

## **Vorwort**

In der sechsten Ausgabe des Newsletters der Plattform Technologieevaluierung sind zwei aus unserer Sicht sehr wichtige Beiträge enthalten: Terttu Luukkonen aus Finnland stellt aktuelle Trends der Forschungs- und Technologieevaluierung in ihrem Land dar; aufgrund der sehr weit entwickelten Evaluierungskultur in Finnland lassen sich viele best practices für Österreich ableiten. Klarerweise handelt es sich dabei sowohl um formale Punkte als auch um ein gemeinsames informelles Grundverständnis.

Ein weiterer Beitrag vergleicht drei kürzlich fertiggestellte österreichische Technologieprogrammevaluierungen. Es handelt sich dabei um die Schwerpunkte Verkehrstechnik, Energietechnik und Umwelttechnik des Innovations- und Technologiefonds (ITF). Diese Ex post-Evaluierungen wurden einander gegenübergestellt und vor allem nach methodischen Gesichtspunkten von den Evaluatoren in der letzten Sitzung der Plattform Technologieevaluierung vor einigen Wochen präsentiert. Auf diesen Vorstellungen und einer umfassenden Diskussion baut der Beitrag auf, der auch einige Lehren für zukünftige Programmevaluierungen enthält. Damit soll kein Schlußpunkt in dieser Arbeit gesetzt werden, vielmehr sollen in den nächsten Monaten weitere konzeptive Arbeiten erfolgen.

Der Newsletter steht übrigens für alle einschlägigen Beiträge und Arbeiten offen (wenn möglich auf Englisch)! Eine interessante Lektüre wünschen Dorothea Sturn, Oliver Fritz, Gernot Hutschenreiter und Michael Stampfer.

## **Preface**

The sixth issue of our "Plattform Technologieevaluierung"-Newsletter contains a most interesting contribution by Terttu Luukkonen, one of Finland's leading RTD evaluators. In her article (based on an OECD conference paper) she describes trends in Finland's RTD evaluation culture. The second contribution, written by Oliver Fritz, Gernot Hutschenreiter and Dorothea Sturn, compares three recently finished evaluations of Austrian RTD programmes, "Energy Technology", "Environmental Technologies" and "Transport Technologies" within the Innovation and Technology Fund (ITF). Their synopsis shows that there is still room for institutional learning, though considerable methodological and strategic progress has already been made. Conceptual work on RTD programme evaluation will continue.

März 1998; Dorothea Sturn, Oliver Fritz, Gernot Hutschenreiter, Michael Stampfer

# Trends in research evaluation in Finland

*Terttu Luukkonen*<sup>1</sup>

## ***Stages of research evaluation in Finland***

There have been three stages to research evaluation in Finland. Research evaluation – in the sense of science and technology policy evaluation – was introduced in Finland in the early 1980s. In the early phases, it was largely a question of quality evaluations of both basic and mission-oriented research using panels of external experts.

Evaluation was introduced by the Academy of Finland, which is a system of research councils. The Academy started evaluating fields of research in 1982. The idea of evaluating research and the method used were borrowed from the practices of the Swedish Natural Science Research Council, although the method was applied in a slightly different way, *i.e.* to whole areas of research irrespective of whether or not they were funded by the Academy. A central feature of the evaluation procedure was the use of panels of external experts. In a small country like Finland, “external” means that the experts were researchers from other countries, mostly from outside the Nordic countries. A major evaluation criterion was the quality of research in relation to the international state of the art. To date, the Academy has carried out 23 evaluations of research fields and research institutes, and one evaluation of international research collaboration with an international institute (IIASA). A further evaluation is currently underway. Twenty-five evaluations, either completed or ongoing, is an impressive figure, although fairly modest considering that the first one was initiated 15 years ago. The number of evaluations is also much smaller than that envisaged by the Cabinet level advisory body, the Science and Technology Policy Council when, back in the early 1980s, it suggested that all research fields be subject to evaluation every ten years in a rolling system of evaluations.

The use of external – in practice, foreign – experts in quality evaluation was soon adopted by other research funding agencies in Finland. The ministries responsible for mission-oriented research, such as the Ministry of Environment and the Ministry for Agriculture and Forestry, adopted the procedure of the Academy in the late 1980s and applied it to the research funded by their respective ministries. Similarly, TEKES, the Technology Development Centre, a major governmental agency responsible for funding industrial technology programmes, evaluated its first generation of technology programmes in 1988-91. The first evaluations by the mission-oriented agencies and ministries were also quality evaluations. TEKES introduced the criterion of relevance from a practical point of view and commissioned professional external evaluators, for example from Technopolis in the United Kingdom, to carry out the evaluation exercises.

