



Judging Research on its Merits

Evaluation and Valorisation of publicly funded research

Jack Spaapen
ESF memberships forum on evaluation
Berlin, 21-23 October 2007



Royal Netherlands Academy of Arts and Sciences



2 parallel developments

Will they merge?

1. SEP = national system of evaluation since 2003

what is it and does it work?

2. Valorisation debate

background: European policy goal to close the 'knowledge gap'

questions: how to get research geared more towards the needs of society, and how to evaluate societal quality or impact of research

→ **Towards a new SEP in 2009**





Before 2003

- ❑ Universities had disciplinary evaluations, geared towards the evaluation of scientific quality
- ❑ Academy, Research council, have research institutes, own protocols, broader evaluations
- ❑ Other public research institutes, no formal regulations





Since 2003

Committee recommends overall national system:

- ❑ **Mission oriented and comprehensive**

(it may include societal goals, assessment of public functions, f.e. library and collections, and also the institute's policy and its management)

- ❑ **Evaluations no longer at the national level per discipline but local per 'institute'**

(concept 'institute' used in a broad way: an 'organisational unit covering a more or less coherent area of research', criteria and indicators may differ to some extent)

- ❑ **The governing boards of universities, institutes fully responsible for the evaluation process**

(They appoint review committees, f.e. of mixed composition, and comment on the evaluation reports, and show the (possible) effects)





Some important implications

- ❑ Boards decide about when and how to evaluate (was disciplines)
- ❑ Self evaluation provides opportunity to focus on what's found important in different areas of research
- ❑ Institutions to focus on their broader mission, including societal goals, applications etc.
- ❑ Peer review committee → allowed for mixed expertise
- ❑ Information process easier (thanks to the new national information system (METIS))
- ❑ (Hopefully) no double work in gathering and using the information





Major consequences

- Overall more flexibility, more room for alternative approaches
- No national comparison, government not happy





Standard Evaluation Protocol (SEP)

- **Universities, Academy, Research Council agree to one overall system → SEP**
- **All academic research evaluated according to SEP, and more, f.e. national research programs (NGI), and on a voluntary basis: public research bureaus (SCP)**
- **1st period: 2003 – 2009**
- **meta evaluation committee (MEC)**
- **1st MEC report March 2007**





4 main assessment criteria

Quality: international recognition, innovative capacity

Productivity: output in journals, or other media [focus on sci]

Relevance: impact on science, polity and society

Vitality and feasibility (flexibility, management aspects, leadership)

verdicts in five categories (excellent (5), very good (4), good (3), satisfactory (2), unsatisfactory (1))





Outline assessment procedure

- **Information**

- Self evaluation report (mission, management, results, future policy etc.)
- Key publications
- Input and output data

- **Focus**

- both group level and institute as a whole
- Four criteria (quality, productivity, relevance and management)
- SWOT analysis

- **External site visit**

- every 6 years, internal mid-term in between
- Assessment is both retrospective and prospective, accent on the latter

- **Meta evaluation**





Main differences

	OLD	NEW
1. overall approach	Disciplinary evaluations, Jury model	Mission oriented (SEP), Coach model
2. Responsibility	Unclear: Ministry? Discipline?, universities?	Board of institution, faculty, institute
3. Format	Disciplinary review , national comparison	Review of 'institute', int. comparison, self evaluation statement of the board, meta evaluation
4. goals	Accountability, assessing scientific quality of research	Accountability, assessing research quality in a broad sense, but also other aspects and research management
5. focus	Past performance, scientific quality, narrow criteria	Future mission, scientific quality, socio-economic relevance broad range of criteria





Critique on evaluation system(s) continues

- **Growing dissatisfaction with the traditional ways of evaluating from soc sc., humanities, medical and health, technical disciplines, etc.**
- **Too much geared towards norms and values of natural sciences
→ Focus on publication figures and impact → ‘knowledge paradox’**
- **Consequently, bad scores in evaluations and grant applications for soc sc, hum,, but also for many techn. areas, MIT research (=way more than half of all research)**
- **Costly, inefficient and time consuming (too many different evaluations: disciplines, institutes, research schools, program grants), unclear what the benefits are**





Coincides with valorisation demands from government

- **National and European governments discover ‘knowledge gap’; political to become no.1 economy in the world**
- **Mounting pressure from many research fields to develop another system, different criteria and indicators**
- **Search for alternatives, studies, reports, conferences, etc.**





Examples

- **RMW 2002:** societal impact of applied health research
- **SWR, RGW 2005:** judging research on its merits
- **AWT 2006:** 'alpha rays'
- **Research councils:** various (Zon/MW, RGO)
- **ERiC 2005:** evaluating research in context





Judging research on its merits

- **Self evaluation focus on both scientific reputation and communication, and on other audiences**
- **List of target groups: peers, students, professionals, policy makers, business, broader public**
- **List of indicators per target group: publications, citations, but also text books, reviews, grants from policy, collaborations with business, professionals, awards, popular publications, etc.**
- **Benchmarking as a critical process**





ERiC – evaluating research in context

www.eric-project.nl

- **ERiC project endorsed widely by research community, academic, but also professional schools, and government**
- **Goal: to follow developments, coordinate where necessary, and further development and use/implementation of methods to evaluate research in its scientific and societal context**
- **Activities: experimental studies; a guide; a ‘wiki’ website, workshops, internationalisation**
- **9 November: international expert meeting in Amsterdam**



Royal Netherlands Academy of Arts and Sciences



Meta Evaluation Committee

- 36 evaluation reports of 2005 and 2006
 - Medicine, Pharmaceutical sciences, Physics, Social Medicine, Architecture, Computer sciences, Humanities
 - Research from all 13 universities, and institutes from Academy and NWO
 - Discussions with universities, boards, research coordinators, researchers, etc.
- 1st report and list of assessment issues





MEC assessment issues

- **Disciplinary vs institutional eval.**
- **Goals and ambitions**
- **Point of view board**
- **Consequences of evaluation reports**
- **Organizational context**
- **Societal vs scientific quality**
- **Financial aspects**
- **Collaboration**
- **Peer review committees**
- **Benchmarking**
- **Inflation of scores, verdicts**
- **Relation between assessments and recommendations**





some plus issues

- **Overall: SEP flexible enough, leaves room for alternatives**
- **Goals & ambitions: Self evaluation is seen as very positive, but should become more concrete**
- **Peer review: Accent on forward looking, mission orientation, mixed composition**
- **Organizational context: decision making process more transparent, information better organised**





some minus issues

- **Societal relevance: more attention, but no method, no knowledge how to measure this**
- **Peer review: still geared towards assessment of scientific quality**
- **Verdicts: inflation of scores, not critical enough → group & network dynamics (too much solidarity)**
- **Follow up → still unclear what happens with the results**





MEC report 2007

'Trust, but verify'

First conclusion: relation between universities and government still characterized by mutual distrust

- Governmental interference to improve quality and relevance leads at best towards reallocations based on fashions**
- Universities' inertia to be more sensitive to the needs of society and politics, and an unwillingness to be transparent about its own input and output**





Towards a new SEP in 2009

Universities, Academy, Research Council work together on a new SEP

- **Include MEC comments**
- **Include comments of academic community [work load]**
- **Include comments of ERiC [valorisation, societal goals]**
- **Include comments of government [effect, consequences]**

→ Goal: lighter, more efficient, focus on consequences, valorisation



Royal Netherlands Academy of Arts and