PLATTFORM fteval
Forschungs- und Technologieevaluierung

Evaluated – what next?

Nr. 33

Nov. 2009

BUNDESMINISTERIUM FÜR WISSENSCHAFT UND FORSCHUNG
Minoritenplatz 5, A-1014 Wien
Mag. Irene Danler
irene.danler@bmwf.gv.at
Mag. Simone Mesner
simone.mesner@bmwf.gv.at

JOANNEUM RESEARCH FORSCHUNGSGESELLSCHAFT MBH
Haus der Forschung, Sensengasse 1, A-1090 Wien
Mag. Wolfgang Polt
wolfgang.polt@joanneum.at
Mag. Michael Dinges
michael.dinges@joanneum.at

BUNDESMINISTERIUM FÜR VERKEHR, INNOVATION UND TECHNOLOGIE
Renngasse 5, A-1010 Wien
Dr. Rupert Pichler
rupert.pichler@bmvit.gv.at

KMU FORSCHUNG AUSTRIA
Gusshausstraße 8, A-1040 Wien
Dr. Sonja Sheikh
s.sheikh@kmuforschung.ac.at

BUNDESMINISTERIUM FÜR WIRTSCHAFT, FAMILIE UND JUGEND
Stubenring 1, A-1010 Wien
Mag. Sabine Pohoryles-Drexel
Sabine.Pohoryles-Drexel@bmwfj.gv.at

LUDWIG BOLTZMANN GESELLSCHAFT
Operngasse 6, A-1010 Wien
Mag. Claudia Lingner
c.lingner@ludwigboltzmanngesellschaft.at

AQA – ÖSTERREICHISCHE QUALITÄTSSICHERUNGSAGENTUR
Wickenburggasse 26, A-1080 Wien
Mag. Alexander Kohler
alexander.kohler@aqa.ac.at

RAT FÜR FORSCHUNG UND TECHNOLOGIEENTWICKLUNG
Pestalozzigasse 4/DG 1, A-1010 Wien
Mag. Peter Lindner
p.lindner@rat-fte.at

AIT AUSTRIAN INSTITUTE OF TECHNOLOGY GMBH
Donaucitystraße 1, A-1220 Wien
Mag. Petra Wagner-Luptacik
petra.wagner@arcs.ac.at
Dr. Susanne Giesecke
susanne.giesecke@arcs.ac.at

TECHNOPOLIS FORSCHUNGS– UND BERATUNGSGESELLSCHAFT MBH
Rudolfplatz 12/11, A-1010 Wien
Mag. Katharina Warta
warta@technopolis-group.at

AWS – AUSTRIA WIRTSCHAFTSSERVICE GESELLSCHAFT MBH
Ungargasse 37, A-1030 Wien
Dr. Sonja Hammerschmid
s.hammerschmid@awsg.at

WIFO – ÖSTERREICHISCHES INSTITUT FÜR WIRTSCHAFTSFORSCHUNG
Arsenal, Objekt 20, PF 91, A-1103 Wien
Dr. Rahel Falk
rahel.falk@wifo.ac.at

CDG – CHRISTIAN DOPPLER FORSCHUNGSGESELLSCHAFT
(Haus der Forschung) Sensengasse 1, A-1090 Wien
DI Brigitte Müller
mueller@cdg.ac.at

WWTF – WIENER WISSENSCHAFTS– FORSCHUNGS– UND TECHNOLOGIEFONDS
Währinger Straße 3/15a, A-1090 Wien
Dr. Michael Stampfer michael.stampfer@wwtf.at
Mag. Klaus Zinöcker klaus.zinoecker@wwtf.at

FFG – ÖSTERREICHISCHE FORSCHUNGSFÖRDERUNGS-GESELLSCHAFT (Haus der Forschung)
Sensengasse 1, A-1090 Wien
Mag. Klaus Schnitzer klaus.schnitzer@ffg.at
Mag. Mariana Karepova mariana.karepova@ffg.at
Mag. Leonhard Jörg leonhard.joerg@ffg.at

ZIT – ZENTRUM FÜR INNOVATION UND TECHNOLOGIE GMBH
Ebendorferstraße 4/DG 1, A-1010 Wien
Robert Mayer-Unterholzner
robert.mayer@zit.co.at

FWF – FONDS ZUR FÖRDERUNG DER WISSENSCHAFTLICHEN FORSCHUNG (Haus der Forschung)
Sensengasse 1, A-1090 Wien
Dr. Rudolf Novak
novak@fwf.ac.at

ZSI – ZENTRUM FÜR SOZIALE INNOVATION
Linke Wienzeile 246, A-1150 Wien
Dr. Klaus Schuch
schuch@zsi.at

Preface

Between 2003 and 2007 more than 60 evaluations were conducted in Austria. These exercises are collected and published in the book "Evaluation of Austrian Research and Technology Policies", published by the Austrian Council and the Platform fteval. This helped to increase transparency, both in methodology as well as in results. Many say this was a big step forward in the culture of evaluation, others call it “evaluities” – such as “we are evaluated to death”.

Fact is that in times of crisis and shortages in budget, evaluation gains importance. Evaluation is today well embedded in the Austrian R&D and innovation system. However, an international report (CREST) stated that evaluations are of little utility if their results fail to inform future policy and practice. More thought should therefore be given to the mechanisms needed to ensure that the results of evaluations do feed back into policy formulation and implementation.

“Evaluation – what next?” was a one day conference in March 2009, which was jointly hosted by the Platform and the Council. How can evaluation results and recommendations be integrated into the policy cycle to shape research politics and the use of public money? What is the role of evaluation as a learning instrument and in the context of the governance system? Which factors can improve the integration of evaluation results in programmes and in design and policy of programmes?

The opening speech of the conference was given by Prof. Dervilla Donnelly, Member of the Austrian Council. Further contributions in this newsletter comprise a summary of Paul Cunningham’s speech by Petra Wagner, who asked some further questions to the chief scientific advisor of the Irish government. Martin Weber from the European Court of Auditors gives an insight in evaluations in the EU context from an auditor's perspective. Julia Schmidmayer summarises the contributions of all participants. Hans Peter Lorenzen was participant of the Round Table discussion lead by Klaus Schuch and gives in his article recommendations for clients of evaluation. In the discussion Brigitte Tiefenthaler, Dorothea Sturn, Hans Peter Lorenzen participated as discussants, the open chair was used by Mariana Karepova, Rudi Novak, Wolfgang Polt. Michael Dinges contributes his statement on InnoApraisal during the discussion in an essay to this newsletter.

Authors:

Michaela Topolnik

Co-ordinator of Evaluation day 2009

Former Austrian Council

michaela.topolnik@spoe.at

Klaus Zinöcker

WWTF Vienna Science and Technology Fund

Währingerstr. 3/15a, 1090 Wien

klaus.zinoecker@wwtf.at

Announcement **4. 12.2009, 9:00 – 16:00**
Die Systemevaluierung. Anstoß zu einer Neuausrichtung der angewandten Forschungspolitik?
Camineum der Österreichischen Nationalbibliothek, Josefsplatz 1, 1010 Wien

Dervilla Donnelly

"Evaluated – what next?"

Welcome Address given at the conference "Evaluated – and then".

The conference was jointly hosted by the Platform Research & Technology Policy Evaluation and the Austrian Council. The opening speech was given by Prof. Dervilla Donnelly, Member of the Austrian Council (Rat für Forschung und Technologieentwicklung) since 2000.

The main focus of our discussions today is to talk about and take advantage of the results obtained through the evaluation of Austria's RTI-Policy.

What are the possibilities? What are the experiences in other countries, especially EU-countries? Is Austria ahead of its neighbours? In times of deficit spending, will RTI policy be able to compete with other areas like education or health policy? These are essential questions that seek answers in events like this conference.

Evaluation is a key tenet of the Austrian Council's mission and it attaches great importance in advancing and establishing a culture of evaluation. The Council has the duty to make more demands in this area and actively cooperates with its partner the Platform Research and Technology Policy Evaluation. The Platform FTEval was founded in 1996, has so far presented methods and approaches to evaluation - especially benchmarking - in an international comparison, and has together with the decision makers developed a culture of evaluation. Evaluation is to achieve increased improvement and more transparent assessments in order to assist optimal strategic planning of R&D policy. When the Platform thinks about evaluation it thinks beyond project selection and the evaluation for journals. While peer review is an important method to assess scientific quality, the platform sees and offers a lot of different methodological possibilities in evaluation such as logic charts, cost-benefit-analysis or questionnaires to name only a few examples.

The strategy of the Council was geared to the global objective of strengthening competitiveness and dynamism of the Austrian economy in order to lay out the basis for sustainable economic growth and an increase of the level of employment. An assessment of the progress Austria has made in terms of its research, technology and innovation policy prior to the publication of the

Council's 2010 Strategy was based on the success of the Council's National Research and Innovation Plan. The picture was gratifying when we looked at the results of the plan – the research quota had risen from about 1.0% in 1981 to a projected 2.73% in 2009 and was above the EU average. The structural reforms like the establishment of Austria Wirtschaftsservice Gesellschaft (AWS), The Austrian Research Promotion Agency (FFG) and the reform of the Austrian Science Fund (FWF) resulted in increased efficiency in the innovation system. The time has now arisen to assess or evaluate the outcome of the 2010 strategy plan published by the Council who considered it necessary to improve the quality and efficiency of the national innovation system and thus increasing the return on government investment. Let me shortly summarize the strategic policy recommendation laid down in the Strategy:

- Promote quality on a broad level and excellence at the top
- Strengthen networking and co-operation between science and industry
- Improve the efficiency and effectiveness of the promotion system

Additionally, globalisation and technological advances are continuously changing the business environment and the demand for skills of the workforce in virtually all countries. As knowledge and creativity increasingly become the basis of competition, high skill levels are vitally important to economic performance and thus living standards. Without a world class education and training system any country will have great difficulty succeeding economically in the face of this intensified competition.

We do know that a well educated and highly skilled population will contribute to a competitive, innovation driven, knowledge based, participative and inclusive economy. It is thus essential to exhaust the skills of the resident population. It is also important to increase participation in the workforce and continue to attract highly skilled migrants.

At the first Evaluation Day in 2006 we posed the question "Excellence: to pick or to foster?", we looked at evaluation in the context of projects and peer review. Now we have the evaluation assessment of programmes, programme portfolios or institutions in mind when we wonder: "Evaluated – and then?". To ask the question "Evaluation – and then?" is another step in integrating evaluation and its results in the policy cycle. In Austria we already have established a respectable culture of evaluation. In between 2003 and 2007 more than 60 evaluations were conducted. They were collected in the book titled "Evaluation of Austrian Research and Technology Policies", and published by the Council and the Plattform FTEval. All these reports can also be downloaded from the website of the platform. This all is just the peak of an extensive and intensive culture of evaluation. It helped to increase transparency, both in methodology as well as in results. You can say that it is almost stylish to order an evaluation. Politicians always ask, whether it has been evaluated. What an evaluation is or should be used for, however, is still under discussion. In my opinion it should be for transparency, comparability, legitimacy and above all it should be used as a learning instrument.

