



HOW TO EVALUATE

This pamphlet is part of a four-pamphlet series aimed at setting out the importance and role of evaluating research, and providing guidance about how research evaluations can be carried out. The first two pamphlets discuss the rationale for evaluating and the objects of research evaluation. The purpose of this third pamphlet is to introduce some key research evaluation techniques and to highlight important factors to consider when selecting evaluation approaches. A fourth pamphlet provides examples of current research evaluations undertaken by foundations.

PAMPHLET 3: HOW TO EVALUATE RESEARCH?

A variety of evaluation methodologies are available to assess research. Individual methods have their own strengths and limitations. Evaluating research effectively and efficiently therefore requires considering which methods are most appropriate for a specific evaluation context. Methods tend to fall into two broad categories: (1) broad and shallow, and (2) narrow and deep. Broad and shallow approaches aim to capture the large-scale effects or quality of research and often 'survey' large samples for information. Narrow and deep evaluations focus on understanding research processes in more depth, and zoom into questions such as how one can improve research funding decisions or accelerate the translation of research findings into new products, technologies, services and practices. Below we provide an overview of some key evaluation methods.

A RESEARCH EVALUATION TOOLKIT

The techniques associated with research evaluation and evaluation more generally are applications of traditional social research methods such as surveys, key-informant interviews and statistical analysis of quantifiable data. Key techniques that are used in research evaluations include:

1. **Bibliometrics:** allows measurement of scientific outputs and outcomes, drawing on information on publications and citations by means of statistical methods.
2. **Case studies:** They are generally based on multiple sources of evidence, which all feed into deriving conclusions from an evaluation, and are used to test confidence in the conclusions. The main sources of evidence include peer-reviewed literature and so-called 'grey literature' (publications by government,

industry and academia that are distributed outside the normal publishing channels) and archival documents, semi-structured key informant interviews which can also be complemented by surveys, and at times bibliometric databases or focus groups.


3. **Peer review (and expert panels):** a process of evaluation involving qualified individuals within the related field to reflect on research(er) outputs and impacts ex-post or ex-ante to funding decisions. Peer review is employed with the goals of maintaining research standards and providing credibility in research decision-making.
4. **Surveys and consultations:** used to collect quantitative and/or qualitative information about items in a population; they may focus on opinions or factual information depending on their purpose; they often involve administering questions to individuals.
5. **Economic analyses:** analyses relying on economic indicators to assess the outputs, outcomes and impacts of research:
 - a. **Micro-econometric analysis and modelling** allow estimation of outputs, outcomes and impacts at an individual or sector level;
 - b. **Macro-economic analysis and modelling** allow estimation of broader socioeconomic impacts of policy interventions at an aggregate or national level.

The table below describes the characteristics, advantages and disadvantages of these standard research evaluation methods.

TABLE 1: SOME KEY RESEARCH EVALUATION METHODS

EVALUATION METHODS	CHARACTERISTICS	ADVANTAGES	DISADVANTAGES
BIBLIOMETRIC ANALYSES	Can be narrow and deep or broad and shallow	<ul style="list-style-type: none"> > Quantitative measuring of volume output > Can be used to indicate quality of output > Enables analysis of global trends 	<ul style="list-style-type: none"> > Estimates of quality based on citations alone can be misleading > Data must be normalised to enable comparisons across research fields and journals > Does not measure future potential
CASE STUDY ANALYSES	Narrow and deep	<ul style="list-style-type: none"> > Provides in-depth analysis of the process of discovery > Can demonstrate pathways from basic science to application > Widely applied > Broadly accepted > Flexible and adaptable 	<ul style="list-style-type: none"> > Selection bias: how to know that the chosen cases are representative > Highly resource-intensive to do well > Can be difficult to generalise from
SYSTEMATIC PEER REVIEW	Narrow and deep	<ul style="list-style-type: none"> > Well-understood component of research management > Widely accepted 	<ul style="list-style-type: none"> > Time-consuming for experts involved > Concerns over the objectivity and reliability of findings
SURVEYS AND CONSULTATIONS	Can be narrow and deep or broad and shallow	<ul style="list-style-type: none"> > Can identify outputs and outcomes associated with particular pieces of funding research > Provides qualitative analysis of outcomes 	<ul style="list-style-type: none"> > Dependent on contact details being available for researchers in question > Poor response rate can limit findings
ECONOMIC RATE OF RETURN 1: MICRO-ECONOMIC ANALYSIS	Broad and shallow	<ul style="list-style-type: none"> > Can be applied to different sectors > Comparative potential, e.g. cost-benefit analysis 	<ul style="list-style-type: none"> > Difficult to put a financial value on many of the influences involved
ECONOMIC RATE OF RETURN 2: MACRO-ECONOMIC ANALYSIS	Broad and shallow	<ul style="list-style-type: none"> > Quantitative > Provides 'big picture' and context of research 	<ul style="list-style-type: none"> > Difficult to identify the contribution of an individual sector or funder

