It’s not how you measure – it’s what you measure

Fiona Wood

Presentation to the ESF Member Organisation Forum on Peer Review

6 December 2011

Why invest in research?

To change the world

- Economically
- Socially
- Environmentally

Where to invest?

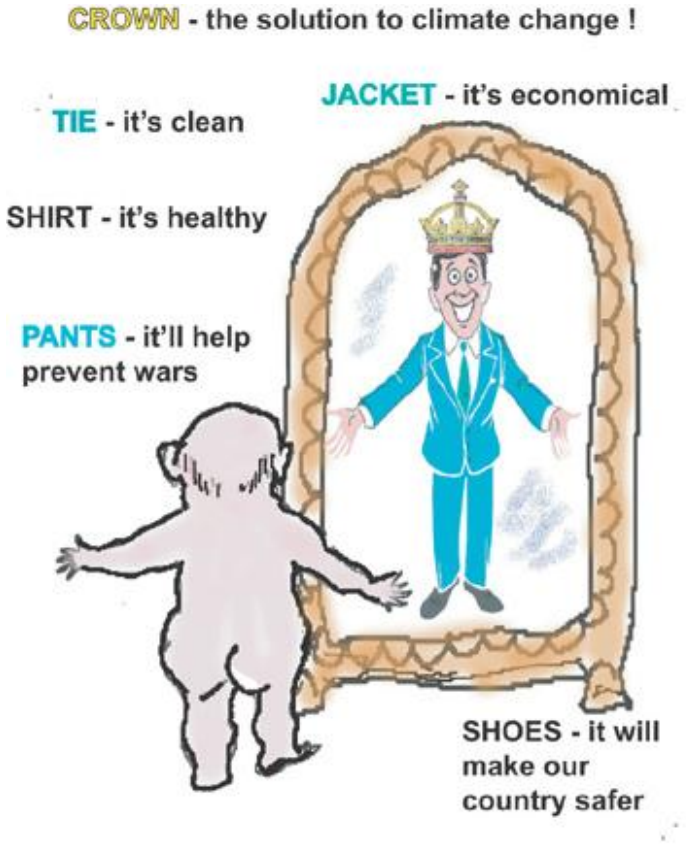
- High risk-high return curiosity driven research (internet to lasers to the sex life of the screw worm)
- Research geared to the grand societal challenges
- Creation of new innovation and entrepreneurship skills

How to invest?

- Support basic/frontier and applied research separately
- Develop performance measures that match the different expectations

How well do bibliometrics measure the value of investment in these areas?

Not well…



http://nuclear-news.net/2010/10/16/the-emperors-new-nuclear-clothes/

Use and abuse of Bibliometrics

- Count publications, patents, citations to develop S&T performance indicators
- Treat them as proxies for performance
- Use them to guide policy development & the allocation of funds, grants, promotion etc

OECD Frascati Manual; RN Kostoff Use and Misuse of Metrics in Research Evaluation, Science and Engineering Ethics (1997) 3, 109-120

BUT….

‘…bibliometrics does not properly take into account the originality of the academic’s work, conceptual innovation, research applications, scientific and industrial utility’

Institut de France Académie des sciences 2011

…and furthermore…

- Citation counts are a measure of quantity, not quality
- Bias towards English language, established researchers and low risk research
- Induce conformity to the ‘exigencies’ of the measure (salami publications, risk-averse research, fraud)
- Demoralises and disenfranchises researchers

….leading to….

- an enormous fragmented data sets of dubious meaning
- a massive disaggregated literature
- an aggressive band of metrics zealots

HOW NOT TO FIX IT…

Excellence in Research for Australia 'ERA' - Objectives

- establish an evaluation framework that gives government, industry, business and the wider community assurance of the excellence of research conducted in Australia's institutions;
- provide a national stocktake of discipline-level areas of research strength and areas where there is opportunity for development in Australia's higher education institutions;
- identify excellence across the full spectrum of research performance;
- identify emerging research areas and opportunities for further development; and
- allow for comparisons of Australia's research nationally and internationally for all discipline areas.

