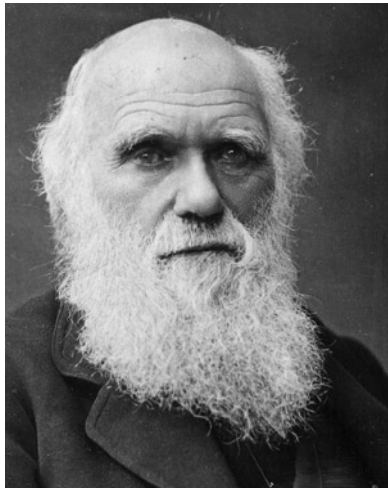


Quality of Ex-post Evaluation Studies of Funding Schemes





Why me?

 <p>JOANNEUM RESEARCH</p>	<p><u>www.joanneum.at/rtg</u></p>
 <p>W W T F</p>	<p><u>www.wwtf.at</u></p>
 <p>PLATTFORM fteval Forschungs- und Technologieevaluierung</p>	<p><u>www.fteval.at</u></p>

Terminology

- ▶ Evaluation of funding schemes \neq project selection
 - ▶ Those rules, that are just and equitable in project selection, are not automatically transferrable to programme evaluation
- ▶ Funding schemes = programmes
 - e.g. Excellence Programmes, Mobility Programmes, Thematical Programmes
- ▶ A Programme is a combination of interventions where the underlying intentions refer to each other. A programme usually is made up for a fixed time period, has its own centrally administered budget and a clear structure.

Four Dimensions of Quality

- ▶ Standards, Good Behavior
- ▶ Utility
 - ▶ Did the stakeholders learn?
 - ▶ Has the programme improved after the evaluation?
- ▶ Evaluation Planning
 - ▶ Is the evaluation timely?
 - ▶ Is all necessary data available?
- ▶ Contribution to the further Development of Evaluation
 - ▶ Did the evaluators learn?
 - ▶ Did the evaluators contribute to the evaluators' discussion on evaluation? ...on methods?

Standards and Good Behaviour

- ▶ Evaluation Standards aim to provide a set of guidelines
 - ▶ For evaluators
 - ▶ For those who are commissioning evaluations
 - ▶ For those to be evaluated
- ▶ There are different, but quite similar Standards
 - ▶ International professional associations of evaluators like American Evaluation Association
 - ▶ DeGEval: German Evaluation Society
 - ▶ Austria: Research and Technology Policy
- ▶ Sources
 - ▶ <http://www.eval.org/EvaluationDocuments/Progeval.html>
 - ▶ <http://www.fteval.at/home.php?lang=en>
 - ▶ <http://www.degeval.de/calimero/tools/proxy.php?id=72>



Behave, Evaluator!

- ▶ Utility Standards
 - ... intend to ensure that an evaluation meets the information requirements of intended users.
- ▶ Feasibility Standards
 - ...intend to ensure that an evaluation is realistic, prudent, diplomatic, and frugal.
- ▶ Propriety Standards
 - ... intend to ensure that an evaluation is conducted legally, ethically, and with due regard for the welfare of those involved in the evaluation, as well as those affected by its results.
- ▶ Accuracy Standards
 - intend to ensure that an evaluation will reveal and convey technically adequate information about the features that determine worth or merit of the programme being evaluated.

Some Words on Utility and Feasibility

- ▶ While a 'code of conduct' is a necessary prerequisite for high quality evaluation, 'utility' is it's most important characteristic
 - ▶ Evaluations should be planned, conducted, and reported in a way that encourages follow-through by stakeholders
- ▶ Horror! 500 pages and 2 weak conclusions
 - ▶ Evaluation should help to formulate the next steps
- ▶ Unrealistic advice
 - ▶ „Ways have to be found to overcome the legal issues regarding the programme“
- ▶ A useful evaluation must base on sound scientific methods
- ▶ An evaluation done on the highest scientific level is not automatically useful

Propriety and Accuracy Standards

- ▶ are some sort of „Code of Conduct“ for Evaluators
 - ▶ Evaluations should be complete and fair
 - ▶ Evaluations should respect and protect the rights and welfare of human beings
 - ▶ The sources of information should be described detailed enough
 - ▶ The conclusions reached in an evaluation should be explicitly justified, so that stakeholders can assess them
 - ▶ Conflict of interest should be dealt with openly and honestly

Low Quality in Evaluation: Whom to blame?