Still one wonders whether and how we could have done better. As a matter of fact there is always room for improvements and thus the existing evaluation systems should also be further developed. More knowledge improves your performance and should lead to better results and better outputs. To improve the situation we need to have the following questions in mind: Did we educate people properly to get the best results and to make the best use of results? Are we asking the right questions? Are we approaching a problem correctly? How can we assess the value of what we are doing? What is requested to be better next time?

The final aim of funding research is to create jobs and solve problems in the private as well as the social live of people. In Austria it is the same as in any other European country. At the end of day jobs should be created both in the areas of science and research, in universities and the industrial sector, as well as in the production and the service of the newly developed products. The successful conversion of results from research and development into applications is important. Going through the process of research, what needs to be done to promote that? Adequate funding of science and research is essential to educate people to be qualified to carry out research and evaluation. Therefore the government needs to guide the process of knowledge generation through increased research funding in higher education and industry.

When you assess the results of funding, you have to pose some very important questions. Has the money been spent well? Could we have improved the output? What are "products" of research? Does it have the "side effect" of producing jobs and improving industry? Did or could we change to the right structures? Were our measures leading to the right structures?

In this conference today, we hopefully get answers to some of these questions. In order to address these issues we have to consider the impact of the level of education, the level of sustainable development and at the level of enterprise.

In the end, I would like to stress that politics decides what happens. Are the recommendations implemented? To use evaluation is a political process. They are an important foundation of the actions that an administration takes and consequently the basis of political decisions. In the light of the political implication of evaluations the conference today is very important. We want to take a critical look at Austria's situation; however we should not be negative but look positively into the future and try to improve the system. Evaluation should be used as a learning instrument. Although not legally binding it should be an integrated assessment in a programme cycle and should be designed from the beginning.

In a recent publication, Widmer and his colleagues say that Austria lacks a system of quality control and that there is no common approach to evaluation with respect to initiating and pursuing evaluations. So we have to get the system right. Already the knowledge that something will be or even could be assessed has at least an indirect effect on the subject of evaluation and the persons involved. “Use of evaluation results and processes does not come automatically. It must be planned and cultivated throughout the evaluation process.” (Horton et al. 2003)

In this conference we are going to hear some very brilliant presentations from Patrick Cunningham who is already for many years the scientific advisor to the Irish Government. Patrick Cunningham will talk about the Irish perspective on science policy and evaluation. Following is Martin Weber from the European Court of Auditors who tells us what concepts of intervention logic and feedback loops the Court recommends. This part is titled “Evaluations- and what could be done in the EU?” In the afternoon we are going to have Bernd Ebersberger from the Management Center in Innsbruck and he looks at meta-evaluations and how to make evaluations comparable.

Author:

Dervilla Donnelly

c/o Austrian Council - Rat für Forschung und Technologieentwicklung

Pestalozzigasse 4, 1010 Wien

donnelly@stp.dias.ie, donnellydk@eircom.net

Petra Wagner-Luptacik

Science Policy and Evaluation: the Irish perspective Interview with Prof. Patrick Cunningham

Where does Ireland stand with regard to the impact of its science policy?

Ireland has come through a dramatic period of economic expansion in the past 15 years. Whereas traditional manufacturing - half of which is food industry – is growing at a modest trajectory, modern manufacturing is now broad based in Ireland with dramatic increase in output. More recently, due to a deliberate policy shift, internationally traded services have also experienced dramatic growth.

Growth in value of the economy has followed that: In the 1980s, Ireland's economy partly lost ground due to mismanagement and excessive debts. Starting in the early 1990s there was rapid acceleration in economic growth, statistically ahead of the OECD average. In the ten years 1998-2008, the economy grew at 6.5% per year, twice the average of the EU 15.

Science policy was designed to drive and support that growth. From a low base a programme of new investment in science began in 2000. Between 2000 and 2008, the Government Science Budget increased in real terms from €400m to almost €1,000m. This was twice the rate of growth in the economy. Business investment increased in parallel.

What is the impact of the current global economic crisis on Ireland?

Ireland is particularly strongly hit by the global recession with a projected 8.3% decline in GDP for 2009. The problem is more severe than in other European countries. This very sudden reversal is mainly due to the construction boom in the past years (15% of GDP), which has come to a dramatic stop now. Moreover, in the boom years government expenses for public services expanded - a process which cannot easily be reversed.

What has been the role of policy in this transformation?

When we look back at the success of the preceding 15 years, Ireland was able to take advantage of factors such as globalization and cultural factors like the traditional links with the US. Policy has also played a part. Investment in education is a major element of policy. Over the last 30 years in the OECD as whole, the proportion of the age group entering the work force (25 to 34 yrs) with tertiary education has been increasing at 0.5 percentage points per year. In Ireland, by contrast, the increase has been 0.8 percentage points per year.

This up-skilling of our work force has been a very important element in driving the transformation of the industrial base, which would not have been possible without the steady supply of well trained people. The further transformation since 2000 includes a significant contribution to the generation of knowledge – with growing investment in R&D.

Recent statistics show that 1,025 Irish firms are doing significant R&D. Some 345 foreign-owned firms (over 60% American) are operating, increasingly with an R&D base, in Ireland. Of the total business investment, the biggest proportion is in fact conducted by foreign firms.

The strong growth in public investment in R&D began in 2000, following policy decisions made in the mid 1990s. Two National Development Plans were launched: the first running from 2000 to 2006, the second from 2007 to 2013. The first national development plan shows public investment growing strongly - from 0.33% GNP in 2000 to twice that level (0.58%) in 2007. That was the result of very consistent policy-making – the results promised have been delivered.

Where does investment leave Ireland in international rankings today? Total expenditure on R&D is 1.66% of GDP. Compared to the Lisbon target of 3%, Ireland has a more modest national target of 2.5%, as the Lisbon target would be unrealistic given the low starting position. Austria is already beyond that (2.6%). So, Ireland has still some way to go.

What are Ireland's science governance structures?

There were no traditional structures to manage R&D investment at national level with the exception of some sectors like agriculture or food research. Ireland's current science governance structures were put in place following a series of reports in the 1990s. Eight government departments are directly involved. In order to have a coherent science policy a Cabinet Committee chaired by the Taoiseach (Prime Minister), was established. It meets 3 times a year and takes significant decisions at a strategic level. One of the decisions last year was to compete to host the *Euroscience Open Form* ("European City of Science") in Dublin in 2012 – our bid was successful.

The Cabinet Committee is shadowed by an Inter-Departmental Committee of senior civil servants from the eight government departments, which in turn is served by the Office of Science, Technology and Innovation (STI) in the Department of Enterprise, Trade and Employment.

As Chief Scientific Adviser, I report to the Cabinet Committee**.** A separate Advisory Science Council represents major stakeholders such as industry etc. and prepares regular reports on different S&T topics.

Part of the new structure was the establishment of the Science Foundation Ireland, which provides competitive support for basic research.

What about science promotion and funding in Ireland?

Total Government allocation to research in 2008 was about 1 billion Euros. The majority of government funds flow through two government departments: Enterprise, Trade and Employment (49%) and Education and Science (39%), followed by Agriculture and Food (14%) and modest contributions by Health, Energy and Environment.

Is there best practice in advising government on strategic choices?

We try to learn from international best practice.

John H. Marburger III, former Director of The US Office of Science and Technology Policy, made the point that we formulate policy with a lot of implicit assumptions, but do not articulate them very well. A "Science of Science Policy" is needed for informed decisions. Major strategic policy questions revolve around the need for better benchmarks - and evaluations. How much should a nation spend on science? What kind of science? And, how much from private versus public sectors? These and related science policy questions or assumptions deserve closer scrutiny. A new "science of science policy" is emerging which should provide input for evidence-based guidance for policy decisions.

How much should a nation spend on R&D?

There are no absolute targets. Research promotion is an act of faith – faith in the long-term profits to the economy. The EU target of 3% GDP has been set with the USA in mind, but South Korea now targets 5 %.

On average, expenditures in R&D can take 10 years to return measurable benefits to the economy. With basic research, it may take up to 30 years before an application is found and turned into business. Applied research however often turns out marketable products within three or five years.

How about measuring the benefits of public investment in research and innovation?

The wealth of nations is captured in various reports. The INHOLLAND report with its 38 indicators is interesting in this respect. It emphasises intellectual capital (people, knowledge, institutions). This is where the investments should go in the future.

Ireland ranks 8th in the INHOLLAND Report and 11th in the Lisbon review 2008. The summary Innovation Index of the European Innovation Scoreboard 2008 (with 29 indicators) shows that Austria and Ireland are both among the "followers group", ranked 7th and 8th, respectively. The leaders were Switzerland, Sweden, Finland, Germany, Denmark, and United Kingdom. These metrics now show Ireland close to EU-15 average.

What are the challenges for Irish science policy-making in the future?

Monitoring the delivery of policy against correct indicators is a crucial issue. Moreover, we need to establish more effective linkages to the dual business sectors. Ireland will need to pay close attention to the needs of foreign-owned firms and make special efforts to lift indigenous firms. And last but not least, keeping on target to match the leading countries, which is realistically achievable within 10 years.

Author:

Petra Wagner-Luptacik
AIT - Austrian Institute of Technology GmbH
Techgate Vienna, Donau-City-Str. 1, A-1220 Wien
petra.wagner@ait.ac.at

Martin Weber

Embedding evaluations in the EU context - an auditor's perspective

Introduction[1]

Many different types of appraisal or assessment activities are referred to as evaluations. Formal evaluation is carried out routinely in many public administrations across the world, including the European Commission. Although there is no single agreed definition of evaluation, there is a general consensus that evaluation is a process that makes evidence-based judgements about the actual or likely effects of public interventions such as expenditure programmes, financial instruments and regulatory measures. These judgements are formed on the basis of a disciplined inquiry involving procedures which draw heavily on techniques from the fields of social science and public policy for collecting and analysing quantitative and qualitative information.

This paper

- explains the use made at the European Commission of prospective and retrospective evaluation;
- provides some information on the diversity of evaluation arrangements in the different EU policy areas;
- suggests three measures to make evaluations more relevant and robust for policy formulation and decision-making.

Prospective and retrospective evaluation at the European Commission

The Commission's evaluation policy requires that the Directorate-Generals evaluate their activities on a regular basis. Evaluation is seen as an integral part of the Commission's Strategic Planning and Programming (SPP) policy and its Activity Based Management (ABM) initiative, and is acknowledged in the Financial Regulation as part of the Commission's internal control system for ensuring the sound financial management of EU funds. In addition, prospective

---

[1] The content of, and any views expressed in, this article are those of the author and are not those of the European Court of Auditors (ECA).

evaluation, particularly in the form of impact assessment, has been allocated a key role in the Commission's efforts towards "Better Regulation" and the reduction of administrative burden.

