



PLATTFORM TECHNOLOGIE EVALUIERUNG

Nr. 1

Eine Initiative von

Bundesministerium für Wissenschaft, Verkehr und Kunst
WIFO Österreichisches Institut für Wirtschaftsforschung
JOANNEUM RESEARCH Forschungsgesellschaft mbH

Plattform Technologie-Evaluierung - ein Vorwort

Die Plattform Technologie-Evaluierung ist eine lose Abfolge von Veranstaltungen / Workshops sowie eine informelle Publikationsreihe zu Fragen der Evaluierung v.a. von Technologieprogrammen, -projekten und -konzepten. In diesem offenen Prozeß, der im wesentlichen Impulse für die österreichische Diskussion geben soll, geht es uns vor allem um folgendes:

- Die Evaluierung von staatlichem Mitteleinsatz in der FTE-Förderung ist noch nicht selbstverständlich genug. Es wurde eine Reihe von einschlägigen Arbeiten in den letzten Jahren durchgeführt; ein zweiter Schritt ist nun das Erreichen eines besseren gemeinsamen Verständnisses zwischen Planern, Förderern und Evaluatoren.
- Damit verbunden ist die Diskussion und Entwicklung von gemeinsamen *best practices* und die Diskussion von Methodenfragen, um zu einer besseren Vergleichbarkeit und zu avancierteren Untersuchungstechniken zu gelangen.
- Durch die Verbreitung von Expertenwissen aus anderen europäischen Ländern soll ein Dialog auch über unsere Grenzen hinaus geführt werden; damit ist auch die kritische Übernahme von Methoden und Erfahrungen möglich.
- Schließlich setzt die Europäische Union mit ihrer eigenen Evaluierungsmaschinerie Standards, deren Geltung / Bedeutung für Österreich zumindestens überprüft und diskutiert werden sollte.

Die Wirtschaftssektion des Bundesministeriums für Wissenschaft, Verkehr und Kunst bildet für die Veranstaltungen die räumliche Plattform; die Einrichtung eines "Redaktionskomitees" ist bei entsprechendem Interesse geplant. Der Newsletter selbst als offene inhaltliche Plattform wird - meist angekoppelt an die erwähnten Workshops - an Vertreter der Technologieressorts, an Förderinstitutionen und in- und ausländische Wissenschaftler versendet. Wir hoffen damit - mit Ihrer Mitwirkung - eine spannende Debatte zu verstärken und einen Erfahrungsaustausch anzuregen. In Zeiten knapper Mittel scheint dies uns ein geeigneter und gemeinsamer Schritt zur Reflexion zu sein.

In dieser ersten Ausgabe berichtet der britische Evaluierungsspezialist ERIK ARNOLD über seine Erfahrungen mit der Technologieevaluierung und leitet daraus *best practices* - Vorschläge ab. Demnächst erscheinen Beiträge zur österreichischen Situation und zur FTE-Evaluierung der Union.

English abstract: This newsletter is the first of a planned serial about problems and methods in the field of technology policy evaluation. The aim is to improve the discussion between policy makers, researchers and funding institutions. In this issue Erik ARNOLD, a Brighton - based evaluation specialist, reports about his experiences in technology policy evaluation.

Dorothea Sturn, Joanneum Research
Gernot Hutschenreiter, WIFO
Michael Stampfer, BMWVK

Developing an Evaluation Culture in Austrian Science and Technology Policy¹

Erik Arnold
Technopolis
September 1996

Summary

Evaluation is an important practice in the management of science and technology, as it is in other activities of the state. Actions should be evaluated both to provide accountability to the taxpayer and to learn how to improve performance over time. In Europe, the larger and more northern economies make the most extensive use of evaluation while the southern countries do comparatively little. Austria occupies an intermediate position. Moving towards best practice depends on relating the evaluation function to the national science and technology policy system and creating elements of a 'learning organisation' among responsible arms of government.

What is evaluation?

The terminology of evaluation is evolving and is not yet standardised either within or between European languages. The growth of a community of evaluators over the past decade has created a group with a professional interest in clarifying definitions - but also in extending the definition of 'evaluation' as widely as possible.