Source: Ismail, Nason, Marjanovic and Grant (2009), adapted from UK Evaluation Forum (2006)



The above advantages and disadvantages need to be considered when selecting appropriate methods for your evaluation purposes and context. For example, case studies may be less suitable when time and funding for evaluation is limited, and bibliometrics may not fit well for evaluating research in fields with poor journal coverage in bibliometric databases. Since individual methods have limitations, it is worth considering whether combining some of them is feasible and how this may improve the overall quality of the research evaluation.

Some other techniques used in research evaluations include:

- > Benchmarking - allows performance comparisons based on a relevant set of indicators
- > Cost-benefit analysis - allows establishment of whether a policy, programme or project is economically efficient by appraising its economic and social effects
- > Public value assessment - review of the perceived value to different stakeholders of the knowledge generated
- > Network Analysis - allows analysis of the structure of cooperation relationships which are sometimes a measure of outputs from a research project that are interesting to capture. These networks can be with diverse stakeholder groups
- > Logic modelling - used to capture the logical flow between inputs, processes, outputs, outcomes and impacts of research (inherent in some research evaluation frameworks such as the Payback framework for health and biomedical research evaluation)
- > Foresight type assessment - used to identify potential mismatches in the strategic efficacy of projects, programmes and/or policies

GLOSSARY OF TECHNICAL TERMS

> **ATTRIBUTION:** Attribution involves drawing causal links and explanatory conclusions about the relationship between observed changes (whether anticipated or not) and specific interventions. In evaluation it is typically a matter of attributing outcomes and impacts to a research project or programme.

> **BIBLIOMETRICS:** A generic term for data about publications, including which publications are cited by other publications (citation data), the number of times they are cited and so on.

> **CROSS-SECTIONAL EVALUATION:** A cross-sectional study is one that takes place at a single point in time. In effect, we are taking a 'slice' or cross-section of whatever it is we're observing or measuring.

> **INDICATOR:** A quantitative or qualitative factor or variable that provides a simple and reliable means to measure achievement, to reflect changes connected to an intervention, or to help assess the performance of an actor.

> **IMPACT:** In these guidelines, impact refers to any changes in academia, economy and the wider society attributable to a research project and programme and its outputs. It is used interchangeably with outcome.

> **IMPACT FACTOR:** In its simplest form, this is a score assigned to academic journals based on the average number of citations an article in that journal receives over a fixed period of time.

> **LONGITUDINAL EVALUATION:** A longitudinal evaluation is one that takes place over time - we have at least two (and often more) waves of measurement in a longitudinal design.

> **MONITORING** is the systematic collection and analysis of information as a project progresses. It is aimed at improving the efficiency and effectiveness of a project or organisation. It is based on targets set and activities planned during the planning phases of work. It helps to keep the work on track, and can let management know when things are going wrong.

> **OUTCOME:** In these guidelines outcomes refer to long term changes to an academic field or discipline or to impacts in wider economy and society, such as improved productivity or health benefits, that can be attributed to a research project or programme and its outputs. In some evaluation literature, outcomes are an intermediate result that can be identified independently of longer-term impact, but we make no such distinction here.

> **OUTPUT:** In these guidelines outputs refer to the direct, measurable results of a research project or programme such as publications, conferences and patents. This is primarily concerned with evidence of new knowledge produced by the scientific community. In addition, 'outputs' can refer to additional resources or research capacity that is a consequence of the funding, such as increased finance or new PhD students.

WANT TO KNOW MORE?

Suggested further reading on research evaluation

The following list contains all the references and sources material used in the writing of these guidelines, as well as suggested reading for those who wish to go into more detail. We have favoured material that is explicitly aimed at those with no prior knowledge of the subject area, but some of the texts are aimed at a more specialist audience.