The second wave of evaluations was related to the evaluation of research organisations, research institutes and research funding agencies. At this stage, the research funding agencies which had been commissioning evaluations were themselves subjected to outside

---

<sup>1</sup> This is a shortened version of the author's paper entitled "The Increasing Professionalisation of the Evaluation of Mission-oriented Research in Finland: Implications for the Evaluation Process", published in: *Policy Evaluation in Innovation and Technology: Towards Best Practices*, OECD Proceedings, OECD, 1997, pp. 347-356. The author would like to express her gratitude to Tarmo Lemola for his useful comments and to Pirjo Niskanen and Sasu Hälikkä for their help in procuring the statistical information.

evaluation. Thus the Academy of Finland (Blume *et al.*, 1993), TEKES (Guillaume and Zegveld, 1995) and the Technical Research Centre of Finland (VTT) (Tomner and Zegveld, 1993), were evaluated by outside experts. Large public research institutes such as the National Public Health Institute (1995) and the Institute for Occupational Health were also evaluated in the 1990s. In these cases, evaluation concerned the extent to which these organisations had fulfilled their missions and, in the case of the Academy, the evaluation was initiated with the aim of achieving organisational reform. In all cases, evaluation was carried out by panels of external experts who obtained a wide range of background information; in the case of TEKES, the information included a survey among recipients of TEKES funding about the usefulness of that funding.<sup>2</sup>

By the mid- to late 1990s, a third phase of evaluation was emerging; evaluations have now become commonplace. All major public programmes for funding research and technology development and for the development of research infrastructure are subject to external evaluation. For example, the impacts of a major increase in public funding of research in 1997-99 will be monitored and evaluated on a real-time basis. The purpose of this increase in funds is, in particular, to strengthen the national system of innovation and to support enterprises, employment and the economy as a whole. The increase over three years corresponds to one-quarter of the total appropriations for government R&D funding in 1997. The money will be obtained through the sale of shares of government-owned enterprises. An example of an initiative on a much smaller scale is the evaluation of public support to regional centres of expertise<sup>3</sup> carried out in 1996 (Ahola and Kortelainen, 1997). At this latest stage, different type of evaluations, described above, are carried out simultaneously.

Why did research evaluations start and why were they fairly widely adopted? Originally, the Academy of Finland implemented the evaluations to provide arguments for obtaining more money for basic research by showing the uses to which funding was put and demonstrating that the research carried out was of fairly high quality. The initiative originated in the late 1970s when, after the recession of the early 1970s, there was a decrease in the funding of the Academy and the universities. In addition, most evaluations by the Academy were commissioned for specific reasons, such as a perceived need to develop a particular field or to assess major investments in equipment.

With regard to the mission-oriented agencies,<sup>4</sup> one of the reasons behind the implementation of evaluations was the need to improve quality control; another, a wish to re-orient re-