<u>Box 1:</u> *Evaluation comprises retrospective and prospective assessments*

The European Commission defines evaluation as a "*judgement of interventions according to their results, impacts and needs they aim to satisfy*"[2]. In addition to retrospective evaluation (such as ex-post or interim evaluations), the Commission carries out two types of prospective analysis: (regulatory) impact assessments are defined as *"... the process of systematic analysis of the likely impacts of intervention by public authorities*"[3]. They are required for all items on the Commission's Legislative and Work Programme (CLWP) with an economic, social or environmental impact, which in general includes all regulatory proposals, "White Papers", expenditure programmes and negotiating guidelines for international agreements. Ex-ante evaluations are required by the Financial Regulation for all proposals that are likely to have a significant budgetary impact. There is considerable overlap between them and in practice, where both are required, impact assessments and ex-ante evaluation tend to be combined in a single exercise.

Diversity of EU policy areas and evaluation arrangements

The European Union Treaties form the legal basis for all measures taken by the European Union and organise the way it is run. In accordance with the Treaties, the European Parliament and the Council can adopt legislation, generally based on a proposal by the Commission. The powers and responsibilities centralised at EU level may however differ from one policy domain to another.

The role of the European Commission and the management method applicable and the type of intervention used in a particular policy field for achieving the EU's policy objectives largely determine the monitoring and evaluation arrangements in place. In general, this is specified in the legal acts underlying each policy and programme.

---

2 European Commission, DG Budget, Communication to the Commission from Ms. Grybauskaitė in agreement with the President, "Responding to strategic needs: reinforcing the use of evaluation ", SEC(2007) 213, 21 February 2007

3 European Commission, "Communication from the Commission on Impact assessment", COM(2002)276 final, 2.6.2002

Partly in response to this diversity and in order to encourage more widespread use of evaluation, the Commission's framework for retrospective evaluation activities was reviewed in 2000 and 2002. These communications were updated in 2007 taking into account the consequences of the Commission's administrative reform and the new Financial Regulation[4]. Similarly, the Secretariat-General introduced a number of reforms in 2002 and 2006 to streamline the Commission's system for (regulatory) impact assessments[5].

Decentralised evaluation system

The Commission is the sponsor of most evaluations of EU activity, with the exceptions of the ex-ante (and interim) evaluations in the area of Structural Funds (which are the responsibility of the recipient Member States). Within the Commission, evaluation is decentralised to the Directorates-General, with the central services (in particular DG Budget and the Secretariat-General) providing support and coordination. As a result, the evaluation systems can vary significantly according to the policy area and between Directorates-General.

***Box 2*: Sector-specific arrangements for retrospective evaluation**

In some areas the Commission carries out evaluations of activities for whose management it is directly responsible (such as the EU RTD framework programmes). In most policy areas, however, the Commission evaluates the effects of activities that are implemented by others (for example, national administrations, international organisations, joint undertakings, non-governmental organisations or private-sector entities), but for which - according to the Treaty and the Financial Regulation - the Commission has an overall responsibility.

The "evaluation paradox" - and some reflections on what to do about it

Within the European Union, but also in Member States and elsewhere, one can observe a large discrepancy between the high expectations placed in evaluation, as expressed in many policy documents and academic papers, and its implementation in practice. While the quality of

---

4 European Commission, DG Budget, Communication to the Commission from Mrs Schreyer, "Focus on Results: Strengthening Evaluation of Commission Activities", July 2000, SEC(2000) 1051; European Commission, DG Budget, Communication for the Commission from the President and Mrs Schreyer, "Evaluation Standards and Good Practice", C(2002) 5267 23.12.2002; European Commission, DG Budget, Communication to the Commission from Ms. Grybauskaitė in agreement with the President, "Responding to strategic needs: reinforcing the use of evaluation ", SEC(2007) 213, 21 February 2007

5 European Commission, "Guidelines on Impact assessment", SEC(2005)791, 15.5.2005, as amended on 15.3.2006

individual evaluations is very heterogeneous, one can also observe that very good pieces of analysis sometimes play only a limited role in decision-making.

The question is how to make evaluation procedures and results more relevant and robust, and thereby provide incentives to make more and better use of evaluations as a potentially powerful tool for an evidence-based policy. It could be argued that solving this "evaluation paradox" requires the following:

- first, legislators need to be more involved in setting the evaluation agenda;
- second, evaluative activities must come full circle, i.e. prospective assessments must be designed so that they can provide a basis for subsequent retrospective evaluations;
- third, the evaluative work carried out within the Commission must be subjected to external scrutiny to ensure its adequacy and reliability.

An enhanced role for legislators in setting the evaluation agenda

Within the European Union, as in all democratic systems elsewhere, the European Parliament and the Council exercise their function as legislators primarily through the promulgation of regulatory measures. A system such as evaluation which poses constraints on this process, whether by requiring a more thorough assessment of risks, a more careful quantification of costs and benefits, an exploration of potential consequences and an assessment of intended (or unintended) effects may therefore seem counter-intuitive.

Evaluation will not change the basic dynamics of political decision-making, which includes negotiation and bargaining. Moreover, legal acts often involve hard-won compromises that none of the Institutions subsequently want to unpick. The static allocation of funding to programmes in the European budget within the financial perspective's ceilings only provides an opportunity to introduce systemic changes (involving a shift of budgetary resources) when legislation comes up for renewal every seven years.

Only if legislators and regulators believe that evaluations will help them to produce better designed, more successful and more acceptable measures, will the prospects for the system be positive. Legislators therefore must be given a greater stake in setting the evaluation agenda, both for prospective and retrospective evaluations. In the future, setting up formal interinstitutional groups (composed of representatives of the Commission, the Council and the Parliament) could provide a forum to jointly define the design of evaluations. Basically, it would ensure that legislators have a possibility to indicate the kind of questions that in their view evaluations should answer and the type of information that they would consider adequate. Opening up the system would allow the legislators' expectations to be reflected in evaluation planning from the very beginning, rather than leaving it to the Commission alone.

This is of course a delicate (but not impossible) proposal to put into practice in the EU context, because it might be seen as an infringement of the Commission's exclusive right to make legislative proposals (as enshrined in the Treaty) and the Commission's final responsibility for implementing the budget in accordance with the EU Financial Regulation.

Linking prospective and retrospective evaluation

If the purpose of evaluation is to increase the space for evidence-based policy formulation, it should provide comprehensive and consistent evaluative information throughout the whole policy cycle (policy design, legislation, and implementation). Prospective evaluations should therefore be seen as a framework against which legislation and regulation can be evaluated and reviewed and useful lessons learnt for the future.

Retrospective evaluations for spending programmes are required by the Financial Regulation, but also in each of the legal acts underlying EU programmes or activities. With regard to prospective evaluation, the Commission has imposed on itself a requirement for a comprehensive system of impact assessments. In the current situation, however, the two systems merely coexist and possible synergies are not sufficiently exploited.

As a way ahead, the Commission should make it a formal requirement for all retrospective evaluations of policies and programmes to look back to impact assessments to see whether and to what extent the purposes initially intended have been achieved.

Using the estimations provided in prospective evaluations as benchmarks of intentions it is possible to attempt a more informed review of whether policy aims have been achieved, of whether costs and benefits have been broadly in line with expectations (or significantly different from them) and whether unanticipated effects or developments have intervened to change the outcomes from those anticipated or desired.

In addition, such an approach would bring about a further alignment and standardisation in the methodologies used within the Commission for prospective and retrospective evaluations. As a side effect, it would also start a reflection process on the suitability of the Commission's current organisation and procedures in the area of evaluation.

External quality assurance of evaluations - a role for public audit bodies?

Over the last years, the Commission has made uneven progress in institutionalising its internal quality review of evaluative activities.

For prospective evaluations, an "Impact Assessment Board" under the authority of the Commission's President was set up as recently as 2006 to scrutinise the quality of all impact assessments prepared by the operational Directorates-General and services (but without the power to sanction poor quality assessments or to require impact assessments to be carried out on specific issues).

For retrospective evaluations, there is not even such a centralised quality review procedure within the Commission. Here DG Budget merely has an advisory role, in line with the function of the Secretariat-General for impact assessments. This situation is in stark contrast with what one would find in the system for evaluation and regulatory impact analysis in the United States of America, where the Office of Management and Budget (OMB) has a much stronger role in ensuring quality assurance.

A robust quality assurance system would also benefit from having an external component, in addition to internal arrangements. Such a system for auditing whether or not the Commission is complying with its mandatory (or self-imposed) evaluation requirements could focus on the quality of individual assessments, or on the system for producing them, or on both.

Evaluations or evaluation-related activities are often used as valuable sources of information by the European Court of Auditors in its work, and this is general practice in most public audit bodies.

In addition, through audits carried out on an "ad hoc" basis on its own initiative, the Court can decide to assess a specific evaluation (in particular where its findings are highly relevant to the Parliament and Council) or an evaluation system in a specific field[6]. So far, however, this is not done in a comprehensive or consistent manner across all policy areas. In particular, it is not done for regulatory measures without direct budgetary incidence.

***Box 3:* Recent example of an audit of an evaluation system**

A summary of the European Court of Auditors' approach to auditing the Commission's evaluation system for the EU RTD framework programmes has been presented in a recent issue of the Plattform Newsletter[7]. The full audit report submitted to the European Parliament and the Council has been published in the Official Journal as Special Report N°9/2007[8].

---

6 European Court of Auditors, "Performance audit manual", December 2006 (see http://www.eca.europa.eu)

7 Plattform Newsletter N°31: Martin Weber & Gareth Roberts, "Evaluating the EU Research and Technological Development (RTD) framework programmes - could the Commission's approach be improved?", 2008

8 European Court of Auditors, Special Report N° 9/2007 "Evaluating the EU Research and Technological Development (RTD) framework programmes - could the Commission's approach be

As a result, in the current evaluation system of the European Union, an independent and external function which scrutinises the operation of the Commission's different evaluation systems is lacking. Such a “compliance” body should verify on a regular basis whether evaluations have been carried out where and when they ought to have been, whether they have been conducted to a suitable standard of quality and whether their conclusions have been properly fed into the Commission's regulatory or policy-making processes and its legislative proposals.

One way to achieve this would be to entrust the external auditor with this function, as is the case in the USA, where the Governmental Accountability Office (GAO) plays an active role in particular in the area of prospective evaluations. In contrast with the USA, up to now the European Court of Auditors has been given no explicit role by the legislators in assessing the quality of evaluations.

