



THE AUSTRALIAN NATIONAL UNIVERSITY

Policy Perspective: The Politics of Evaluating Science Impact

Dr Claire Donovan

Research Evaluation & Policy Project

Research School of Social Sciences

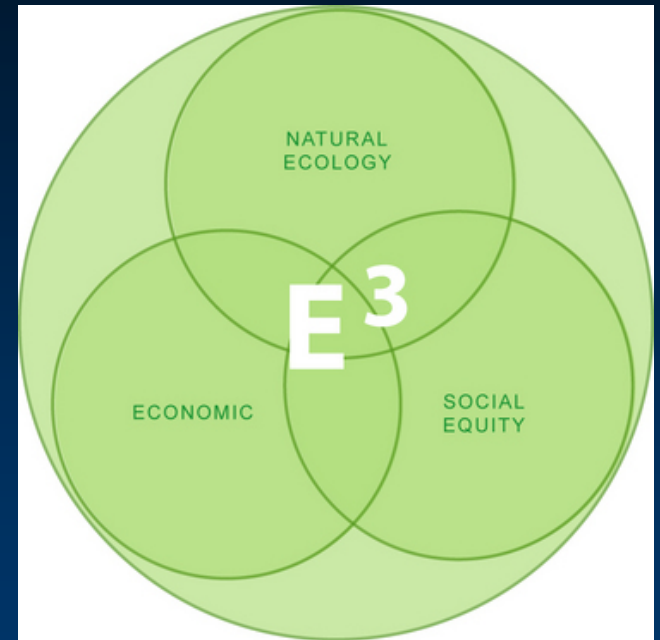
- Special edition of *Science and Public Policy*: 'Future Pathways for Science Policy and Research Assessment: Metrics vs. Peer Review, Quality vs. Impact' (October 2007)
- 'The Australian Research Quality Framework: A Live Experiment in Capturing the Social, Economic, Environmental and Cultural Returns of Publicly Funded Research', in C. L. S. Coryn and M. Scriven (Eds.) *New Directions for Evaluation: Reforming the Evaluation of Research*. (Los Angeles: Jossey-Bass forthcoming December 2007)

Policy dilemmas (1): economic rationalism

- Science policy driven by wealth creation, international competition & technological advance
- Simple metrics based on natural science activity
- Simple 'impact' metrics do not measure research impact:
 - **Technometrics** ⇒ economic returns ⇒ low level impact ⇒ private over public interest
 - **Sociometrics** ⇒ macro social statistics ⇒ no credible causal link to particular research efforts
- Simple impact metrics ignore wider public value
- Holistic measures rely on peer (and 'end-user') judgements, and are necessarily complex

Role of impact evaluation: public value

- Meaningful impact assessments must strive to measure the wider public value of research
- Triple bottom line accounting:
 - beyond neo-liberal NPM
 - centre-left, centre-right
 - economic, social, environmental gains
- Redefines purpose of STI policy
- Public value eludes quantitative approaches \Rightarrow simple metrics increasingly detached from STI policy imperatives



Policy dilemmas (2): the 'Pushmi-pullyu'

Politics of constructing research evaluation exercises is like a 'Pushmi-pullyu':



External audit ↔ internal peer-based appraisal

Broader relevance ↔ scientific autonomy

Interests of industry & commerce ↔ broader public benefits

Policy dilemmas (2): the 'Pushmi-pullyu'

Key points of tension:



'Quality' vs. 'impact' (the 'relevance gap')

Metrics vs. peer review (simplicity vs. complexity)

Isolation from developments in scientometrics

Lack of policy learning from other exercises & nations

UK Research Assessment Exercise

- Two decades of 'quality' assessment, peer and discipline-based
- Post-2008 'quality' evaluation:
 - metrics only for STEM
 - light touch peer review & discipline-specific metrics for HASS, maths and statistics
- 'Impact' defined as maximising the economic impact of research
- New system to reward 'user-focused' research, i.e. 'the relative amount of research universities undertake with business'
- HEFCE to distribute £60m (€85m) 2007-08
- Simple impact metric likely (revenue generated?)

UK Research Assessment Exercise: dilemmas

- Simple income metrics/technometrics do not measure 'impact'
- Public subsidy for private value
- Economic rationalism: the research society most needs?
- Against the grain of international developments in 'impact' measurement
- UK Research Councils to assess 'economic impact' of funding applications (includes social and cultural)

- 'Quality' and 'impact' assessment begins 2008: research group focus with discipline 'clusters' of peer and 'end-user' panels
- Qualitative approach for 'impact', with minimal role for metrics
- 'Impact' defined as social, economic, environmental and cultural benefits from research
- DEST to distribute AU\$600m (€385m) per year, but proportion for 'impact' unknown
- Evaluation of impact statements and case studies (plus 'end-user' testimony if required)
- Complex contextual approach, with groups rated against a controversial 'impact scale'

- Collapse of 'impact' metrics when applied to impact scale and to broader public value of research
- Resource intensive process
- Does high 'impact' mean low 'quality' research?
- Follows *some* developments in scientometrics, but complexity seized on by political opposition

Policy dilemmas (3): implementation

- Dominance of science policy network:
 - desire for simple metrics
 - 'engineering'-based impact scale
- Compromises in design:
 - partial adoption of scientometric advice
 - complexity a political weakness
- Bureaucrats:
 - high turnover & loss of policy learning
 - cycle of conversion to qualitative processes
 - hostility to non-metrics approaches
 - confused guidelines
- 'Science' default setting (economic rationalism, excluding HASS)

The RQF Impact scale

Rating	Description
A	Adoption of the research has produced an outstanding social, economic, environmental and/or cultural benefit for the wider community, regionally within Australia, nationally or internationally.
B	Adoption of the research has produced a significant social, economic environmental and/or cultural benefit for the wider community, regionally within Australia, nationally or internationally.
C	Research has been adopted to produce new policies, products, attitudes, behaviours and/or outlooks in the end user community.
D	Research has engaged with the end user community to address a social, economic, environmental and/or cultural issue regionally within Australia, nationally or internationally.
E	Research has had limited or no identifiable social, economic, environmental and/or cultural outcome, regionally within Australia, nationally or internationally.