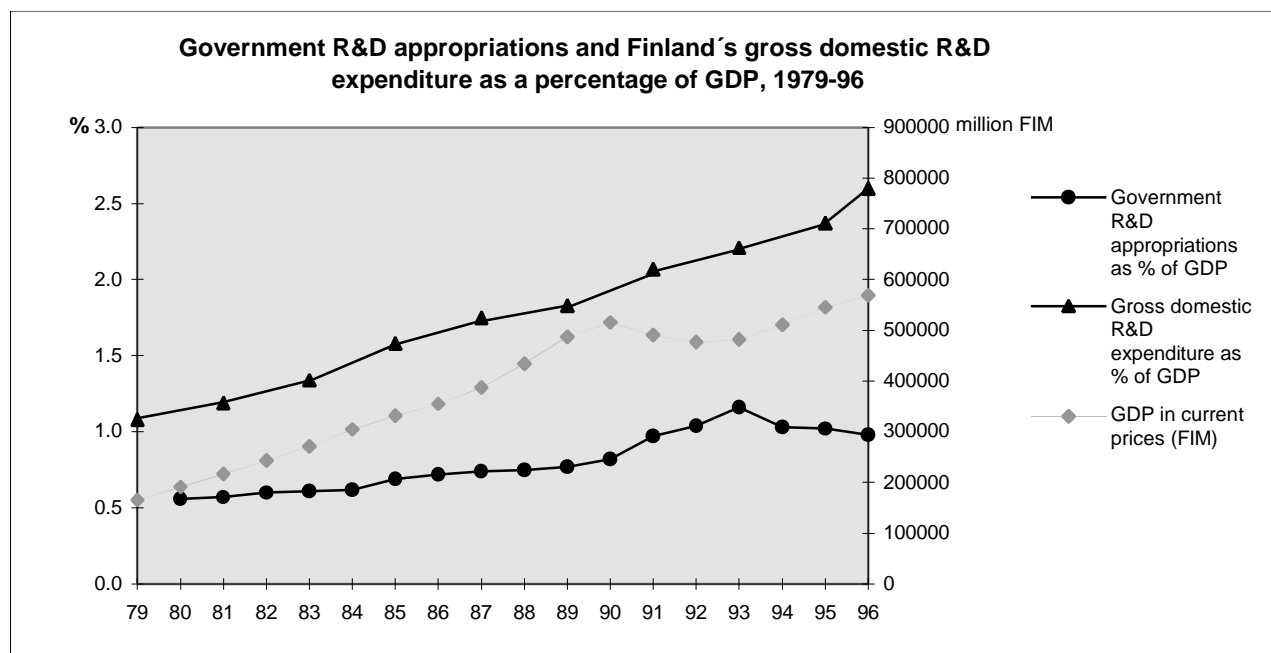
---

<sup>2</sup> Parallel to the evaluation efforts described above, there has been a movement towards quantitative evaluation of university research. The quantitative evaluation movement started in the mid-1980s simultaneously with, but independently of, a movement to improve effectiveness and increase accountability in the government sector. Quantitative evaluation was from the start related to efforts to apply evaluative criteria to the allocation of university resources. Alongside quantitative evaluation, there have been qualitative evaluations of whole universities. These have been learning exercises rather than evaluations related to any decisional needs. The most recent development is a move towards qualitative evaluation, such as to define centres of excellence by qualitative criteria, and at the same time, to increase the proportion of funds allocated by performance criteria. As an indication of the importance of evaluation as a management tool for universities, the former Council for Higher Education was reformed and its name changed to Evaluation Council for Higher Education at the beginning of 1996.

<sup>3</sup> The main purpose of the regional centres of expertise is to promote the establishment of new knowledge-and technology-based firms.

<sup>4</sup> Mission-oriented agencies refer to the ministries and units of central administration which finance research to attain different socio-economic objectives, such as health, transport, agriculture, defence, etc. The Ministry of Education has a prime responsibility to promote the "general advance-

search activities. Since the mid-1980s, the Cabinet-level Science and Technology Policy Council has paid particular attention to measures to improve the quality, effectiveness and relevance of research financed and/or carried out by the mission-oriented agencies. Evaluation has provided these bodies with a new management tool. The Science and Technology Policy Council of Finland has systematically endorsed evaluations and recommended that governmental research funding agencies commission them to assess the effectiveness and efficiency of the research activities funded. The deep recession of the early 1990s caused tight budgets in all fields of public administration and highlighted the importance of accountability in public funding of R&D. Government appropriations for R&D decreased slightly in 1994-96 after a long period of growth (see figure).<sup>5</sup> However, total national R&D expenditure has grown continuously throughout the 15-year period. Evaluations have been part of the overall strategy to build up a national research system, in terms of both quality and quantity.



### ***The professionalisation of evaluation***

The professionalisation of evaluation implies that professional evaluators or evaluating institutions carry out the evaluations in addition to, or as a substitute for, peer review/expert panels. The latter have traditionally been used to evaluate research quality in science. Even though peers are the best experts for this task, they are not professionals in evaluation. In professional evaluation, panels may also be used to gain a better understanding of the subject matter of the area to be evaluated and/or to provide the evaluation with more

---

ment of knowledge” through the funding of the universities and Research Councils of the Academy of Finland. The latter are excluded from the concept of mission-oriented agencies, although this chapter draws on their evaluation experiences.

<sup>5</sup> The sharp peak and subsequent decrease in the percentage of government R&D appropriations in GDP are caused by the trends in GDP. When measured in fixed prices, the changes are small.

credibility and prestige. There has been a trend towards more professionalisation in evaluation, both in Finland and, in particular, in other countries.

Along with the professionalisation of evaluation, evaluation criteria have shifted from those pertaining to quality of research towards a stronger emphasis on strategic aspects and factors pertaining to the relevance of the project from a practical point of view. The type of information used has shifted away from qualitative assessments by experts to more systematic, often quantitative, data gathering and to the application of standard social science methods for the exercise. Evaluations have become more professional, in particular where they concern mission-oriented research. A group of professionals, consultants and evaluating institutions engaged in evaluation activities has emerged. This has occurred in different ways in different countries. In Finland, it has happened on a modest scale.

Finland has aimed at independent and external evaluation by commissioning outside experts to carry out the evaluations. Often this has meant the use of experts from abroad either as members of expert panels or as evaluation consultants or a combination of both. Whether the experts and professionals come from abroad or from Finland, external evaluation implies that those commissioned to carry out the evaluation do not have a vested interest in it.

### ***Openness and publicity***

The Nordic countries have a tradition of openness in most areas of public administration. These principles pertain to the evaluation of institutions, programmes and activities funded by public money. The research evaluation movement, from its implementation in the early 1980s, adopted from the first a principle of openness and public reporting of evaluation findings. Publicity was important for both the transparency of the evaluation exercise and the legitimation of the activities to be evaluated. In some cases, this has given rise to unwanted and unexpected effects, such as negative media attention to criticised individuals, well beyond the intrinsic merits (demerits) of the cases considered. Another drawback is the reluctance of the evaluators to explicitly voice very critical opinions. This leads to indirect and hidden criticism and increases the interpretative flexibility of the reports – which is not usually in the interests of those commissioning the evaluation. Publicity does not only act as a sanction: researchers or research groups confident of their performance have in some instances sought to be evaluated in order to use the evaluation to support their status and to obtain further funds.

