometric measure with different normalisations

Measure: average number of citations per publication (Rafols et al., 2012)

- **a** Number of citations per publication
- **b** Number of citations weighted by average citations in the journal of publication
- **c** Number of citations weighted by average citations in field of publications – e.g. condensed matter, computational biology, atomic physics, business, management, economic finance, etc
- **d** Number of citations weighted by the number of reference in the citing article
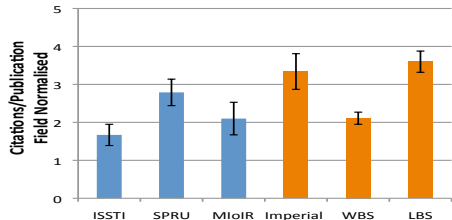
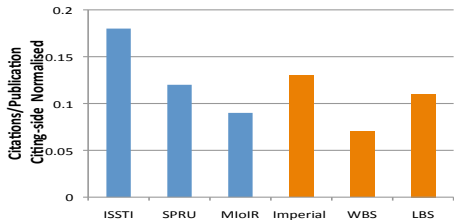# Difference in rankings (Innov VS BS) changing normalisation
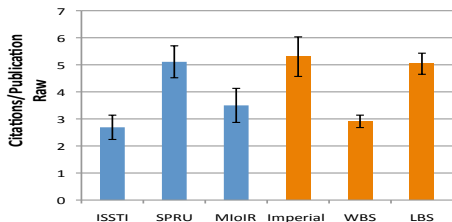
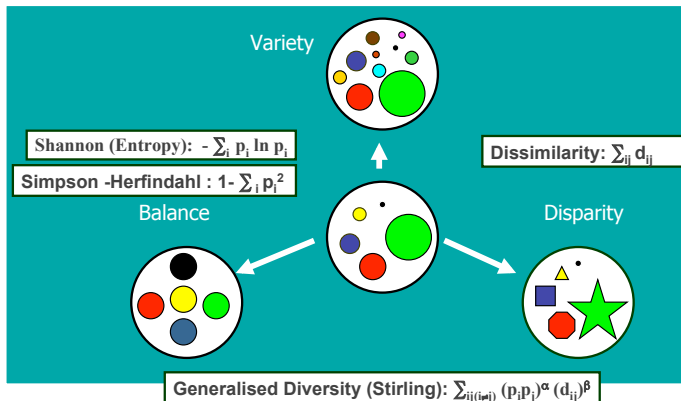

(a) Raw citations
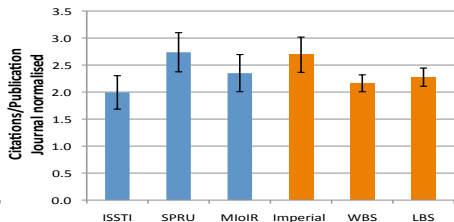
(b) Weighted by Journal

(c) Weighted by Field

(d) Weighted by References

Source: Rafols et al. (2012)

Heuristics of diversity

Variety

Shannon (Entropy): $-\sum_i p_i \ln p_i$

Simpson -Herfindahl : $1 - \sum_i p_i^2$

Dissimilarity: $\sum_{ij} d_{ij}$

Balance

Disparity

Generalised Diversity (Stirling): $\sum_{ij(i\neq j)} (p_i p_j)^{\alpha} (d_{ij})^{\beta}$

d: distance between categories; p: share
Source: Stirling (2007)

- ▶ Variety: Number of distinctive categories
- ▶ Balance: Evenness of the distribution
- ▶ Disparity: Degree to which the categories are different.