But why should this task be entrusted to the public audit body? First of all, the Court of Auditors is the only EU Institution which - according to the Treaty - is independent of the Commission, the European Parliament and the Council at the same time. It also reports to both legislators. Moreover, the Court has the necessary resources and expertise in the field. Finally, entrusting this role to the Court would not necessitate the setting up of a new body.

What would be needed to change the current situation? In practice, the European Parliament and the Council would have to include in all legal acts regarding spending programmes or other forms of public intervention a provision whereby the responsibility to carry out a performance and compliance review of the respective evaluation system is to be given to the European Court of Auditors.

Moving the Court towards a more expansive view of its mandate would not only be in keeping with modern governance in the public sector, it would also lend force to the view that efficient, effective and proportionate legislation is a key criterion for achieving "value for money" in public administration and increasing the competitiveness of the European Union and its Member States.

---

improved?”, together with the Commission’s replies, 20.12.2007 (Official Journal 2008/C 26/01), 30.01.2008

CONCLUSION

This paper indicates some thoughts on how evaluations in the EU context could be made more effective. In particular, it is argued that three requirements must be met to achieve this:

- legislators should be involved in the evaluation design through formal interinstitutional groups (composed of representatives of the Commission, the European Parliament and the Council);
- the Commission should make it a formal requirement for all retrospective evaluations to evaluate against the impact assessment;
- the European Court of Auditors should be consistently asked by the European Parliament and the Council to scrutinise the performance and compliance of the Commission's various evaluation systems.

Author:

Martin Weber
Head of Cabinet
European Court of Auditors
12, rue Alcide de Gasperi, L-1615 Luxembourg
martin.weber@eca.europa.eu

Julia Schmidmayer

Evaluated. So What!? Summery of the second "Evaluation Day"

Austria's history of evaluation

The establishment of the Platform Research & Technology Policy Evaluation (FTEval) in 1996 was a major impulse for developing a decent evaluation culture in Austria. FTEval has been highly successful in strengthening the role of evaluation in the public policy discourse within Austria. The handbook 'Evaluation of Austrian Research and Technology Policies' edited by the FTEval (published in July 2007) documents nearly 60 evaluations carried out from 2003 to 2007. This fairly large number of evaluations shows the remarkable position of Austria with regards to evaluation. Besides the efforts of the Platform FTEval the Austrian Council for Research and Technology Development attaches also great importance in advancing and establishing an Austrian culture of evaluation. The Council aims to improve the use of evaluation as a tool for transparent assessments of technology policy interventions in order to assist optimal strategic planning of RTD policy.

Why do we need an 'evaluation day'?

What main conclusions can be drawn after having carried out that many of evaluation studies? Did these evaluations actually change the way technology policy is organised in Austria? Did we learn from possible mistakes detected by evaluation studies and did we implement the recommendations given by evaluation studies? Questions like these motivated the Platform FTEval and the Council to organise a specific workshop (called 'evaluation day') in order to shed light on the implementation of evaluation results. It was held on March 23rd, 2009 and aimed at raising the significance of evaluations as learning and governance instrument for RTD interventions as Michaela Topolnik (Austrian Council) emphasised in her introduction as moderator and initiator of the day. As she pointed out the Evaluation day 2009 addressed the following topics:

- How can we use evaluations in organisations and in the governance system to initiate learning processes?
- What are the requirements to make the results of evaluations comparable and utilisable?

- Which implications can be derived for the implementation of recommendations?
- Which factors should be considered when integrating evaluation results in the design of RTD programmes and governance processes?

Rupert Pichler (Federal Ministry of Transport, Innovation and Technology) opened the event and asked for the function of evaluation. He stressed the demands of the evaluation community for the implementation of recommendations even if they are inconvenient. Evaluators should communicate their results in a way appropriate to the specific information needs of policy makers in order to have them taken seriously. In reference to that Rupert pointed out the importance of the evaluation of government funding in RTD from a systems perspective in Austria. First results of this so called 'systems evaluation' are to be expected in May 2009[9].

After that *Devilla Donnelly*, who is a member of the Austrian Council, stressed the necessity to improve the policy and efficiency of the Austrian innovation system. Specifically, she emphasized the role of evaluations for increasing the rate of return of public policy interventions. One tool to generate and maintain RTD policy is continuous evaluation which offers orientation, contributes external knowledge and opinions, and provides documentation of how policies are implemented.

Raising the question 'Evaluated – and then?' goes one step further to demand for the integration of evaluation and its results into policy cycles. Looking back at a bulk of 60 evaluations, Devilla speculated that today ordering an evaluation is trendy in Austria but evaluation should be no process for its own sake. In the end, it is the policy maker who decides what is to be done with the results and recommendations of an evaluation. Devilla Donnelly asked for a critical view on the Austrian situation during the 'evaluation day'. She raised in her introductory talk a lot of very ambitious questions and raised the bar high for the speakers of the workshop to answer these questions.

Metrics to counterfeit policy options

After the welcome address of Rupert Pichler and the introductory talk of Devilla Donnelly the first presentation was held by *Patrick Cunningham*, Chief Scientific Adviser to the Irish Government. He gave an insight into the Irish science policy and evaluation. His starting point was a discussion about the definition of the term 'evaluation'. In English-speaking countries evaluation is usually understood in a quite narrow sense, namely as mere project assessments. With the quest for improving the Irish science policy an urgent need for benchmarks and comprehensive international comparisons has been arising which go beyond the narrow

[9] For more information please refer to http://www.bmvit.gv.at/en/innovation/policy/evaluation.html

approach of evaluation. Patrick then discussed recent results of benchmarking endeavours of the Irish science and technology landscape with the Austrian one.

In addition to these comparisons between countries, he looked for comprehensive metrics in order to counterfeit policy options. Being spoilt for choice with a variety of indicator sets he concentrated his endeavours on the EU Innovation Scoreboard. Even though it has certain deficiencies it is probably the best general metric to be found for policy evaluation in Europe. In total, a ranking based on 29 different indicators is made for all EU member states. Ireland and Austria are comfortably placed as leaders in the so called group of innovation 'followers'. By adhering to these indicators, progress of Irish policy can be monitored and relevant indicators found. If Ireland keeps on track, Patrick predicts that the country's efforts matching the leading countries can be achieved in approximately ten years.

In Europe evaluation is fashionable …

Subsequently, *Martin Weber*, who works for the Head of Cabinet at the European Court of Auditors, held the second presentation. He gave an overview on the evaluation system of the European Commission. It is of prior interest how knowledge gained through evaluation of European Framework Programmes (FP) can be utilised for policy design of future generations of these programmes. Today, evaluation is a surprisingly fashionable topic in the European context. With regard to a conference held in Prague in February this year 300 people participated in the presentation and discussion of results of the ex-post evaluation of FP6.

The European Commission drew up an elaborated evaluation system paired with methodological approaches which was gradually implemented during the last three FPs on the European level. In 2007 the Court of Auditors undertook an audit assessing the Commission's evaluation system of the FP. Among other things the Court recommended that one of the main deficiencies of FP is the lack of a comprehensive logic for intervention. Such a framework is fundamental when assessing the outcomes of various policy measures. The existence of a rationale when designing a policy should lead to targeted and better structured programmes and would improve the effectiveness of evaluations and programme controlling.

As the Commission's budgeting process regulates that every policy or programme has to be evaluated, two instruments were introduced – impact assessment and evaluation of FP. The definition of both is often mixed up. Evaluation is used in the context of reports and analysis carried out retrospectively during a programme cycle. Impact assessment is conducted ex-ante before a FP is carried.

Regarding the Commission's system of evaluation a comment was raised by a person in the audience who noted that the draft of the FP is put together without knowing the results of the ongoing one. As a consequence the design for the next FP is based on results of the second last

FP which might be already outdated. Nevertheless, the situation has improved since the implementation of FP7 as duration of the programme cycle was extended to seven years (from formerly four years). Thus, more information about the implementation and management of the last FP is available.

Evaluation on a conceptual level

After the lunch break *Bernd Ebersberger*, working at the Management Center Innsbruck (MCI), talked about evaluation on a systemic level. He presented a concept which tries to improve policy understanding by means of secondary analyses of policy evaluation. There is a demand for meta-analysis as politics is becoming more complex due to influencing variables, a growing number of stakeholders etc. Policy interventions are interacting or even complementary, as not one single measure can remedy failures. Impact assessment of bundles of policy interventions is needed on a systems level to identify weaknesses and an overlapping of policy measures. Comparability of evaluations should be guaranteed in order to better understand individual interventions, to locate them in the policy mix and to better describe policy strategies. Thus, more measures have to be evaluated and a bigger number of stakeholders integrated.

By combining a large number of evaluations, effects of certain policy interventions can be assessed in order to shed light on their interplay at the level of innovation systems. A meta-analysis would answer questions in regards to networks of stakeholders, institutions, framework conditions and other policies that together determine the creation and exploitation of knowledge and innovation. Bernd Ebersberger and his colleagues drew up a systematic conceptual framework for utilising and analysing existing evaluation data. It consists of three modules:

a) meta evaluation (data collection, evaluation of existing evaluations),

b) meta-analysis (comparison and in-depth understanding of individual types of measures), and

c) evaluation synthesis (interpretation of specific evaluations of individual interventions against the backdrop of a new interpretation of existing measures).

The discussion of secondary analysis of evaluation is characterised by an academic discourse as e.g. INNO Appraisal tries to pick up one or another idea of the presented concept. INNO Appraisal is a large scale study financed by the European Commission. Results will be obtainable by the end of 2009[10].

---

[10] http://www.proinno-europe.eu/index.cfm?fuseaction=page.display&topicID=53&parentID=53

Ready to rumble in the policy arena!

Klaus Schuch, Centre of Social Innovation, opened the panel discussion which addressed the interaction of evaluators and clients of evaluations.

Klaus Zinöcker broke the first ground with a very entertaining impulse talk about the Austrian evaluation platform FTEval and its achievements. The Platform FTEval was set up to contribute to the field of policy development and especially improve innovation policies. One way to shape innovation policy is the implementation of recommendations of evaluations. But evaluation should not serve as an end in itself; but as means to an end. There are a lot of guiding principles that evaluators have to bear in mind: Evaluations are reputable endeavours. Luke Georghiou referred to evaluation as a social process[11]. Stefan Kuhlmann remarked an increasing interest in learning-oriented evaluation approaches and aligns evaluation as mediation instrument in the policy arena[12].