The fact that evaluations have become part of the everyday management routine of programmes and research institutions has decreased public and media interest in them. This is not wholly to be regretted as too much media attention is not beneficial to the communities being evaluated. In professional evaluation, there may be other, less desirable, threats to publicity and openness: professional evaluators may face pressure not to publish their findings if the evaluation is critical. Panels of prestigious experts face fewer problems in this respect.

### ***The use made of evaluations***

The use made of evaluations by the (so-called) decision makers and by the researchers themselves was analysed in the first wave of science policy evaluation in Finland and in the corresponding period in the other Nordic countries (Luukkonen and Ståhle, 1990; Luukkonen, 1995). I will not report the findings in detail, but I would like, however, to highlight that the everyday notion of the use of information – that is, instrumental use – is not the only way to utilise evaluation exercises. Instrumental use means that the information or, in our case, evaluation results, will be used to solve a concrete policy or decision-making

problem. Evaluation findings may be used in various, less direct, ways, such as to provide new insights into problems at hand, to formulate questions not previously asked or to throw light on the scope of problems. A further category of use – the use of evaluation findings to advocate a previously adopted standpoint or as ammunition to support a stand – is often morally reproved as a mis-use of evaluation findings. It is, however, more fruitful to acknowledge that any type of information can be used to promote various goals if there are people who have strong commitments to advance these goals. The same applies to evaluation findings: many evaluations have special purposes and hidden agendas which may be related to an advocative use, and they are used in such a way.

An example of this last category follows. The evaluation of the Academy of Finland was aimed at organisational change. There had been an initiative to merge the seven Research Councils into three, in the hope of increasing the effectiveness of the Research Councils and in response to management problems with an increasing number of research projects in interdisciplinary areas. The initiative had been blocked in Parliament for reasons irrelevant to the major initiative. After the 1991 Parliamentary elections, the organisational reform was given a new impetus by the Ministry of Education who commissioned a group of foreign experts to carry out an evaluation of the Academy of Finland in 1992. Not surprisingly, in their report, the group of experts came to a conclusion similar to that of the original reform initiative. In the end, the number of Research Councils was reduced but, after successful lobbying by the medical scientists, who managed to retain their own Research Council, the number of the Research Councils was eventually reduced to four. Why did the experts come to the same conclusion as those advocating the organisational reform? Obviously, they were largely dependent on the information provided by the major players in science policy who were, by and large, in favour of the reform. Another factor is that merging separate Research Councils into one large Council or into a smaller number of larger units seems to be the trend in this decade – at least in the Nordic countries, which have a tradition of following each other's examples.

The consequences of the evaluation of the VTT were similar to those of the Academy; that is, they led to organisational change. However, the situation was quite different. The evaluation (1992-93) was not initiated with any particular aim in mind, but rather as a response to a vague dissatisfaction with the ability of the institute to respond to the needs of industry – and also because evaluation was becoming increasingly popular as a management tool. The initiative for the evaluation came from the Ministry of Trade and Industry. VTT is a large contract research institute with approximately 2 500 employees. It operates under the Ministry of Trade and Industry, which set up a two-man evaluation committee (Sigvard Tomner and Walter Zegveld). The two wise men recommended, among other things, a radical reorganisation of the institute: merging 34 research laboratories into ten units for reasons of efficiency, synergy and strategic management. The question of potential organisational change had been under discussion within the institute, but it was not actively on the agenda. It is, therefore, somewhat surprising that after the evaluation the recommendation for radical organisational reform was promptly put into effect. The director general of the institute became an active "champion" for change. The starting point was more open-ended than in the example of the Academy. The outcome may be related to many factors; among others, the need for the management of the institute to assume the initiative after the Ministry-led evaluation process.

In drawing attention to the uses and impacts of evaluations, too little attention is paid to the process impacts: the fact that activities, institutions, etc., are being evaluated does have an impact on those involved in the process. It may be simply that they are forced to think of their activities over a longer time perspective than would routinely be the case, or that the evaluation provides new contacts for the people involved in the activities being evaluated.

We do not have enough information on these impacts. The process effects may be useful to the evaluated institutions in helping people reorient their activities or reinforce their current plans. However, process impacts can also be negative: the evaluation process is a source of uncertainty for staff in the institution being evaluated and may be detrimental to their work motivation.

We have to acknowledge that evaluation plays a limited role in decision making; providing only part of the background information. Demands that evaluation recommendations be always put into practice are unrealistic and represent an unduly rationalistic view of the decision processes. This view contains an implicit faith in the rationality of the evaluation process and the superiority of the knowledge thus acquired compared with the knowledge previously held by the people concerned. Local decision makers have the responsibility for carefully judging the evaluation findings and recommendations and coming to their own conclusions. The evaluators, especially if they are outsiders, are not (at least, not always) in a position to come up with the best possible or most realistic practical solutions for perceived problems or performance failures.