In principle, actions can be evaluated

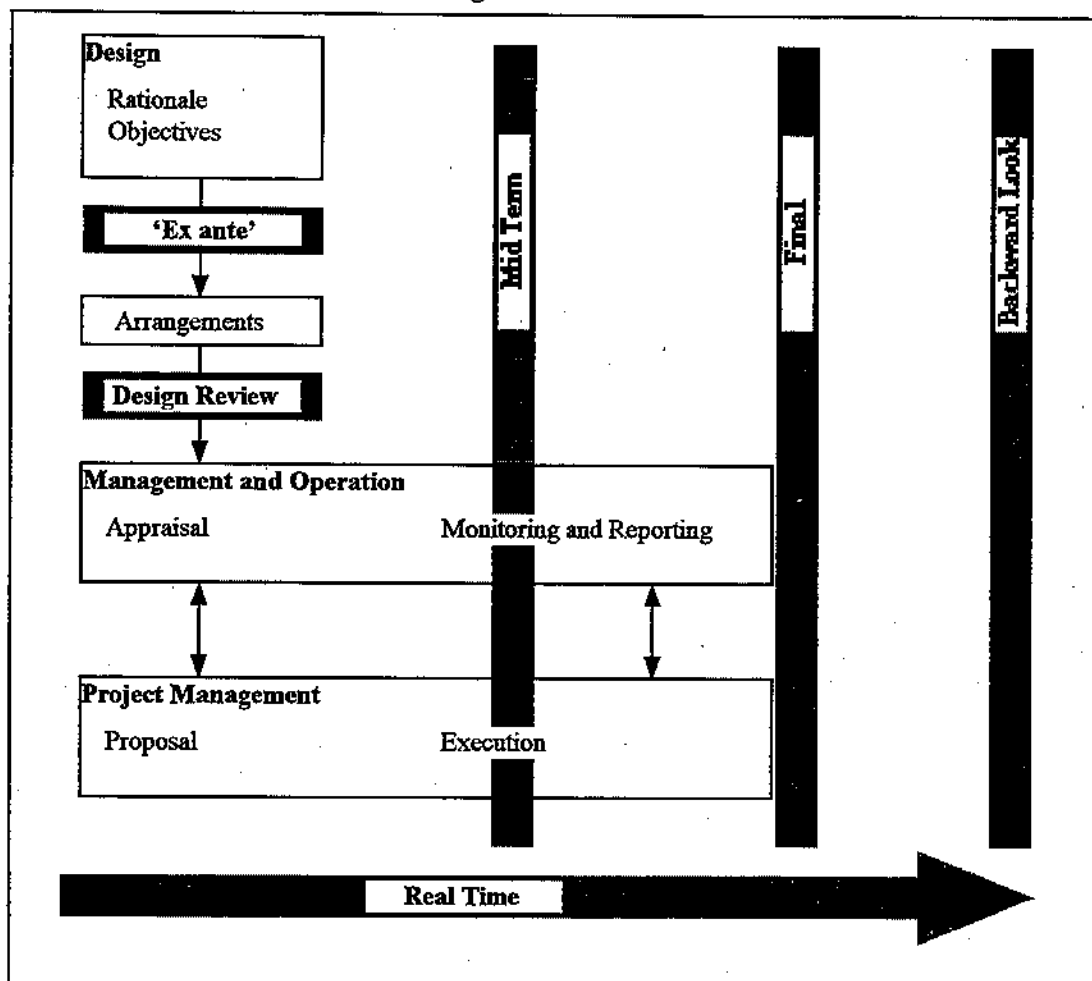
- **Ex ante:** testing whether the action is appropriate - that is, based on a correct problem diagnosis and checking whether the objectives proposed for the action are consistent with meeting the need identified
- **Interim:** testing whether the action is well implemented and identifying opportunities for operational improvement. Interim evaluation may also cover appropriateness of the action
- **Ex post:** reviewing the appropriateness, implementation efficiency, effectiveness and impact of the action

¹ This paper is based on an informal presentation given at the Federal Ministry of Science, Transport and the Arts in Vienna on 4 September 1996

Quality judgements (especially judgements about the quality of the science involved) are often important in interim and ex post evaluations.

Exhibit 1 shows how, in practice, evaluation can relate to the life cycle of an action - typically a programme or an institution.

Exhibit 1 Evaluation in the Management of RTD Actions



The most natural time to evaluate is at the end of an action. However, since a key objective is usually to decide what to do next or whether to change the course of existing policy, interim or mid-term evaluations are probably the commonest. Their timing means it is hard to say much about impact - and therefore about whether the actions evaluated are meeting their objectives. Therefore, they often focus on process questions - 'how to perform better' - rather than on questioning the rationale or effectiveness of actions. They may also examine whether actions continue to be appropriate, given the passage of time since they were initiated.

