

Building a Contextual Evaluation Culture

Why and how

European Foundation Centre
Jack Spaapen 020610
ERiC / SIAMPI projects

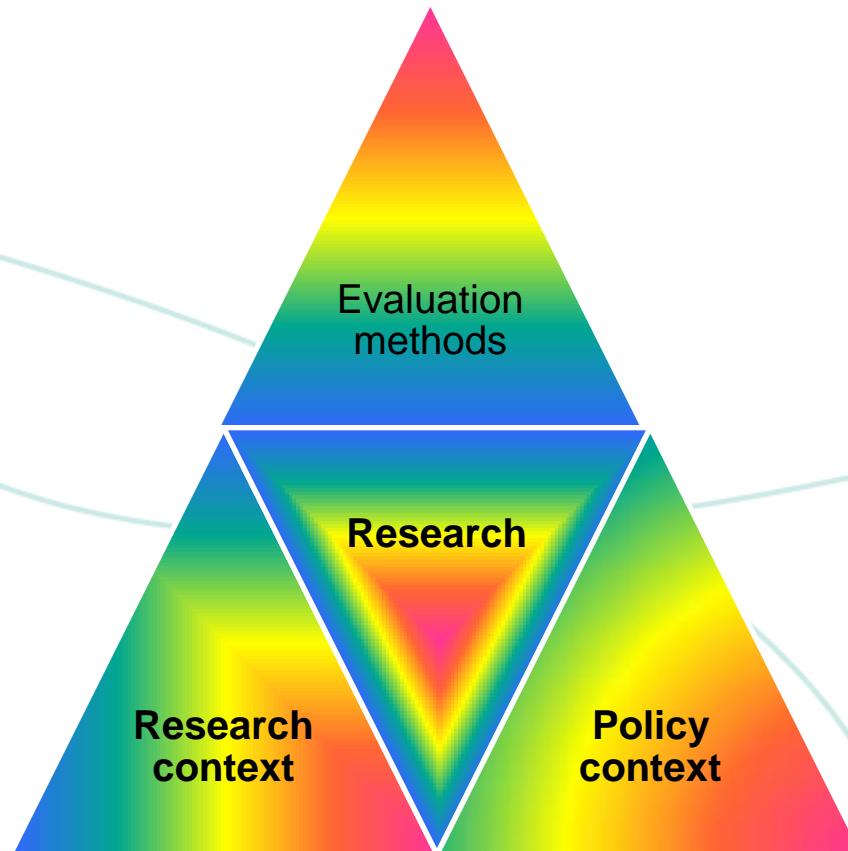
Changing policy context, demanding efficiency and effectiveness

- Growing pressure from policy context, (inter)national and local, since 1970s, focus on accountability, priority setting
- Growing internationalization, European context (FPs), bigger programs, more collaboration with external parties (a.o. for financial reasons)

Changing research context, need for broader evaluation

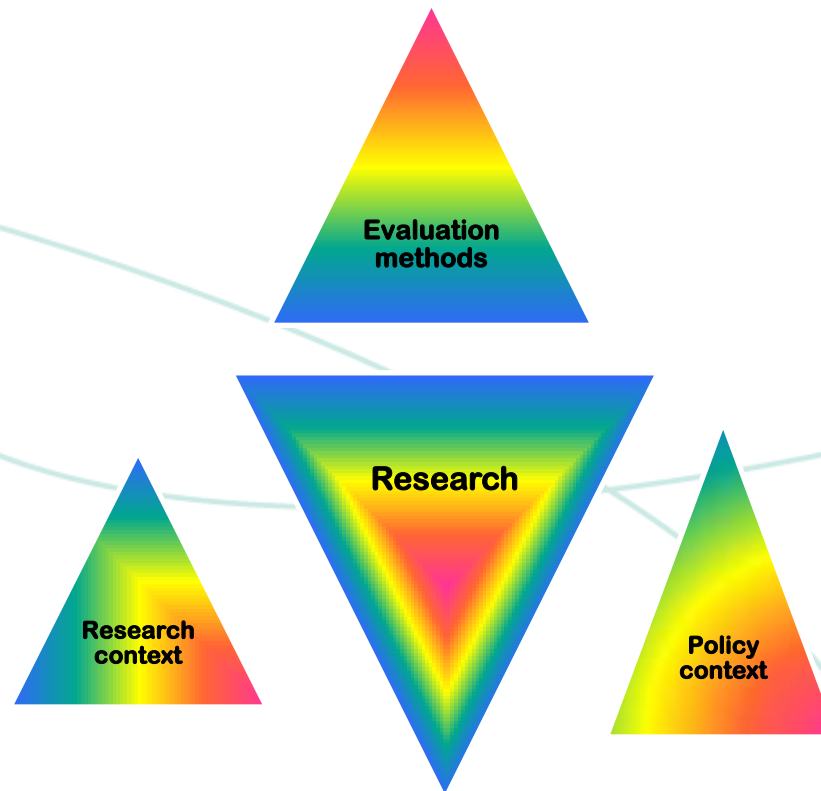
- More research in the context of application, more collaboration between disciplines, pressure from EU (Lisbon agenda 2002)
 - Mode 1: disciplinary competition / int. comparison, focus on int. journals, IF etc.
 - Mode 2: problem solving, context of application, focus on productive networks, socially robust knowledge
- Traditional evaluations, based on peer review / bibliometrics, often do not fit the new situation