Interdisciplinarity as diversity

Bibliometric comparison of interdisciplinarity in different academic organisations using overlay maps (Rafols et al., 2012)

Indicators: journal attributes, publications and references

Distinguish different measures of diversity

- ▶ Variety: number of disciplines: n
- ▶ Balance: Size of each discipline: $-\frac{1}{\ln(n)} \sum_i p_i \ln p_1$
- ▶ Disparity: distance between the categories, computed using the Global Map of Science ↪: $\frac{1}{n(n-1)} \sum_{i,j} d_{i,j}$
- ▶ Shannon entropy: $-\sum_i p_i \ln p_1$
- ▶ Rao-Stirling diversity: $\sum_{i,j} p_i p_j d_{i,j}$
  where $d_{i,j} = 1 - s_{i,j}$, $s_{i,j}$ is the cosine similarity between categories i and j, and p_i the proportion of elements in category i

Different measures of diversity are uncorrelated (Yegros et al., 2010)

ISSTI Edinburgh – Disciplines of publication
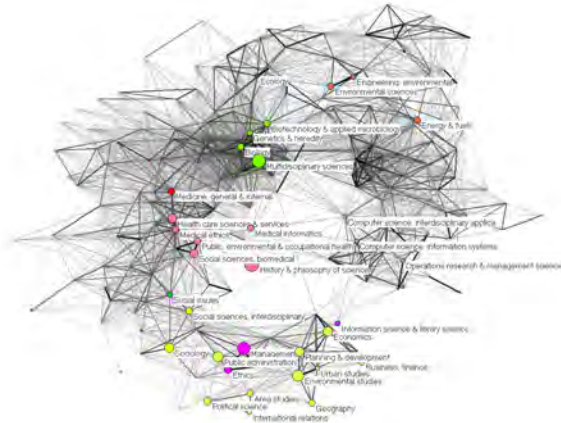
Source: Rafols et al. (2012)

Extremely diverse `Global map of Science`

Social sciences, from sociology to political sciences and economics, health
services, biological sciences, environmental sciences, and computer sciences

London BS – Disciplines of publication
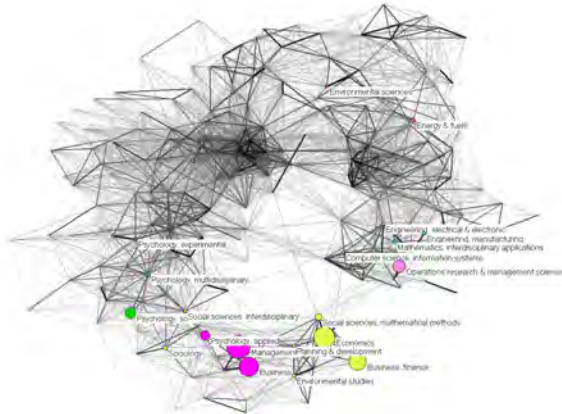
Source: Rafols et al. (2012)

Four disciplines `Global map of Science`

Management, Business, Economics and Finance (some Psychology and Operations research).

ISSTI and LBS compared

(a) ISSTI   **(b)** LBS
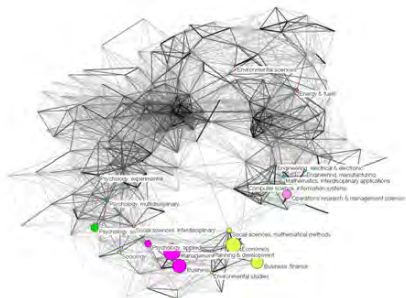
Source: Rafols et al. (2012)

Using a graphic visualisation we can study the different measures of diversity in one figure, without having to compromise as with composite indicator

MIoIR and WBS compared

(a) MIoIR Manchester   (b) Warwick BS
Source: Rafols et al. (2012)

Which one is more interdisciplinary?

Comparing diversities

	ISSTI	MIoIR	WBS	LBS
Variety	28	19	20	9
Balance	0.653	0.543	0.46	0.37
Disparity	0.832	0.817	0.77	0.768
Entropy	3.558	2.966	3.078	2.343
Rao Stirling	0.81	0.726	0.68	0.603

Source: Rafols et al. (2012)

**Which measure of diversity should we use to assess
interdisciplinarity? (and relate it to performance)**

Strategies for opening up

Work in progress...

Presenting contrasting perspectives

Simultaneous visualisation of multiple properties / dimensions

- ▶ Allowing the viewers/policy makers to take their own perspective
- ▶ Unveiling the assumptions and the properties of the indicators and variables (distribution?)