Decision makers and those commissioning the evaluations will benefit from an open-minded reading of the evaluation reports. The information provided by the evaluation may provide them with new and unexpected insights; it may challenge their traditional ways of thinking and help formulate new solutions to problems. However, it is *their* responsibility to draw the conclusions: if the end result does not follow the recommendations of the evaluators, this may be unsatisfactory from the evaluators' point of view, but it is part of the local research and technology policy context.

The evaluation exercise and the evaluated activities gain in credibility if the evaluation is transparent for all the parties concerned. This includes informing both the evaluators and those being evaluated as to follow-up, potential changes in the activities evaluated, and plans concerning the monitoring of these changes.

## **References:**

- AHOLA, Eija and Sami KORTELAJAINEN (1997), *Osaamiskeskusarviointi* (The Evaluation of the Centres of Expertise), Sisäasiainministeriö, Aluekehitysosaston julkaisu 1/97.
- BLUME, Stuart, Leif Arne HELOE, Peder Olesen LARSEN, and Michael V. POSNER (1993), *The Academy of Finland - An International Evaluation 1992*, Opetusministeriö, Koulutus- ja tiedepolitiikan linjan julkaisusarja, Helsinki.
- EVALUATION OF THE NATIONAL PUBLIC HEALTH INSTITUTE OF FINLAND - KANSANTERVEYSLAITOS (1995), *Report of the Evaluation Panel*, Publications of the Academy of Finland, 9/95, Painatuskeskus, Helsinki.
- GEORGHIOU, Luke (1994), "Impact of the Framework Programme on European Industry", Office for Official Publications of the European Communities, Luxembourg, EUR 15907 EN.
- GUILLAUME, Henri and Walter ZEGVELD (1995), *The Technology Development Centre of Finland, TEKES - An International Evaluation*, Finnish Ministry of Trade and Industry, Publications 5/1995.
- LUUKKONEN, Terttu and Bertel STÅHLE (1990), "Quality Evaluations in the Management of Basic and Applied Research, *Research Policy*, Vol. 19, No. 4, pp. 357-368.
- LUUKKONEN, Terttu (1995), "The Impacts of Research Field Evaluations on Research Practice", *Research Policy*, Vol. 24, pp. 349-365.
- TOMNER, Sigvard and Walter ZEGVELD (1993), *The Technical Research Centre of Finland (VTT) - An International Evaluation*, Finnish Ministry of Trade and Industry, Publications 2/1993.

**Terttu Luukkonen:**

Senior Researcher at VTT Group for Technology Studies, P.O.Box 1002, FIN-02044 VTT, Finland.



# **Drei österreichische Programmevaluierungen im Vergleich: Energietechnik, Verkehrstechnik und Umwelttechnik**

*Dorothea Sturn  
Oliver Fritz, Gernot Hutschenreiter*

Das zeitliche Zusammenfallen der Evaluierung von drei recht ähnlichen Schwerpunktprogrammen - bei allen dreien handelt es sich um nationale Programme des österreichischen Innovations- und Technologiefonds (ITF), - war Anlaß, in einer Plattform den Vergleich dieser drei Evaluierungen zum Thema zu machen. Dafür wurden die Evaluatoren gebeten, ihre Ergebnisse anhand einer vorgegebenen Themenliste in vergleichbarer Weise zu strukturieren und vorzustellen. Der vorliegende Artikel stellt das Ergebnis von Präsentation und Diskussion dieser drei Ansätze dar. Wir möchten an dieser Stelle den Evaluatoren für ihre Präsentation danken, die gleichermaßen die Basis für die folgenden Ausführungen darstellen: Claudia Pichl (Wirtschaftsforschungsinstitut) für die Umwelttechnik, Fritz Ohler und Leonhard Jörg (Österreichisches Forschungszentrum Seibersdorf) für die Verkehrstechnik und Manfred Bruck (Kanzlei Dr. Bruck) für die Energietechnik.

Der ITF ist das derzeit einzige Instrument der österreichischen FTE-Förderung, in dem planmäßig Ex-post Evaluierungen und Ex-ante Programmplanungen und -evaluierungen vorgenommen werden. Dabei wird dieser Planerrolle auch dadurch entsprochen, daß ein gemeinsamer Rahmen für solche Bewertungen erarbeitet und der Arbeitsfortschritt auch öffentlich diskutiert wird.

## **1. *Kenndaten der evaluierten Programme***

Da es sich bei allen Programmen um ITF-Schwerpunktprogramme handelt, ist ihnen eine Fokussierung auf bestimmte Themen und Schirmprojekte gemeinsam. Dennoch sind Art und Ausmaß dieser Fokussierung unterschiedlich: Sie reichen von einer starken Konzentration auf relativ wenige Projekte und Firmen im Bereich der Verkehrstechnik bis zu einem stärker diffusionsorientierten Ansatz mit nur wenigen Schirmprojekten im Bereich der Umwelttechnik. Lediglich bei der Energietechnik gab es ein externes, begleitendes Programm-Management.