Three different types of actors encounter in this policy arena: the evaluator, the client of the evaluation (in Austria it is typical one or more of three ministries) and the programme agency which is responsible for the management of R&D programme to be evaluated. In theory an ideal interaction of actors is shown in the policy cycle in which several levels of evaluation leads to better policies. But in practice every actor has a hidden agenda to follow, e.g. the evaluator wants to put forward a certain set of methodologies, the client likes to have the evaluated R&D programme proven to be successful etc. The question ‘Why do we actually need evaluations?’ is often posed by stakeholders and this attitude leads to a bad handling. As the evaluator has done the job by completing the evaluation the client is responsible for the communication of evaluation results and implementation of recommendations. Feedback is often only accepted from the headship in the ministry but not from the evaluator. Klaus Zinöcker made some very rough suggestions of how to bring evaluators and clients together:

- Evaluators have to take great pains to bring forth useful recommendations.
- Clients have to take evaluations and their results more seriously.
- Steering and controlling function of evaluation ought to be improved.

Dorothea Sturn, University of Vienna, commented on the impulse talk by appreciating the entertaining delivery but countering that the situation is slightly different than it was illustrated.

---

[11] Georghiou, L. (2003): Evaluation of Behavioural Additionality. Concept Paper. Presented at the Meeting of the TIP Working Group of the OECD, June 2003

[12] Kuhlmann, S.: Evaluation as a source of ‚strategic intelligence’, in: Shapira, P., Kuhlmann, S. (2003): Learning from Science and Technology Policy Evaluation – Experiences from the United States and Europe

In the end, it is the client's decision whether recommendations of evaluators are accepted and implemented. A vast number of evaluations were carried out during the last 20 years. Austrian RTD policy is characterised by innumerable governance structures with even a much bigger number of RTD funding programmes which ought to be reduced. However, the current wave of evaluations in Austria did not result in this much-wanted reduction of different RTD programmes. Furthermore, the optimisation of monitoring systems and the collection of data would be a desperate topic which is demanded for ages.

Subsequent to Dorothea Sturn's comments *Hans-Peter Lorenzen*, board member of the German Association of Evaluation (DeGEval), rose to speak and reminded that policy makers have to take recommendations in hand. According to the standards of the Platform FTEval, evaluations provide knowledge which should lead to concrete action. For those involved, evaluation results should be made 'easily digestible' to make best use of them. In addition, it is important to create supporting structures and conditions to facilitate the utilisation of evaluation results. It is often not easy to implement results as there is a gap between the 'request' and the 'ability' of implementation. The underlying problem is the comprehensibility of recommendations which needs to be elaborated and is often disregarded. Furthermore, not everything requested is evaluable due to methodological problems. Thus, it is time to have realistic expectations.

In the second discussion round *Brigitte Tiefenthaler* from Technopolis Austria raised the question of which ingredients a good evaluation shall be composed. In her opinion only a good evaluation provides a basis for decision making. Naturally, it is only one element but a crucial one. However, evaluation needs the commitment of the clients to take the results seriously and to utilise them. She raised the question why something has to be evaluated if results and recommendations are not utilised.

The timing of evaluation was another important topic to be discussed. Hans-Peter Lorenzen stressed that one has to be aware of the aims of an evaluation beforehand. The more complex the evaluation is, the more difficult it is to implement recommendations. Thus, the Platform FTEval suggests in its standards that a realistic planning process is required to ensure the availability of evaluation results before a decision is taken. Evaluation has to be taken into account throughout the whole policy cycle and integrated into it at best. The instrument of evaluation ought to be an integral part of planning and developing RTD policy. Brigitte Tiefenthaler added that courage for change is also needed. Policy makers have to take their time to utilise evaluation results.

Dorothea Sturn summarised that there are two possibilities after an evaluation has been carried out: an evaluation without implementation or an actual policy change as recommendations are considered. She raised the question if change does always need an evaluation. Naturally, difficulties can always occur when trying to implement the recommendations of an evaluation.

Sometimes implementation is not possible due to general conditions, it was tried but failed or it was not requested from the beginning. *Mariana Karepova*, working for the Austrian Research Promotion Agency (FFG), added some reasons why the implementation of an evaluation can fail from an agency's point of view. In total, the FFG implemented nearly 90% of all recommendations made in evaluations:

- Recommendations are not specific enough (not good enough prepared to the information needs of those affected);
- Implementation of recommendations is beyond the power of the agency;
- Financial and personnel resources are not given;
- Agency does not share the opinion of the evaluators

Wolfgang Polt from Joanneum Research sounded the bell for the last discussion round for that day. He argued that there was a 'cultural' revolution ten years ago and now there is a demand for a second one. A new approach towards evaluation is needed. At the moment an evaluator takes a role as a trainer advising the client. In his opinion the distribution of roles should change towards the evaluator being more of a juror. Otherwise the evaluator runs into danger to make too cautious recommendations by courtesy which are not far reaching and effective enough. In order to ensure that 'undesirable' evaluation results are put forth, he demands to 'sacrifice' the disclosure of evaluation reports. Furthermore, the role of financier and commissioner of an evaluation should be separated.

What next?

Klaus Zinöcker made a brief wrap-up of the workshop. A lot of questions were posed but only some of them could be answered at the 'evaluation day'. Evaluation means first learning then steering. A good evaluation contributes to a good decision but also 'fig leafs' have their right to exist. We should face up with 'evaluability' to think about what is evaluable. As evaluation results support decision-making processes, evaluators should communicate them in an appropriate way to policy makers. But at the end of the day, policy makers have to take in hand evaluation results and decide on follow-up activities. The credibility and effectiveness of evaluation would be enhanced by disseminating them widely and ensuring a suitable level of publicity.

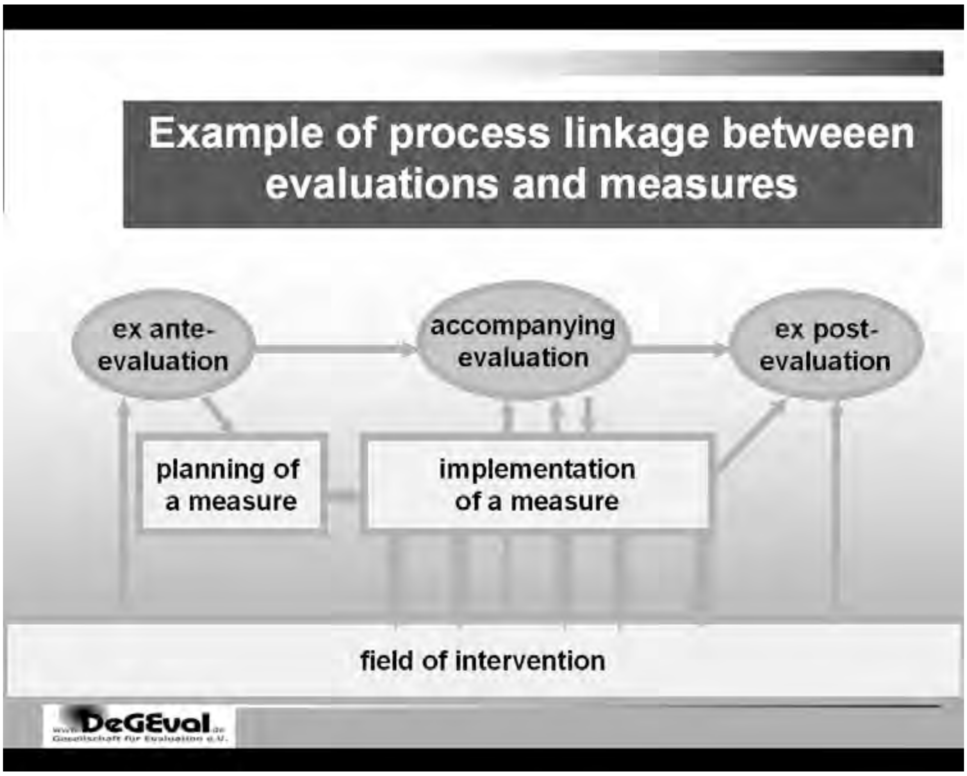
Author:

Julia Schmidmayer
Joanneum Research ForschungsgesmbH
Sensengasse 1, A-1090 Wien
julia.schmidmayer@joanneum.at

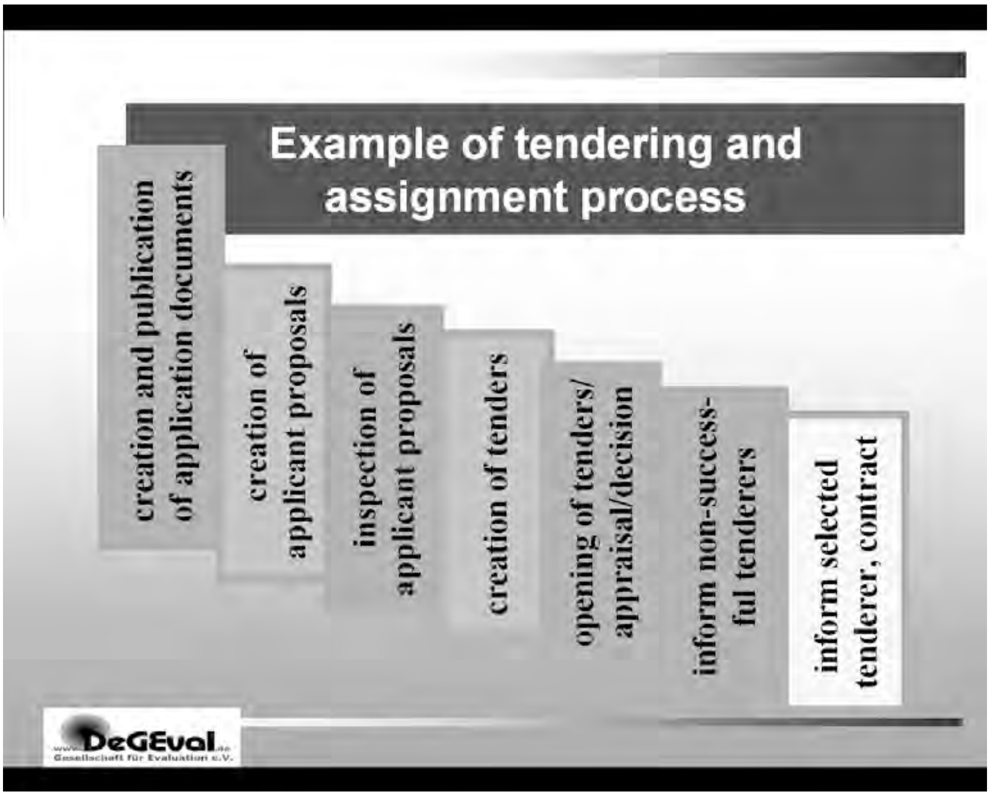
Hans-Peter Lorenzen

Recommendations for Clients of Evaluation

Commissioning agencies and clients of evaluation need recommendations that concentrate on their role in the evaluation process and enable them to discuss relevant questions with evaluators while planning or accompanying an evaluation.