The other types of evaluation are all less usual. Final evaluations are comparatively rare. Even rarer are 'backward look' evaluations of programmes. These are conducted long enough after the end of the programme to allow at least some impacts to be realised in practice. Ex ante evaluations are also rather unusual - even more so when they incorporate a 'design review' which examines the intended mechanisms for implementing the action and compares it with experience and good practice in existing actions.

There are three distinct traditions in science and technology evaluation

- **Peer review** where scientists make judgements about each other's work or proposals, based on their scientific knowledge. This is chiefly used for assessing scientific quality
- **Summative** where evaluators try to sum up the performance of an action or an institution. This is a little like being a judge in the ice-dancing Olympics - awarding marks, but not getting involved
- In contrast, **formative** evaluators act more like sports coaches - not only making judgements about performance but also helping to improve it

Scientific peers do all three kinds of evaluation. However, they are often amateurs when it comes to tackling the methodological issues involved in evaluating non-technical aspects of actions such as policy relevance or economic impact. Professional evaluators focus on summative or formative evaluation. In general, they produce quality judgements by delegating these to panels of scientific peers or by using indicators, such as publications or patenting (though using these indicators is full of methodological difficulties). A mixed team involving scientists but led by professional evaluators often works well. The right balance of skills depends on the needs of those who commission the evaluation.

Why evaluate?

The practice of evaluation is spreading, both in order to account for the way the state spends taxpayers' money and to help the policy system to improve its performance by learning from experience.

In most countries, state expenditure is overseen by a national Audit Office, which checks that financial regulations are followed, looks for fraud, and often does some sort of policy evaluation to test the efficiency with which taxpayers' money is spent. Since an Audit Office's core skill is usually accounting, evaluations tend to use cost-benefit or cost-effectiveness approaches. These work adequately with simple policy questions where costs and benefits can easily be measured and modelled. However, science, technology and innovation actions tend to involve significant uncertainties. Their effects are often indirect, and are difficult to measure, so accounting-based approaches are not very useful. Hence, most

evaluations set out to provide 'accountability' in the use of taxpayers' money are done outside national Audit systems.

Evaluation helps policy makers learn from experience at two levels. First, at the policy level, understanding whether an action is reaching its goals is an important input to deciding whether to continue the action or to take another approach. Second, at the operational level, evaluation can provide important management information about the need for 'mid-course corrections' - allowing increases in effectiveness and efficiency by using experience as a basis for changing procedures. More broadly, evaluation documents experience so that this can be fed forward into the design of new actions. Evaluation should be part of the process of building 'intellectual capital' in the organisations which design and deliver policy.














Working with an international evaluation team can bring particular benefits: the evaluators are more easily seen to be independent than national evaluators. An international team can also bring perspectives and experience from a number of countries, so that the evaluation allows policy makers to learn both from analysis of 'their own programmes' experience and from the experience of others. At the national level, this implies that countries should develop their own evaluation professionals, but that these people should work abroad as well as at home.




European evaluation practices

The spread of evaluation has to be understood partly as a process where policy makers assume increasing control over the direction of national science and technology efforts. Often, state funding was in the past targeted at individual projects. In more recent years, projects have been grouped into themes or programmes, allowing clearer policy choices. State laboratories are less and less allowed to determine their own directions in spending state money. Some countries are introducing the 'customer-contractor' principle, which explicitly separates the part of the state which pays for research from other parts - such as the national laboratories - which perform research. In the UK, this logic is being further extended: some state laboratories are being privatised, enabling the state to buy research on more competitive markets. Evaluation is a key tool in this process of clarifying and managing science and technology policy.

Exhibit 2 gives an overview of the spread of evaluation culture in Europe. The Exhibit is of course highly simplified: each country has its own, often complex, history and process of development.