	<b>Verkehrstechnik</b>	<b>Umweltechnik</b>	<b>Energietechnik</b>
<b>Programmziele</b>	<p>Technologische Höherqualifizierung</p> <p>Bildung von Anbietergemeinschaften</p> <p>Auf- und Ausbau nationaler Kompetenzzentren</p>	<p>Umwelt- (besserer Umgang mit Umweltproblemen) und technologiepolitische Zielsetzungen (Wettbewerbsfähigkeit)</p> <p>Produktionsintegrierter Umweltschutz, Vermeidung, Cleaner Technologies</p> <p>ab 1993 zusätzlich:</p> <p>Verbesserung der Diffusion</p> <p>Verbesserung der Teilnahme an europäischen Programmen</p>	<p>spezifische energie- und umweltpolitische Zielsetzungen (Energiesparpotentialnutzung, Nutzung erneuerbarer Energieträger, Verminderung der CO<sub>2</sub>-Emissionen) und allgemeine technologiepolitische Ziele (Wettbewerbsfähigkeit)</p>
<b>Fördergebiet</b>	Österreich	Österreich	Österreich
<b>Zielgruppe</b>	Verkehrsmittelindustrie in Österreich	Anbieter innovativer Umwelttechnologien und Systemlösungen seit 1993 auch Anwender und verstärkt KMUs	
<b>Förderkriterien bzw. förderbare Vorhaben</b>	<p>ITF-Kriterien</p> <p>Konformität mit den zentralen Anliegen der österreichischen Verkehrspolitik</p> <p>Vorliegen von Machbarkeits- und Wettbewerbsanalysen</p> <p>Projektmanagement</p>	<p>ITF-Kriterien</p> <p>höhere Förderquoten bei Kooperations- bzw. Schirmprojekten sowie bei Diffusion und Konzeptförderung</p>	<p>ITF-Kriterien</p> <p>Gefördert wurden auch Pilot- und Demonstrationsanlagen (Vorlage eines Evaluierungskonzepts)</p> <p>betriebliche Energiekonzepte</p> <p>Integration bekannter Komponenten zu neuen Systemen</p>
<b>Programmlaufzeit</b>	15.12.1992 - 14.12.1997	1988-1995	15.12.1992 - 31.12.1997
<b>Programm-Management</b>	Management innerhalb des Ministeriums; externes Management nur teilweise auf Schirmebene	Externes Management war geplant, kam aber teilweise über die Definitionsphase nicht hinaus	Umfassendes externes Management (Durchführung von Workshops, Beratung von Firmen, Betreuung der Schirmthemen, Herausgabe von Newslettern)

<b>Themen</b>	<p>durch Fokussierung auf Systementwicklungen fast ausschließlich Schirmprojekte, zu denen die einzelnen Projekte gebündelt werden:</p> <p>Lärmarme Bahn, Lärmarme Straße, Logistik Forschungsterminal, Logistikleitsysteme, City-Logistik</p>	<p>Konzentration auf vorsorgenden Umweltschutz und Cleaner Technologies</p> <p>seit 1993 Eingrenzung auf Umwelttechnik, Umweltverfahrenstechnik, Cleaner Technologies</p> <p>Erweiterung auf Diffusion und Konzeptförderung</p> <p>Durchgeführte Schirmprojekte: Prepare und Öko-Audit</p>	<p>Konzentration auf 4 Schirmthemen: Integrierte Energiesysteme, Beleuchtung, Kleinf Feuerungsanlagen und Sonnenwärme</p>
---------------	--	--	---

Für die Evaluatoren von besonderer Bedeutung erwiesen sich die beiden Punkte Programmziele und Programmanagement. Bei den Programmzielen sind unklare und unspezifische Vorgaben vielfach ein Problem. Sehr allgemeine Zielformulierungen, typischerweise die Steigerung der Wettbewerbsfähigkeit lassen sich ohne nähere Erläuterungen der intendierten Wirkungszusammenhänge nur schwer überprüfen. Oft auch stehen die formulierten Ziele in keinem erkennbaren Zusammenhang zur inhaltlichen und organisatorischen Ausgestaltung der Programme. Dies geht so weit, daß die Zielformulierungen in verschiedenen Dokumenten (Richtlinien, Arbeitsberichte u.ä.) nicht identisch sind. In keinem der Programme fanden sich quantifizierte Ziele.

Weiterhin schwierig erwies sich das Management der Programme: Top-down-Programme benötigen prinzipiell umfangreiche Management- und Moderationsleistungen (ob nun von externen Einrichtungen wahrgenommen oder innerhalb der Ministerien bzw. Fonds abgewickelt). Diese Programm-Abwicklung ist das zentrale Thema für die Design-Evaluierung. Lediglich bei Energietechnik gab es ein Programm-Management mit wohldefinierten Aufgaben, die eine solide Überprüfung zuläßt.