For this purpose, the DeGEval – Evaluation Society prepared the booklets "Empfehlungen für Auftraggebende von Evaluationen – Eine Einstiegsbroschüre für den Bereich der Öffentlichen Verwaltung" and its English translation "Recommendations for Clients of Evaluations – An Introductory Brochure for the Field of Public Administration" in 2007. Both booklets can be ordered by e-mail at a price of 5,00 Euro plus postage (info@degeval.de). The target group of these recommendations are organisations and institutions of public administration, i.e. people who are responsible for, control, develop or implement measures within their organisation, but who have not yet had much opportunity to get to know the instrument of evaluation and to gather hands-on experience. The aim of this brochure is to show clients when an evaluation makes sense, who should participate, how an assignment is prepared, how an external evaluation should be accompanied and how the evaluation results can be disseminated. It contains examples for all of these aspects.



Most important, though often neglected, is the ex ante evaluation whose task is to ensure that the three elements of a measure (goals, instruments and evaluability) are compatible.

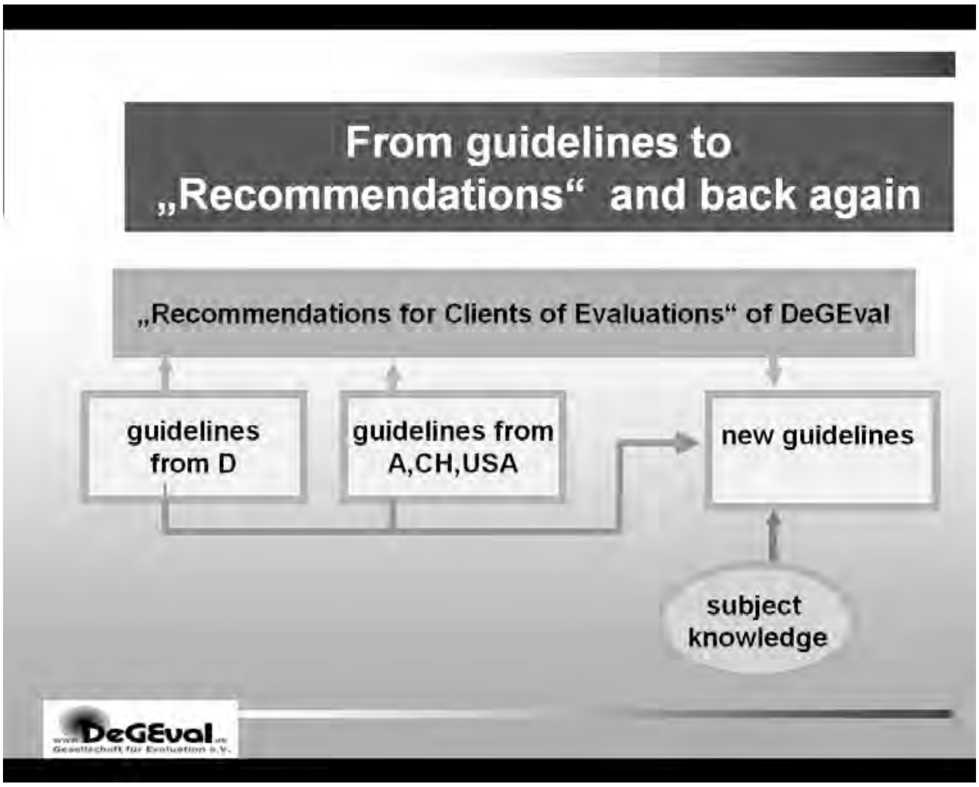
Asking who should participate is a key question that needs to be clarified before starting an evaluation. Schemes depicting the relationship network defined by the measure may be helpful. Within an institution, differentiations should be made between participating work units and levels of management.

External evaluations should be competitively assigned. The tendering and assignment process consists of consecutive phases that do not overlap. The tendering and assignment of an external evaluation by public clients or beneficiaries of public funds are subject to national or even European procurement law if the scope of the assignment exceeds a certain threshold value.

The work schedule during the implementation of an evaluation comprises research on the evaluand and the development of data collection tools, data collection, data input and data analysis. It is advisable to concurrently start developing a dissemination concept as early as possible. Other important elements are the preparation of intermediary and final reports and presentations targeting the different participants in the relationship network as well as the interested public.

The "Recommendations for Clients of Evaluations" are based on a number of guidelines for evaluations in various subject fields to the extent these guidelines could be generalised and were compatible with the DeGEval-Standards. To improve performance, they can be used as checklist in combination with guidelines available for the respective field.

If there is no suitable guideline for evaluations in a specific area, the Recommendations may be helpful in developing a new one by combining guidelines for other subject areas with specific subject knowledge and the Recommendations.



The DeGEval – Working Group "Industry" has proved in 2008 that this process works. The Working Group compiled a "Leitfaden – Evaluation in der Wirtschaft" (guideline for evaluations in industry) involving further scientific and economic experts. It has a lot in common with guidelines in the field of public administration, but includes some additional aspects. Evaluations in industry evaluate processes and structures. They help to optimise them by collective learning and to improve the economic output. Moreover, they take into account social dimensions as well as sustainability. Evaluations in industry tend to be more costly than other methods used to monitor processes and structures. They are useful for solving highly complex problems or studying the reasons for certain trends in the development of indicators. This brochure can be obtained from Dr. Christiane Kerlen, VDI/VDE Innovation+Technik GmbH, Berlin (kerlen@vdivde-it.de).

Author:

Hans-Peter Lorenzen
DeGEval - Gesellschaft für Evaluation e.V.
c/o ZQ Uni Mainz
Forum universitatis 4, D-55099 Mainz
lorenzen@degeval.de

PLATTFORM fteval
Forschungs- und Technologieevaluierung

Michael Dinges

INNO-Appraisal: Some first results of a meta-evaluation on innovation policy evaluation – the case of Austria

The INNO-Appraisal Project is funded by the European Commission Directorate General Enterprise and Industry. The project team comprising Manchester Institute of Innovation Research (MIoIR, formerly PREST), Atlantis, Joanneum Research, Fraunhofer Institute Systems and Innovation Research and Wise Guys Ltd. is coordinated by MIoIR

INNO-Appraisal is one of eight initiatives supporting the European Commission's Directorate General Enterprise in the development and coordination of innovation policy. INNO-Appraisal takes stock of and assesses programme evaluations in the area of innovation policy across Europe. The aim of the INNO-Appraisal activity is to contribute to a better understanding as to how programme evaluation is currently used in innovation policy in Europe, and how evaluation contributes best to policy making. The project shall foster an evaluation discourse in the innovation policy community in Europe and also support the improvement of the Commission's analytical capabilities in the area of innovation policy. Together, this shall improve policy and the administrative environment for innovation in Europe.

For pursuing this aim, INNO-Appraisal collected all available appraisal reports on innovation policy measures that were running and represented within the Trendchart Database. The evaluation reports were made publicly available at the Innovation Policy Appraisal repository of ProInno (www.proinno-europe.eu). This contributes to the dissemination of evaluation results and improves the visibility of evaluations. By making use of a sophisticated and in-depth template, which had been filled in by the INNO-Appraisal team and then checked and complemented by the responsible policy makers in the European Union Member States, all available evaluations were systematically analyzed and put into a common database. This allows the INNO-Appraisal team to gather new knowledge about the evaluation practice in Europe and about characteristics that lead to useful and effective evaluations.

At present Phase 1 of the project, which includes the design of the data collection methodology, the data collection itself, the analysis of evaluation activity across Europe statistically and the set up and running of an online-repository of evaluation, is coming to a close. The data collection is almost finished and first results of the statistical analysis can be presented, although some additional datasets will have to be included and further statistical tests will be run.

In order to provide first illustrations of the information contained in the Inno Appraisal dataset, some indicative results are presented for the case of Austria. The article provides information on evaluation performers, timing and purpose of appraisals, topics covered in the appraisals, data collection and analytical methods used in the appraisals, perceived evaluation quality and usefulness of appraisals and the consequences of appraisals.

Compared to other European countries Austria has a remarkable number of policy measures in the field of innovation policy. In the last decade, about 60 RTI programmes were introduced to address structural weaknesses, foster science-industry relationships and close funding gaps. More than 50 Research, Technology and Innovation (RTI) policy evaluations were carried out between 2003 and 2007. In accordance with the number of existing RTI programmes, this is quite a lot for a country as small as Austria. The existence of a 'programme jungle' is widely discussed and a rising evaluation fatigue raises the question how results and recommendations of evaluations may better feed back into policy formulation and programme implementation (see CREST Policy Mix Expert Group: Country Report Austria, 2008).

In the INNO-Appraisal dataset, which rests upon the Trendchart database, Austria is the leading country in the sample according to the number of appraisals. It has 34 single appraisals for innovation policy measures as listed in Trendchart, which comprises roughly a quarter of the sample. 2 out of the 34 appraisals are portfolio evaluations dealing with more than one policy measure. In order to avoid bias in the statistical analyses, these portfolio evaluations are only considered once in the dataset. The above average representation is not only due to the high number of policy measures, but also due to the activities of the Platform Research and Technology Policy Evaluation (FTEVAL - http://www.fteval.at), which has published a compendium "Evaluation of Austrian Research and Technology Policies" (2007), containing a summary of Austrian evaluation studies. Furthermore the Platform FTEVAL contains an up to date database in which evaluation results of RTDI policy measures are made publicly available. As a consequence, availability of appraisal reports was much better than in other countries.

Evaluation performers

Evaluation performers are mainly national consulting and research companies, whilst international participation is limited. Thus, appraisal reports are usually written in German, but it is becoming common practice to add an English executive summary to the appraisal report. The majority of Austrian appraisals were conducted externally and the most tenders were allocated through a "closed" tender procedure, for which only a limited number of applicants is invited to participate in the tender procedure. The funding of appraisals covered in the database ranges between 10k Euros and 120k Euros; on average it shows a moderate level (48k Euros) which seems to correspond with an overall supportive character of evaluations (see below).

Timing and purpose of appraisals

The majority of appraisals are interim evaluations and have a largely formative purpose. *Ex ante* evaluations performed by external evaluators are rare in Austrian RTI policy, although there are studies and activities within ministries and funding agencies which contain elements of *ex ante* evaluation. These studies often go under the guise of "feasibility studies" introduced before, or at the beginning of, new initiatives. The demand for interim evaluations has risen dramatically over recent years and may partially be explained by the existence of the so-called RTDI directives (FTE-Richtlinien), which state that a "written evaluation concept must be provided, containing the goal, the aims, and the procedures, as well as the dates for controlling the achievement of the advancement aims for all programmes based on the RTDI directives). In this context, the interim evaluations tend to serve a supportive (formative) role aimed at enhancing or readjusting programmes.