- ▶ Beside the incompetence of an evaluator, there might be additional factors that influence the quality of evaluation.
 - ▶ Unclear terms of references
 - ▶ Loosing interest of those who commissioned the study.
 - ▶ Fluffy programme goals
 - ▶ “changing the mindset of the Austrian people”
 - ▶ “enhancing quality of life”
 - ▶ Evaluation is not timely
 - ▶ Measuring impacts of a programme after 2 years
 - ▶ No Programme Documentation
 - ▶ Data is not available

Programme Evaluation: Typical Questions

Issue	Question
Appropriateness	Was it the right thing to do?
Economy	Has it worked out cheaper or more expensive than we expected?
Effectiveness	Has it produced the expected effects?
Efficiency	What's the return on investment (ROI)?
Efficacy	How does the ROI compare with expectations?
Process Efficiency	Is it working well?
Quality	How good are the outputs?
Impact	What has happened as a result of it?
Additionality	What has happened over and above what would have happened anyway?
Displacement	What hasn't happened which would have happened in its absence?
Process Improvement	How can we do it better?
Strategy	What should we do next?

Arnold, E. et al (1995), Good Ideas in Programme Management for Research and Technical Development Programmes, Brighton.

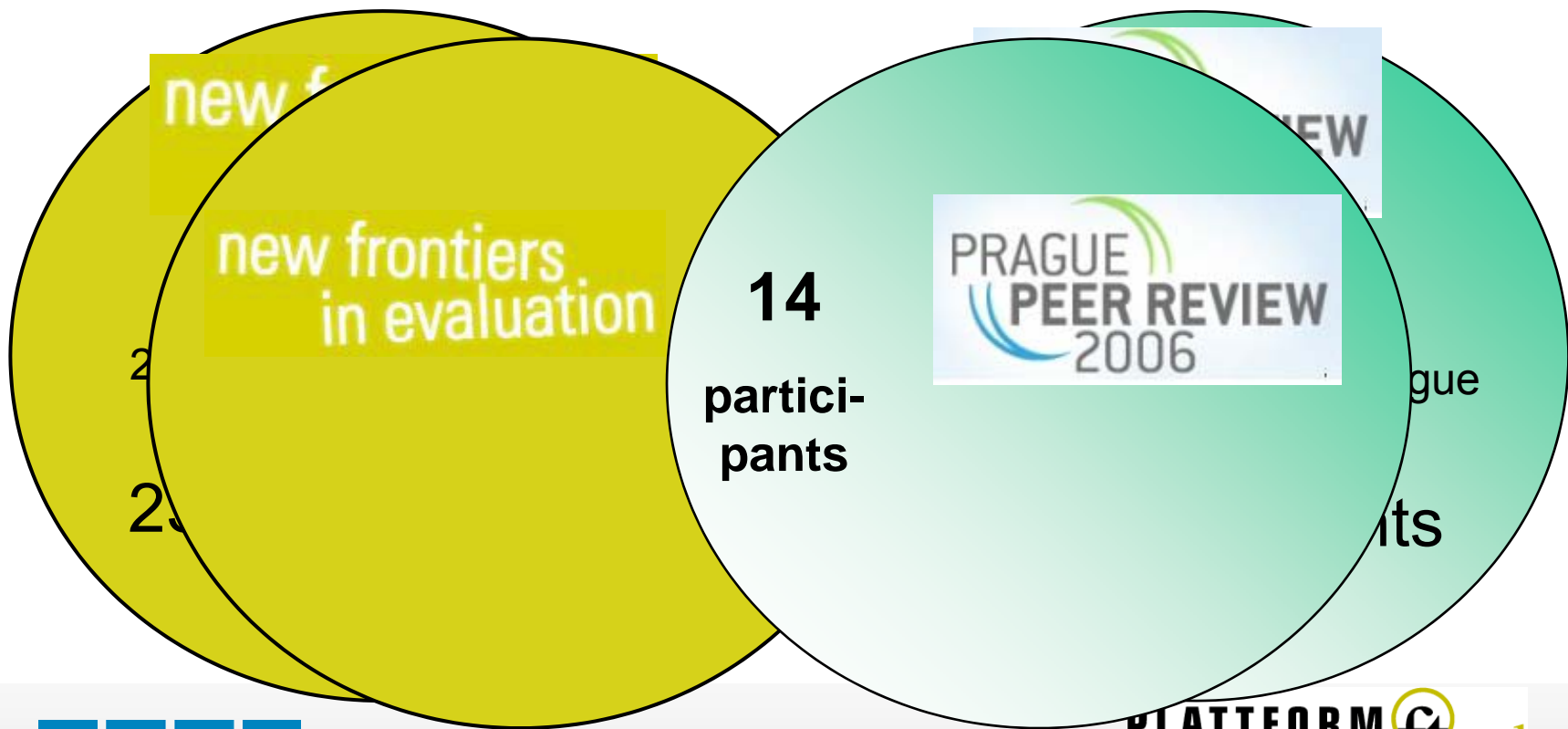
There are more Evaluation Methods than Peer Review