## **2. Beschreibung und Analyse der Evaluierungsstudien**

### **2.1 Allgemeine Angaben**

	<b>Verkehrstechnik</b>	<b>Umwelttechnik</b>	<b>Energietechnik</b>
<b>Evaluierungstyp</b>	überwiegend ex-post, kleine Programmteile ex-ante	überwiegend ex-post, daneben ex-ante	ex-post
<b>Evaluierungslaufzeit</b>	6 Monate (7/97-12/97)	1 Jahr (10/96-9/97)	2 Monate
<b>Anteil des Evaluierungsbudgets am Programmvolumen</b>	0,38% des Förder volumens	0,6%	3,45%
<b>Evaluierungsteam</b>	österreichische und britische Experten	österreichische und deutsche Experten	österreichische Experten

### **2.2 Evaluierungsziele und Vorgaben des Auftraggebers**

In den drei Fällen waren die Evaluierungsziele von seiten des Auftraggebers unterschiedlich eng spezifiziert. Beim Schwerpunkt Energietechnik war die Evaluierung des Programm-Managements expliziter Teil des Auftrags, ansonsten wurde hier - wie bei der Umwelttechnik auch - die Konzeption der Evaluierung den Evaluatoren überlassen. Die Vergabe der Evaluierung des Verkehrstechnik-Schwerpunkts erfolgte mittels Ausschreibung, entsprechend detaillierter und expliziter waren hier auch die Vorgaben vor allem bezüglich Programmdurchführung, Organisation und Abläufe. Eine zu erwähnende Besonderheit beim Schwerpunkt Umwelttechnik war die Intention des Auftraggebers, über die Wahl eines internationalen Evaluierungspartners, im Zuge der Untersuchung allgemeines Evaluierungs-Know-how nach Österreich zu bringen.

Konkret wurden die Vorstellungen und Wünsche der Auftraggeber vor allem dort, wo Empfehlungen hinsichtlich der künftigen Ausgestaltung weiterer Programme gefragt waren, also bei den ex-ante Teilen im Schwerpunkt Verkehrstechnik und Umwelttechnik. In beiden Fällen zeigte sich ein hohes kommunikatives Interesse der Auftraggeber (vor allem in der Umwelttechnik gab es etwa alle 6 Wochen einen intensiven Erfahrungsaustausch zwischen Auftraggeber bzw. dem projektbegleitenden Beirat und Auftragnehmer) gekoppelt mit sehr hohen Erwartungen. Im Schwerpunkt Verkehrstechnik handelte es sich um Vorschläge hinsichtlich "Europäischer Zukunftsbereiche", bei der Umwelttechnik um die Vorbereitung eines Folgeprogrammes "Innovationen für Nachhaltigkeit".

In der Diskussion der Evaluierungspräsentationen wurde das fallweise Fehlen von Evaluierungskriterien und -wünschen des Auftraggebers vielfach kritisiert. Für künftige Programme und Evaluierungen wurde der Wunsch ausgesprochen, einen knappen, klar formulierten Satz an Evaluierungskriterien zu erstellen, der in erkennbarem Zusammenhang zu den Programmzielen steht, aber den Evaluatoren kein zu enges Konzept vorgibt.

### 2.3. Inhaltliche Konzeption der Evaluierung

Folgende Evaluierungselemente kamen jeweils zur Untersuchung:

	<b>Verkehrstechnik</b>	<b>Umwelttechnik</b>	<b>Energietechnik</b>
<b>Konzeptevaluierung</b>	für die ex-post-Teile: integriert in Gesamtstudie, die detaillierte Vorschläge für eine optimale Programmdefinition im Rahmen des "Politikzyklus" enthält für die ex-ante-Teile: Grundsätzliche Fragen der Konzeption	für die ex-post-Teile: Überprüfung der dem Programm zugrundeliegenden Annahmen (z.B. sind finanzielle Engpässe ein Innovationshindernis?) für die ex-ante-Teile: Grundsätzliche Fragen der Konzeption	geringer Anteil, in der Gesamtstudie integriert
<b>Designevaluierung</b>	zentral und integriert in den Gesamtansatz: Rekonstruktion und Evaluierung des gesamten Politikzyklus (Programmdefinition + Implementierung)	nicht im Zentrum, aber folgende Kernfragen: Waren die Förderinstrumente den Zielsetzungen angemessen? Wie ist der Top-down-Ansatz zu beurteilen?	nicht im Zentrum
<b>Prozeßevaluierung</b>	in gleicher Weise integriert	nicht im Zentrum, aber Fragen der Programmdurchführung und des Managements wurden gestellt	gemeinsam mit Wirkungsanalyse im Zentrum der Evaluierung: Die Beurteilung des Programm-Managements, die administrative und operative Durchführung des Programms
<b>Wirkungsanalyse (Zielerreichung, Programmeffizienz, Programmeffektivität)</b>	Effektivität im Zentrum (vor allem direkte Effekte wie Etablierung von Netzwerken, Additionalität, Beteiligung an EU-Rahmenprogrammen) Programmeffizienz wurde sowohl bezüglich der Qualität der Projekte als auch bezüglich Programmablauf und Management untersucht (vgl. integrativer Ansatz)	Zielerreichung wurde vor allem in Hinblick auf umweltpolitische Zielsetzungen untersucht ansonsten Schwerpunkt auf Programmeffektivität (direkte und indirekte Effekte) Programmeffizienz wurde z.B. bei der Untersuchung der Mitnahmeeffekte angesprochen	Zielerreichung (technologie- und umweltpolitische Zielsetzungen) sowie Effektivität im Zentrum