Still, there are relatively few clear examples of *ex post* evaluations in Austria's RTI policy. Although many programmes are evaluated at a certain point in the programme's lifetime, in most cases insufficient time has passed to conduct a full range impact assessment including wider socio-economic impacts. Hence, the majority of evaluations have a formative character; especially interim evaluations are designed to give advice how programme management can be enhanced or readjusted. Only a small number of Austrian evaluations follow a summative purpose in order to judge the rationale and the socio-economic impact of a programme. Summative appraisals are carried out both ex post and interim and they are mainly commissioned at an advanced point in the programme's lifetime.

		Appraisal Purpose				
		summative	*formative*	*both*	*other*	*Total*
Timing	*ex ante*	0	5	0	0	5
	accompanying	1	2	1	0	4
	interim	4	10	4	0	18
	ex post	5	0	0	0	5
	other	0	1	0	1	2
	Total	10	18	5	1	34

Figure 1: Types and purpose of evaluations

Topics and appraisal methods

While the relatively low rate of impact assessment and additionality measurement seems to correspond with the formative purpose of most of the evaluations, outputs, outcomes and impacts are among the most frequent topics considered in appraisal exercises next to issues of policy and strategy development, programme implementation efficiency and consistency (see figure 2).

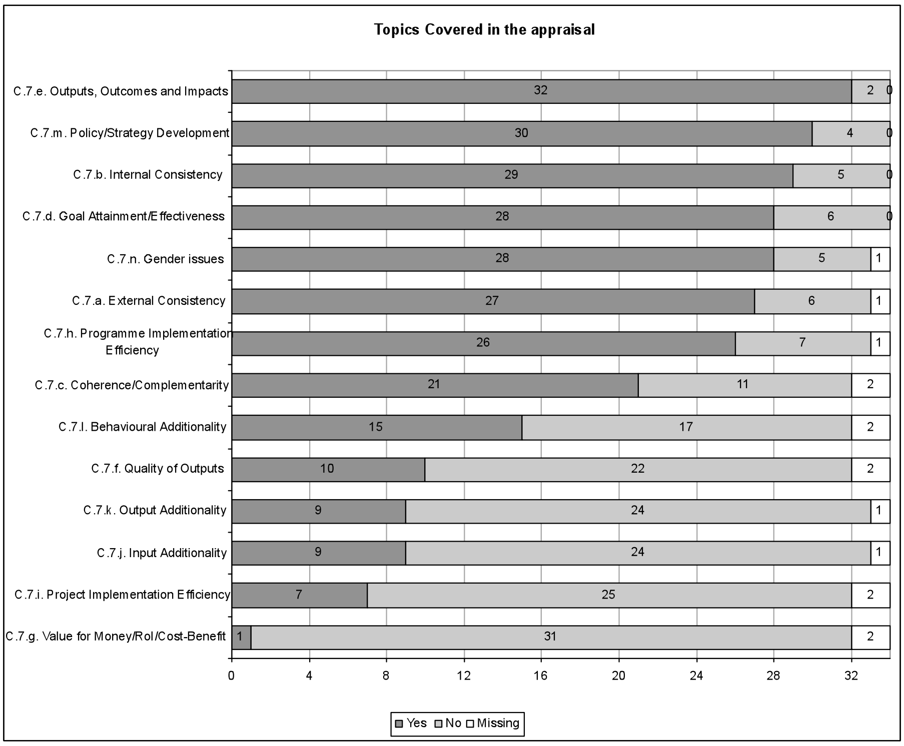


Figure 2: Topics covered in the appraisal

Looking at data collection methods employed, interviews and monitoring data are the most frequently used methods in Austrian appraisals (figure 3). Also monitoring data provided by the programme management is a key data source for evaluations. These results seem to accord with methodologies mainly employed in other European countries. Participant surveys, existing secondary databases and workshops are also heavily used. Regarding data analysis methods (figure 4) the majority of evaluations apply a mixed methodological approach where quantitative and qualitative methodologies are combined. Looking at the total of evaluations performed it is obvious that descriptive statistics, document analysis and context analysis build the core of data analysis methods. A trend to the employment of (social) network analysis is also noticeable. Network analysis is used to assess the programme's impact on collaboration and the relationship between researchers and organisations. More sophisticated quantitative methods (econometric analysis, control group approaches etc.) are only used in very specific cases although it has been acknowledged that RTDI policy measure evaluations have shown strong development in the last years by introducing Logit/Probit Analysis, Matched Pairs and Network Analysis for the first time (cf. Zinöcker 2007).

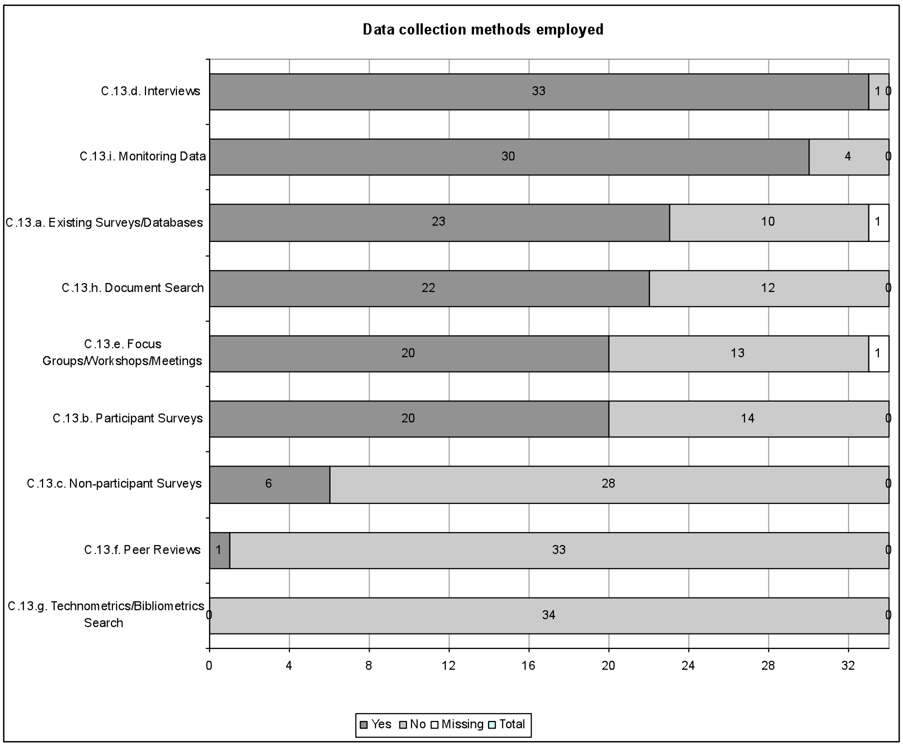


Figure 3: Data collection methods employed

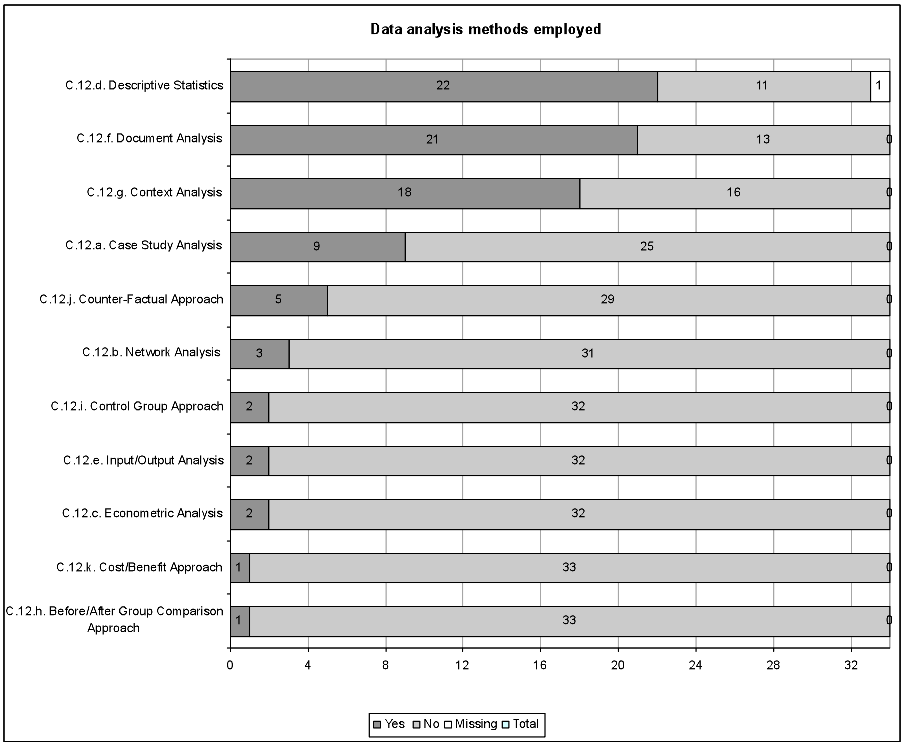


Figure 4: Data analysis methods employed

Overall, the relatively low level of funding and the formative character of Austrian appraisals seem to influence the choice of topics covered and methodologies used. Questions of additionality or impact assessments need to be scrutinised and bring along a sophisticated set of quantitative and qualitative methods which demands more efforts in time and cost.

Evaluation quality and usefulness

Despite the heavy reliance upon descriptive statistics, document analysis and context analyses respondents perceived that the quality of evaluations is high and, given the evaluation purpose, the design and methods employed tend to be considered as appropriate. Evaluation teams seem to choose a balanced set of methodologies which satisfy the purpose of the appraisal. The majority of appraisals focus on the usage of qualitative methods. Thus, the application of quantitative methods is not as satisfactory as of qualitative ones. The coverage of the broader context is not considered extensively in the Austrian evaluation reports rated.

As regards the usefulness of recommendations, policy makers had a bit less optimistic perspective: While recommendations concerning changes to the management and implementation of programmes were perceived to be useful and forward-looking advice was regarded as helpful for the design and implementation of future policy measures, recommendations regarding changes in the design of the measure and to other contemporaneous programmes only show average ratings. In this respect an important criticism was raised by the CREST policy mix review team “more thought should therefore be given to the mechanisms needed to ensure that the results of evaluations do feed back into policy formulation and implementation”. (see CREST Policy Mix Expert Group: Country Report Austria, 2008, p. 17). Although, recommendations of appraisals were discussed within different interest groups, mechanisms are missing to ensure their sustainability.