Interviews	Surveys	Cost Benefit Analysis	Story Telling
Logic Charts	Curricula	Focus Groups	Benchmarking
Case Studies	Control Group Approaches	Peer Review	Foresight
I/O Models	Bibliometric Methods	Social Network Analysis	Logit/ Probit Models

Peers and Evaluation Experts

- ▶ Peers are not (automatically) Evaluation Experts
 - ▶ *“Outside peer review, the principle of triangulation is fundamental to evaluation. However, as one would expect from the fact that the panellists are good and hard scientists, and not specialist evaluators or social scientists, their use of surveys is not especially advanced.”* Erik Arnold et al. *Reviews Reviewed: Lessons from the First Six International Panel Reviews, 1999 - 2004*
- ▶ Example: Evaluation of the Austrian Mathematics
 - ▶ Evaluation of Research and Teaching Programmes of the Faculties of Mathematics at Austrian Universities
- ▶ Is there sufficient exchange?
 - ▶ Professional evaluators are often seen with a lack of respect
 - ▶ If there are two communities, do they know enough about each other?

Is there ,one' Evaluation Community?



What we should know about ‚Professional‘ Evaluators

- ▶ Strong Community
 - ▶ > 2.000 participants at AEA conferences
- ▶ Strong, but highly fragmented
 - ▶ > 20 topical interest groups
 - ▶ Different cultures, different terminologies



- ▶ (Social) Scientist - Evaluator - Entrepreneur

Publications & Evaluations



	SJR (2005)	Total Documents (3 Years)	Total Cites (3 Years)	Cites x Doc. (3 years)
Research Evaluation	0,054	57	38	0,9
Research Policy	0,072	289	790	2,94
Annals of Mathematics	0,116	162	309	1,93
American Economic Review	0,136	522	1.232	3,60
Nature	7,204	7.723	83.517	21,89

Quality of Evaluation: The GEN-AU Example

- ▶ **Austrian Genome Research Programme GEN-AU**
(**GEN**ome Research in **AU**stria)
 - ▶ International Evaluation Team
 - ▶ New Methods, Methodology Mix
 - ▶ Evaluation according to the Austrian Standards

- ▶ Did the stakeholders learn?
- ▶ Is the programme better now?
- ▶ Did the evaluators learn?
- ▶ Did we contribute to evaluators' discussion on evaluation?

Quality of Evaluation: The GEN-AU Example II

- ▶ Did the Ministry learn?
 - ▶ Those who could have left the scene, but ...
 - ▶ Court of Auditors used report to spotlight some aspects afterwards
- ▶ Did the Programme Managers learn?
 - ▶ “The GEN-AU mid term evaluation contributed to
 - ▶ The agentification of the programme
 - ▶ Establishment of efficient routines (e.g. controlling, public relations)
 - ▶ Standardization and transparency
 - ▶ Preparation of an Impact Assessment Tool, starting in 2008”
- ▶ Did the evaluators learn?
 - ▶ Further collaborations
- ▶ Other effects?
 - ▶ “knowledge spillovers”
- ▶ Did we contribute to evaluators’ discussion on evaluation?
 - ▶ Presentation in the US, but, to be honest: No.

Addendum: How to define the Quality of Evaluation Studies? The Role of Meta-evaluation

- ▶ Meta-evaluation is used to assess the quality, relevance, effects, and the usage of evaluations (*Widmer 1996*).
- ▶ It serves as a preparation for and core element of the secondary analysis.
- ▶ Meta-evaluation plays a prominent part when it comes to assessing evaluation culture itself as an element of policymaking in innovation systems.
- ▶ The purpose of these evaluations can be both *formative* and *summative*.

Quality of Evaluation. Some ,hands on' arguments



- ▶ Strong opinions: All Answers in the Proposal
- ▶ “This programme has created 321 jobs and plus 17,43 Mio € turnover for Austrian companies”
 - ▶ There is no safety in numbers
- ▶ No Idea of the context
- ▶ Single Method Approaches
 - ▶ Triangulation
- ▶ Complex Analysis for free
 - ▶ Bibliometric Analyses of 50 individuals for € 3.000.—

Klaus Zinöcker

Vienna Science and Technology Fund WWTF

www.wwtf.at

Plattform fteval

www.fteval.at

klaus.zinoecker@wwtf.at