Exhibit 2 Some European Science and Technology Evaluation Practices

Country	Start of Evaluation	Evaluation Responsibility	Focus	Evaluative Community
France	1982/90	CNE: Universities CNER: Other Report to President	Mainly institutions Increasingly programmes	
Germany	1970s	Wissenschaftsrat: Universities BMFB etc: Other BMFB includes internal evaluation group	Programmes Institutions	
UK	Mid-80s	HEFCE: Universities Policy Departments: Other DTI includes internal evaluation group	Universities Programmes	
Netherlands	1982/86	Universities' Association Policy Departments	Scientific disciplines Institutes Programmes	
Austria	Late-80s	Policy departments	Some university and programme activities	
Ireland	Late-80s	Universities: None Other: Forfás; Department of Enterprise & Employment	Structural Funds programmes	
Greece	1990	Universities: None Other: Policy departments	Structural Funds programmes	
Italy	1994	Universities: Internal Other: Little	Little evaluation done	
Spain	-	Universities: Not required Other: Little	Little evaluation done	
Portugal	-	Legal basis exists, but apparently little practice	Structural Funds programmes	
Sweden	Mid-80s	Research Councils	All institutions, programmes	
Norway	Early-80s	Research Council of Norway	All institutions, programmes	
Finland	Late-80s	Academy of Finland TEKES	Most institutions, programmes	

Key  0 - 5 professionals  6 - 10 professionals  >10 professionals

In many of the European countries which now have established an evaluation culture, the movement started in the universities as an extension of the normal, internal quality control procedures with which scientists work. In some cases such as the Netherlands and Italy, universities or associations of universities have taken control of this process. In others such as France and the UK, the control is more central - which allows evaluation to be more directly connected to the allocation of resources. (The UK has probably gone furthest in this respect.) In the Latin countries, the universities have conceded little or nothing to external evaluation and quality control measures.

Evaluation of technology programmes and institutions is more typically the responsibility of Ministries of Research or Industry, and their agencies. Here, the European Commission has been an important vector of evaluation culture. Structural Funds carry an obligation of evaluation. The Commission has also organised networks of evaluators and policy makers, to spread the message of evaluation by example and persuasion.

Evaluation has become a routine part of science and technology policy making in the northern part of Europe, while there is still little interest and practice in the south. Austria is both geographically, and in terms of practice, in between. A useful indicator of the state of development is the number of professional evaluators based in a country. Our estimates are shown as part of **Exhibit 2**. In this context, by 'professional evaluator' we mean someone with experience of defining and leading high-quality evaluations of science and technology programmes. We have not tried to normalise for the size of the countries. However, it is clear that there are particularly strong cultures in the UK, Germany, the Netherlands, France and Norway. The other Nordic countries also have a strong evaluation tradition, but often deliberately use foreigners in order to ensure independence and this constrains the size of the national community.

Those countries with most experience in evaluation are beginning to question whether all actions should be routinely evaluated. Some funders have taken the position that evaluation effort should be focused on actions likely to produce general policy lessons. Others believe that the need for accountability makes it important to evaluate everything. Overall, there is a movement towards systematising evaluation practices, while recognising the need to leave space for evaluators to innovate. Increasingly, funders are considering how to increase the role and value of routine monitoring, so as to provide earlier and better signals to policy makers about the performance of actions.

Creating an evaluation culture in Austria

Evaluation is not done for its own sake. An important precondition for doing evaluation is that there should be an interest in improving the performance of the science and technology system. Thus, introducing evaluation often goes along with increasing the systematisation of programme planning and creating more explicit missions and objectives for institutions. Next steps in Austria should include

- Understanding good international practice in evaluation and management of programmes and institutions
- Internalising these practices by embedding them in policy design
- Taking the opportunity to move directly to best practice by immediately considering the role of routine evaluation and the desirable balance between monitoring and evaluation
- Documenting, sharing and discussing lessons from evaluation among relevant policy makers and other interested people

Dr Arnold is a co-Director of Technopolis - a research company based in the UK and Netherlands. Technopolis works internationally on questions of science, technology and industrial development policy. Major activities include evaluating science and technology programmes and institutions, programme planning and implementation, regional and industrial development and policy studies.

für den Inhalt verantwortlich:
Bundesministerium für Wissenschaft,
Verkehr und Kunst - Dr. Michael Stampfer
Remgasse 5 - A - 1010 Wien
e-mail: michael.stampfer@bmw.ada.at

weitere Kontaktadressen:

Dr. Dorothea Sturn
Joanneum Research
Wiedner Hauptstraße 76
A - 1040 Wien
e-mail: sturn@pbox.joanneum.ac.at

Mag. Gernot Hutschenreiter
WIFO
Arsenal Objekt 20; PF 91
A-1031 Wien
e-mail: hutsch@wsr.ac.at

"Plattform Technologieevaluierung" ist ein unregelmäßig erscheinendes offenes Forum zur
Diskussion methodischer und inhaltlicher Evaluierungsfragen in der Technologiepolitik.