‘ERA’ indicators

1. research quality (publications, citations, research income, peer review)
2. research volume and activity
3. research application
4. esteem measures

Unit of analysis – the research discipline at each institution

Rating Scale (magically) derived from indicators:

5	well above world standard evaluation
4	above world standard
3	at world standard
2	below world standard
1	well below world standard
NA	research outputs did not meet volume threshold

- No operational defn for ‘world standard’

Problems

- Assessed basic and applied research together
- Difficulties with apportioning MICT to classification codes
- Used a ranked journals list
- Did not measure impact
- Process not transparent

Undesirable outcomes

- Attributing meaning to a meaningless classification system
- Discontinuation of fields of research and and researcher positions
- Poaching of ‘star research teams’

Undesirable outcomes

- Committed just over $35 million, occupied over 600 people's time, and assessed 330,000 research outputs from 41 higher education institutions without addressing any of the real questions

Fiona Wood 2011 *ERA: an ailing emperor's new clothes*. R&D Review Feb-Mar: 12-13. Invited Op. Ed.

HOW TO FIX IT

- Develop measures that reflect the different purposes for investments in curiosity and mission-driven research
- Capture data at sufficiently granulated levels
- Increase global collaboration between funding agencies in developing & trialling new measures

Good Models

- Lattes Platform in Brazil
- SIAMPI pilot KNAW
- STAR-METRICS USA

Lattes Platform Brazil

- Used in making funding funding decision by government and in universities for tenure and promotion
- High quality data on history of scientific, academic and professional activities of researchers and institutions
- Incentives for compliance
- Unique researcher identifiers

http://lattes.cnpq.br/english/index.htm

SIAMPI

- Social Impact Assessment Methods for research and funding instruments through the study of Productive Interactions between science and society
- Offers a framework to assist in the identification and assessment of social impact of research activities, programmes and organizations.
- Case studies in Health, ICT, Nanotechnology and SS&H

NB ESRC in the UK initiatives in capturing impact

http://www.siampi.eu/Pages/SIA/12/643.bGFuZz1FTkc.html

STAR-METRICS

- Science and Technology in America's Reinvestment – Measuring the Effects of Research on Innovation, Competitiveness and Science

- NSF, NIH and OSTP

https://www.starmetrics.nih.gov/Star/Participate#about

Phase 1 – focus on job creation from stimulus

1. Build a clean data base
2. Use university administrative records to calculate the employment impact of federal science spending through the American Recovery and Reinvestment Act and agencies' existing budgets

Phase 2 - focus on impact

1. Economic growth – through indicators such as patents and business start-ups
2. Workforce outcomes – measured by student mobility into the workforce and employment markers
3. Scientific knowledge - measured through publications and citations
4. Social outcomes -measured by long-term health and environmental impact of funding.

STAR-METRIC use of Visual analytics

- Visualising techniques – applied to complex data sets in trying to understand and describe the scientific and innovation enterprise as well as explaining outcomes to policy makers.
- Utilising visual analytics developed by Homeland Security to describe terrorist networks to describe scientific ones.

Where next for European funding councils?

- Important to develop protocols about what data to collect, what they mean and how to use them
- Need a better balance between accountability and risk, requiring major cultural changes in the people and organizations involved in the sponsorship and evaluation of research
- E-Collaborate in the development of a universal template for reporting scientific achievements as proposed by Julia Lane, NSF

ERAB has made a great start

- Recommendations and advice provided by the European Research Area Board in its 2011 contribution to the ‘Common Strategic Framework for Research and Innovation’ consultation

Including…

- Be ambitious and be prepared to take managed risks for the sake of the European economy.
- Divide support between curiosity and mission-driven. The latter to include both high risk enabling technologies and further support for Europeans competitiveness.
- Create a number of independent arms length funding agencies to support and govern different types of excellent research and innovation.

THANK YOU FOR YOUR ATTENTION!