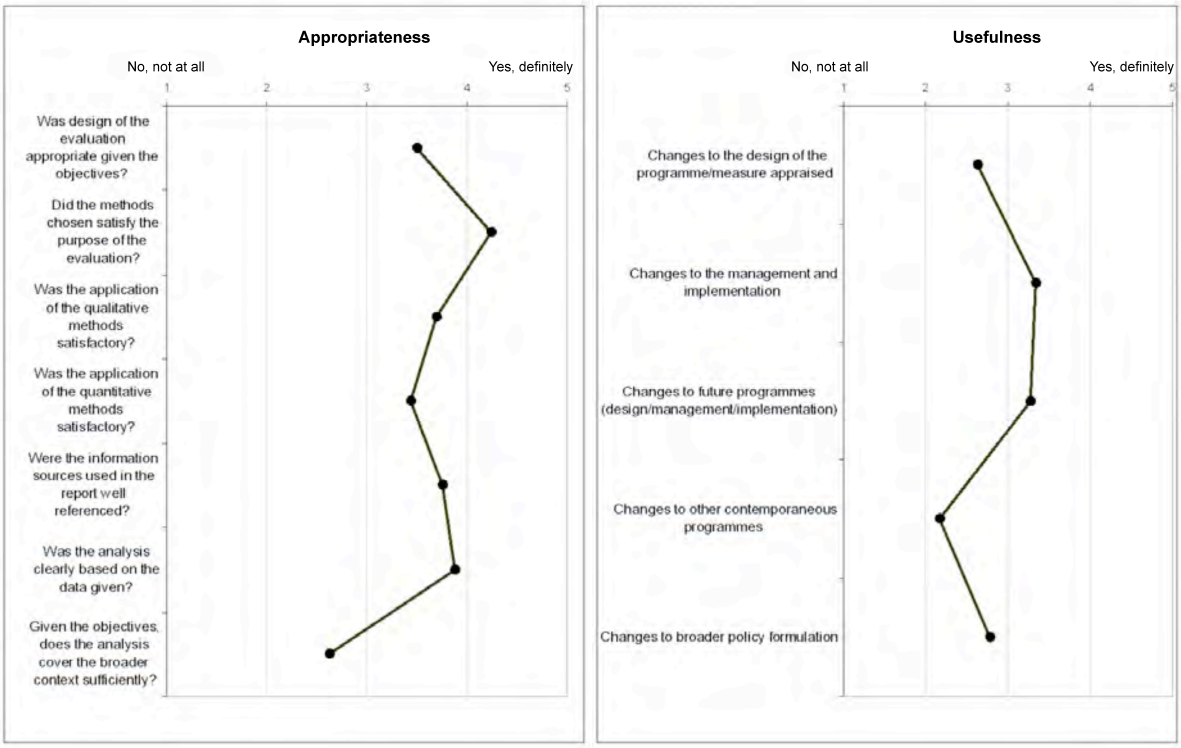


Figure 5: Quality and usefulness of appraisals

Consequences of appraisals

However, contrary to the CREST expert team’s criticism evaluations had consequences and recommendations were implemented and caused at least a minor re-design of policy measures. at least. Interestingly, the few ex post appraisals, which were carried out, convinced policy makers to re-design R&D programmes significantly. Mid-term appraisals led in two cases to a minor re-design of policy measures. Actually, the termination of one R&D programme was caused by an interim evaluation.

		Timing of the Evaluation			
		Ex ante	Accompanying	Interim	Ex post
Consequences	Termination of the measure	0	0	2	0
	Major re-design of the measure	1	0	3	0
	Minor re-design of the measure	1	3	8	0
	Expansion/Prolongation of the measure	1	1	4	0
	Re-design of another measure	1	0	2	1
	Merger of measures	0	0	6	2

Figure 6: Timing and consequences of appraisals

Summary and Outlook

Austria shows a remarkable number of policy measures which were introduced to address structural weaknesses, foster science-industry relationships and close funding gaps. Since the late 1990s, Austria has become a country strongly positioned in the field of policy evaluation.

The majority of appraisals are carried out mid-term during one point in the programme's lifetime. Mainly, a supportive purpose is followed as policy makers respective programme managers need advice how to enhance programme implementation. Thus, appraisals focus on policy/ strategy development and programme implementation efficiency. Looking at the data collection and analysis methods employed, it seems that low cost data gathering methods, e.g. descriptive statistics and direct outputs of evaluations, prevail. Interviews and the usage of monitoring data are the most frequently used data collection methods in Austrian appraisals. In Austria it is common practice to apply a mixed methodological approach where quantitative and qualitative methodologies are combined.

The quality of evaluations is perceived to be high and, given the evaluation purpose, the design and methods employed are widely considered to be appropriate. Recommendations made by the evaluation team were discussed within government circles and with participants and broader stakeholders. Especially, recommendations concerning changes to the management and implementation of RTI programmes were perceived to be useful. Forward-looking advice was regarded as helpful for the design and implementation of future policy measures.

Further statistical analyses are currently performed in order to retrieve information on the linkages between evaluation purpose, methods, and quality and usefulness. Beyond that phase 2 of the INNO Appraisal project will provide more in-depth analyses of specific aspects of evaluation conduct (Behavioural Additionality, Maximising Usefulness of evaluations, Structural Fund evaluation, economic and wider social impact). The studies will make use of the final INNO-Appraisal dataset and a whole range of interviews with policy makers. Final results are to be expected at the time of the end of the project, in March 2010.

Literature

CREST - European Union Scientific and Technical Research Committee (2008), Expert group report on the design and implementation of national policy mixes: Country Report Austria, http://www.era.gv.at/attach/CREST_Austrian_Policy_Mix_Report_-_September_2008-de.pdf

Zinöcker, Klaus (2007), Evaluating Austria's R&D Policies - Some personal comments, in Evaluation of Austrian Research and Technology Policies – A Summary of Austrian Evaluation Studies from 2003- 2007, Platform Research and Technology Policy Evaluation & Austrian Council for Research and Technology Development (eds.), Wien.

Author:

Michael Dinges
Joanneum Research ForschungsgesmbH
Sensengasse 1, A-1090 Wien
michael.dinges@joanneum.at

PLATTFORM fteval
Forschungs- und Technologieevaluierung

Die Systemevaluierung

Anstoß zu einer Neuausrichtung der angewandten Forschungspolitik?
(Oder: Wie kommt der Krampus in den Supermarkt?)

Veranstaltung, 4. 12. 2009, 9.00 – 16.15 Uhr.
Camineum der Österreichischen Nationalbibliothek, Josefsplatz 1, 1010 Wien

Im Mai 2009 wurden die Ergebnisse der Systemevaluierung präsentiert. Ziel der Evaluierung war es, die Forschungsförderung und -finanzierung in Hinblick auf die Leistungsfähigkeit des österreichischen Innovationssystems zu analysieren und allfälligen Handlungsbedarf zu seiner Verbesserung zu identifizieren.

4 Institute (WIFO, KMUFA, Prognos, convelop) wurden mit der Durchführung dieses Unterfangens der Systemevaluierung beauftragt. Im Rahmen der Untersuchung wurden 5000 Unternehmen und 1400 Forschungsinstitute zu ihrer Kenntnis, Nutzung und der Bedeutung von FTI- Maßnahmen befragt.

Das Evaluierungsteam fordert – laut seinem Selbstverständnis – einen radikalen Strategiewechsel der FTI – Politik: (i) einen umfassenden Ansatz der Innovationspolitik, (ii) eine Frontrunner-Strategie, (iii) mehr Koordination der Akteure, (iv) mehr Flexibilität der Instrumente, (v) eine klare Aufgabenteilung der Ministerien und schließlich (vi) eine klare Governancestruktur zwischen Ministerien und Agenturen.

Die Evaluierung hat aber auch einige Bilder unterhöhlt, die bis dato in der innovationspolitischen Diskussion prägend waren; ein solches Bild ist das des „Förderdschungels“. Ersetzt wird der Dschungel nun durch das Bild des Supermarktes „[Companies] do not get lost in a funding jungle, but deliberately pick from the best offers of a funding supermarket“. Obgleich (oder weil?) in diesem Supermarkt eine hohe Kundenzufriedenheit herrscht, bleibt die Forderung nach einem radikalen Systemwandel.

Wie ist ein solcher Wandel anzustoßen? Dies ist das Motiv dieser Veranstaltung der Plattform Forschungs- und Technologieevaluierung (Plattform fteval).

INFORMATION
Der Newsletter der Plattform Forschungs– und Technologieevaluierung ist ein unregelmäßig erscheinendes offenes Forum zur Diskussion methodischer und inhaltlicher Evaluierungsfragen in der Forschungs- und Technologiepolitik.

© Wien 2009 ZVR-Zahl: 937261837 ISSN-Nr.: 1726-6629

IMPRESSUM
Herausgabe und Versand: Dr. Rupert Pichler, Dr. Sonja Sheikh, Mag. Klaus Zinöcker, Mag. Leonhard Jörg
Plattform Forschungs– und Technologieevaluierung, Währinger Straße 3/15a, A-1090 Wien, office@fteval.at
Gestaltung: to-pixelate, A-1170 Wien, carmen@to-pixelate.org
Druck: Remaprint, Neulerchenfelderstrasse 35, 1160 Wien

FÜR DEN INHALT DIESER AUSGABE VERANTWORTLICH
Plattform Forschungs– und Technologieevaluierung: Mag. Klaus Zinöcker
Vorbereitet wurde diese Ausgabe von einer Arbeitsgruppe um Peter Lindner, Daniela Salhofer, Klaus Schuch, Michaela Topolnik und Petra Wagner-Luptacik.

PLATTFORM FORSCHUNGS– UND TECHNOLOGIEEVALUIERUNG

Die Plattform Forschungs- und Technologieevaluierung ist eine Initiative der folgenden Organisationen:
Österreichisches Bundesministerium für Wissenschaft und Forschung (bmwf), Bundesministerium für Verkehr, Innovation und Technologie (bm:vit), Bundesministerium für Wirtschaft, Familie und Jugend (bmwfj), Österreichische Forschungsförderungsgesellschaft mbH (FFG), Fonds zur Förderung der wissenschaftlichen Forschung (FWF), Joanneum Research, KMU Forschung Austria, Austrian Institute of Technology GmbH (AIT), Technopolis Austria GmbH, Österreichisches Institut für Wirtschaftsforschung (WIFO), Wiener Wissenschafts-, Forschungs- und Technologiefonds (WWTF) Zentrum für Innovation und Technologie GmbH (ZIT), Rat für Forschung und Technologieentwicklung, AQA – Österreichische Qualitätssicherungsagentur, Christian Doppler Gesellschaft (CDG), Austria Wirtschaftsservice (aws), Ludwig Boltzmann Gesellschaft, Zentrum für soziale Innovation (ZSI). Im Rahmen der Plattform werden Themenstellungen zur Forschungs- und Technologieevaluierung erarbeitet und – z.T. unter Einbeziehung namhafter ExpertInnen – in einem Fachkreis diskutiert. Der Newsletter beinhaltet Fachbeiträge zu Fragen der forschungs- und technologiepolitischen Evaluierung. Die Herausgabe erfolgt in zeitlicher als auch inhaltlicher Abstimmung mit Plattform-Veranstaltungen, um die Synergiewirkungen eines breiten Austauschforums zu nutzen.

Plattform Forschungs– und Technologieevaluierung
Währingerstraße 3/15a, A-1090 Wien

www.fteval.at