GENERAL BACKGROUND

Brutscher, Ph.-B., Grant, J., Wooding, S.: Health research evaluation frameworks, RAND report RAND/TR-629

Clarke, Alan Evaluation Research: An Introduction to Principles, Methods and Practice, Sage Publications (London 1999)

Evaluating the societal relevance of academic research: A guide, ERiC publication (June 2010) http://www.eric-project.nl/nwohome.nsf/pages/NWOA_6TZJ28_Eng

Marjanovic, S., Hanney, S. and Wooding, S (2010). A historical reflection on research evaluation studies, their recurrent themes and challenges. RAND Report: rand/tr-789-rs

Marjanovic, S (2009) The Payback Framework in: Ling, T and Vilalbia van Dijk, L.(Eds): Performance Audit Handbook. RAND Europe TR-788-RE

Trochim, William M. Introduction to Evaluation in The Research Methods Knowledge Base, 2nd Edition. (2006) <http://www.socialresearchmethods.net/kb/intreval.php>

BEST PRACTICE SURVEYS

Evaluation in National Research Funding Agencies: approaches, experiences and case studies

European Science Foundation, A report of the ESF Member Organisation Forum on Ex-Post Evaluation of Funding Schemes and Research Programmes (September 2009) <http://www.esf.org/activities/mo-fora/completed-mo-fora/evaluation-of-funding-schemes-and-research-programmes.html>

Grant et al Capturing Research Impacts. A review of international practice, HEFCE Documented briefing (December 2009) http://www.hefce.ac.uk/pubs/rdreports/2009/rd23_09/rd23_09.pdf (26.03.2010)

Ruegg, R. and Feller, I., 2003. A Toolkit for Evaluating Public R&D Investment: Models, Methods and Findings from ATP's First Decade.

BIBLIOMETRIC ANALYSIS

Ismail, S., Nason, E., Marjanovic, S and Grant, J. Bibliometrics as a tool for supporting prospective R&D decision-making in the health sciences', RAND Report : RAND/TR-685

Van Leeuwen TN Modelling of bibliometric approaches and importance of output verification in research performance assessment, Research Evaluation, 2007; 16(2):93 - 105

Using Bibliometrics: A Guide to Evaluating Research Performance with Citation Data, Thomson Reuters (2008) http://science.thomsonreuters.com/m/pdfs/325133_thomson.pdf

ECONOMIC ANALYSIS

Exceptional Returns: The Economic value of America's Investment in Research, Lasker Foundation <http://www.laskerfoundation.org/advocacy/pdf/exceptional.pdf>

Ex Post Evaluation of Economic Impacts of Agricultural Research Programs: A Tour of Good Practice, Maredia et al. Paper presented to the Workshop on "The Future of Impact Assessment in CGIAR: Needs, Constraints and Options" Rome, May 3-5 2009; <http://impact.cgiar.org/methods/docs/maredia.pdf>

IMPACT EVALUATIONS

Cox et al, Evaluation of Impacts of Medical Research, Swedish Research Council (2009) <http://www.vr.se/download/18.5adac704126af4b4be2800026596/Evaluation+of+impacts+of+medical+research.pdf>

Making an Impact: A Preferred Framework and Indicators to Measure Returns on Investment in Health Research, Canadian Academy of Health Sciences (2009) http://www.caahs-acss.ca/e/pdfs/ROI_FullReport.pdf

Medical research: assessing the benefits to society, A report by the UK Evaluation Forum, supported by the Academy of Medical Sciences, Medical Research Council and Wellcome Trust (May 2006) http://www.wellcome.ac.uk/stellent/groups/corporatesite/@msh_publishing_group/documents/web_document/wtx031817.pdf

ACKNOWLEDGEMENTS

The European Foundation Centre wishes to acknowledge the generous financial support of the Lundbeck Foundation in the 3 year research evaluation initiative that led to these documents. The drafting of these guidelines and case-studies has benefitted from the input of a wide range of individuals and organisations. However, particular mention needs to be made of the following individuals, whose unstinting generosity with their time and expertise has improved the documents at every step:

- > Dr Liz Allen, Senior Evaluation Adviser, Wellcome Trust
- > Dr Atje Drexler, Deputy Head of Science and Research, Robert Bosch Stiftung
- > Dr Anne-Marie Engel, Director of Research, Lundbeckfonden
- > Dr Sonja Marjanovic, Senior Analyst, RAND Europe

Finally, the EFC would like to acknowledge the contribution of the Adessium Foundation toward the design and printing of the material.



European Foundation Centre, AISBL | 78, avenue de la Toison d'Or | 1060 Brussels, Belgium t +32.2.512.8938 | f +32.2.512.3265 | efc@efc.be | www.efc.be

LUNDBECKFONDEN