the evaluation triangle

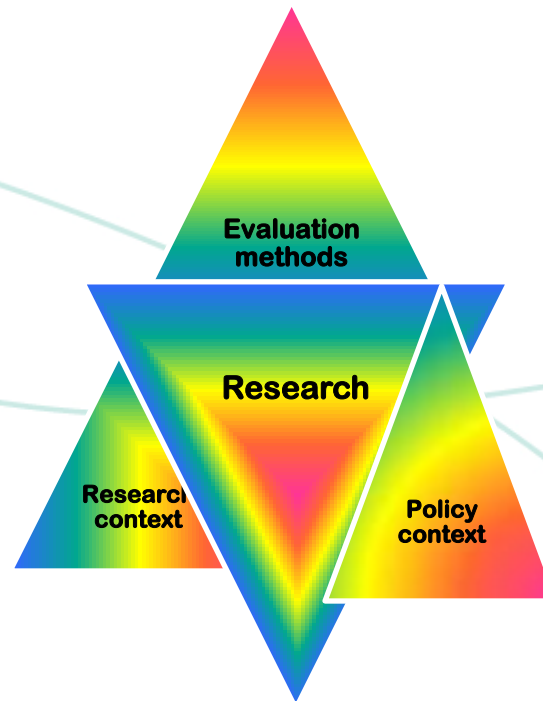


the evaluation triangle

weak interaction



the evaluation triangle
strong interaction



Different research profiles different output / outcome

- **Science oriented: scientific articles, books**
- **Industry oriented: patents, applications**
- **Policy oriented: reports, debates, protocols**
- **Society oriented: stakeholders, end users**
- **Mixture: different combinations of the above, f.e. building of new cities via cradle-to-cradle principles**

Researchers moving between these different profiles have to deal with different evaluation contexts

Different output, different indicators

- **Science indicators: production, impact**
- **Impact indicators: short term, long term**
- **Dissemination indicators: variegated media**
- **Responsiveness indicators: use, uptake**
- **Network indicators: citation, collaboration, funding**
- **Narratives**

Questions about data gathering, robustness, weighing, etc.

New evaluation culture

- From input oriented accountability to output oriented evaluations, (development of bibliometrics), to contextual approaches (stakeholders)
- Focus on wider context, policy demands, application orientation, role of stakeholders
- Sensitive to differences between fields of research, research profiles, MIT research
- Recognition of new indicators
- Keep it simple and doable
- Attunement / Integration at national / EU level

Four generations of evaluation

- 1. measuring performance of individuals
- 2. measuring the close environment of individuals (the curriculum, the research program)
- 3. achievements, are they worthwhile
- 4. stakeholder oriented evaluation

Dutch Standard Evaluation Protocol (SEP)


- Self evaluation report by research unit
review of past performance and forward looks (SWOT)

External site visits every 6 years; every 3 years mid term evaluation

- 4 main criteria:
 - Quality (international benchmarking)
 - Productivity (output, impact)
 - **Relevance (for policy, industry and society)**
 - Viability (policy and management)
- Evaluation both retrospective and prospective
the accent is on the latter



**Evaluating Research
in
Context**



**Social Impact Assessment Methods
through
Productive Interactions
between research and society**



ERiC / Siampi questions and approach

- Comprehensive evaluation, social impact assessment
- Useful in different fields and in different organizational contexts
- Stakeholder oriented
- Mission / profile oriented
- Focus on 3 productive interaction mechanisms (personal, media, money)
- Indicators developed in consultation with stakeholders
- Identifying specific instances of impacts – “outcome”.

Quality and relevance of architecture research

- Scientific Quality
- Scientific Production
- Scientific Recognition
- Responsiveness of research agenda
- Collaboration with (possible) users
- Dissemination and knowledge transfer
- Actual impacts and use of research

Note: in some areas of architectural research scientific quality and recognition are strongly related to societal contribution.



MEC assessment issues

Organizational context

- **Disciplinary vs institutional evaluation**
- **Societal vs scientific quality**
- **Collaboration with other institutes, local goals and ambitions**
- **Financial aspects, administrative burden**

Peer review procedure

- **Evidence based assessments?**
- **Inflation of scores ?**
- **Soft Benchmarking ?**
- **Consequences of evaluation reports, point of view of the board**



some plus issues

- Overall: SEP flexible enough, leaves room for alternatives
- Goals & ambitions: Self evaluation is seen as very positive, has become more concise and concrete in 2009
- Peer review: Accent on forward looking, mission orientation, mixed composition
- Organizational context: decision making process more transparent, information better organized (new system)



some minus issues

- Societal relevance: needs more attention, lack of knowledge how to measure this, ERiC to be implemented
- Peer review: still geared towards assessment of scientific quality, room for mixed committees in 2009
- Verdicts: inflation of scores, not critical enough → group & network dynamics (too much solidarity)
- Follow up → still unclear what happens with the results

Main issues in developing a new evaluation culture (wrap up)

- Balance between fairness and simplicity: fairness re different profiles, disciplines, scores; simplicity re procedures, administrative burden
- Peer review : problems with the assessment of MIT research, socio-economic value, 'referee fatigue', resistance toward innovation (f.e. expert review, mixed committees, stakeholders)
- Bibliometrics: too narrow focus on ISI journals; on biomedical and natural sciences; new measures (google!) are being developed but not easily accepted; ranking and classification debate
- Policy pressure for (national) comparison and simple measures: H-index, G-index etc.
- Specify goals of assessment: judgment/score vs improvement, mission oriented