Interactivity

- ▶ Allowing the viewer to give its own weigh to criteria / factors
- ▶ Allowing the viewer to manipulate visualisation.

Closing thoughts

Keep it complex (Stirling, 2010)

Is 'opening up' worth the effort?

Conventional indicators tend to favour incumbents

- ▶ Incumbents have power and incentive to influence choice of indicators

Important to support diversity in S&T system

- ▶ Manage diverse portfolios to hedge against uncertainty in research
- ▶ Systemic ('ecological') understanding of the S&T
- ▶ Evolutionary understanding of excellence and relevance
- ▶ Open possibility for S&T to work for the disenfranchised
  - ▶ There aren't neglected diseases. There are neglected populations.

Conventional Policy Dynamics

'lock-in' to policy
favoured by incumbent
power structures

GOVERNANCE
COMMITMENTS

narrow scope
of attention

S&T indicators

risk assessment

cost-benefit analysis

*also: restricted options,
uncertainties in participation*

incomplete knowledges

SOCIAL
APPRAISAL

designed)
Res. Excellence

GUIDANCE / NARRATIVE

simple 'unitary'
prescriptions

*expert judgements /
'evidence base'*

"best / optimal /legitimate"

Source: Stirling 2010

Breadth, Plurality and Diversity

dynamic portfolios
pursuing diverse
trajectories

MULTIPLE
TRAJECTORIES

GOVERNANCE
COMMITMENTS

broad-based
processes of
'precautionary
appraisal'
*multiple: methods,
criteria, options, frames,
uncertainties, contexts,
properties, perspectives*

SOCIAL
APPRAISAL

Sustainability ⟶

'opening up' with
'plural conditional'
outputs to policymaking

*viable options under:
conditions, dissonant views,
sensitivities, scenarios, maps,
equilibria, pathways, discourses*

Source: Stirling 2010

Background

Global map of science – 222 SCI-SSCI Subject Categories

Rafols, Porter and Leydesdorff (2010)

Global map of science – 222 SCI-SSCI Subject Categories

- ▶ CD-ROM version of the JCR of SCI and SSCI of 2009
- ▶ Matrix of cross-citations between journals (9,000 x 9,000)
- ▶ Collapse to ISI Subject Category matrix (222 x 222)
- ▶ Create similarity matrix using Saltons cosine (Rafols et al., 2010)

ISSTI

References I

Grupp, H. and Schubert, T. (2010). Review and new evidence on composite innovation indicators for evaluating national performance. *Research Policy*, 39(1):67 – 78.

Hacking, I. (1991). How should we do the history of statistics? In Burchell, G., Gordon, C., and Miller, P., editors, *The Foucault Effect: Studies in Governmentality*. University of Chicago Press, Chicago.

Leach, M., Scoones, I., and Stirling, A. (2010). *Dynamic sustainabilities: technology, environment, social justice*. Earthscan.

Leydesdorff, L. and Bornmann, L. (2011). Integrated impact indicators compared with impact factors: An alternative research design with policy implications. *Journal of the American Society for Information Science and Technology*, 62(11):2133–2146.

References II

Rafols, I., Leydesdorff, L., O'Hare, A., Nightingale, P., and Stirling, A. (2012). How journal rankings can suppress interdisciplinary research: A comparison between innovation studies and business & management. *Research Policy*, 41(7):1262 – 1282.

Rafols, I., Porter, A. L., and Leydesdorff, L. (2010). Science overlay maps: a new tool for research policy and library management. *Journal of the American Society for Information Scienceand Technology*, 61(9):1871–1887.

Stirling, A. (2007). A general framework for analysing diversity in science, technology and society. *Journal of The Royal Society Interface*, 4(15):707–719.

Stirling, A. (2010). Keep it complex. *Nature*, 468:1029–1031.

Yegros, A., Amat, C., DEste, P., Porter, A. L., and Rafols, I. (2010). Does interdisciplinary research lead to higher scientic impact? Conference paper, STI Indicators Conference, Leiden.