Alle drei Evaluierungen enthalten nahezu alle Elemente, allerdings in sehr unterschiedlicher Gewichtung und auch auf unterschiedlichem Niveau: Konzeptevaluierungen waren naheliegenderweise dort gefragt, wo die Studien explizite ex-ante-Teile aufwiesen, d.h. vor allem bei der Verkehrs- und Umwelttechnik. Ein sehr integrativer Ansatz, der grundsätzliche Fragen der Konzept-, Design- und Prozeßevaluierung zusammenführt, wurde bei der Verkehrstechnik gewählt, wo die Rekonstruktion und Evaluierung des gesamten Politikzyklus (das heißt Programmdefinition und Implementierung) im Vordergrund stand. Bei der Umwelttechnik befand sich eine vergleichsweise anspruchsvolle Wirkungsevaluierung im Zentrum, während die Energietechnik aufgrund der expliziten Vorgabe einer genauen Überprüfung des Programm-Managements auch Elemente der Prozeßevaluierung enthält.

## 2.4 Methodik

	Verkehrstechnik	Umwelttechnik	Energietechnik
<b>Firmenbefragungen schriftlich</b>	**	***	***
<b>qualitative und quantitative Auswertung</b>	**	***	***
<b>Kontrollgruppe bzw. Vergleichsgruppe</b>		***	
<b>Interviews „Programm-Manager“</b>	***	**	*
<b>Interviews „Policy Makers“</b>	***	**	*
<b>Auswertung Programmstatistik und Projektdokumentationen</b>	***	**	**
<b>Komparative Analyse (andere Förderprogramme EU oder andere Länder)</b>	*	*	*
<b>Vertragsprüfung</b>			*
<b>quantitative Indikatoren</b>		*	

*(Die Anzahl der Sterne symbolisiert den Stellenwert der jeweiligen Methode im Kontext der Gesamtstudie)*

Bei allen drei Evaluierungen wurde in erster Linie mit Primärerhebungen gearbeitet. Die Auswertung der Firmenbefragung vor allem für Zwecke der Wirkungsanalyse folgte bei der Umwelttechnik durch sehr sorgfältige Konstruktion einer Kontrollgruppe einem vergleichsweise anspruchsvollen Design. Auch wurden hier zusätzlich - und vorsichtig, wissend um die begrenzte Aussagekraft solcher Indikatoren - einige quantitative Indikatoren herangezogen.

Beim Schwerpunkt Energietechnik wurde viel Mühe auf möglichst hohe Rücklaufquoten aufgewendet, die Auswertung der Firmenbefragungen erfolgte dann qualitativ nach den einzelnen Schirmthemen.

Die Verkehrstechnik unterscheidet sich aufgrund des bereits erwähnten spezifischen Ansatzes (Politikzyklus-Evaluierung) auch methodisch etwas, hier stehen die Interviews mit „Policy-Makers“ und Programm-Managern im Zentrum.

### *3. Schlußfolgerungen und Empfehlungen für künftige Evaluierungen*

Von allen Evaluatoren wurden **Datenprobleme** genannt, die ihre Arbeit erheblich beeinträchtigt haben. Hier war nicht nur vom mangelnden Zugang zu bestimmten Daten die Rede, sondern auch von einer mangelnden Strukturierung und Aufbereitung dieser Daten (keine EDV-mäßige Erfassung der Programm Daten, mangelhaft dokumentierte Programmabläufe etc.). Besonders deutlich zeigte sich letzteres beim Schwerpunkt Verkehrstechnik, wo in verschiedenen Dokumenten unterschiedliche Programmziele und Zielgruppen genannt waren.

Ebenfalls einig waren sich Evaluatoren wie auch die Mehrzahl der Diskussionsteilnehmer darin, daß eine **sorgfältige Formulierung der Programmziele** (das beinhaltet auch den Konsens aller Programmteiligten und -verantwortlichen über die konkreten Zielsetzungen) zentrale Voraussetzung für aussagekräftige Evaluierungen ist. Dieses Problem hat mehrere Facetten:

- Erstens wurde das Nebeneinander von sehr langfristigen abstrakten und sehr kurzfristigen, konkreten Zielen kritisiert (Zitat eines Evaluators: „Manche Ziele richten sich auf den Polarstern, andere nach Klosterneuburg“)
- Zweitens - und damit im Zusammenhang - stellt sich die Frage nach **dem optimalen Grad an Operationalität und Quantifizierung der Ziele**. Ungenaue Zielformulierungen stellen zentrale Hindernisse für die Evaluatoren dar: Beispielsweise konnte bei der Evaluierung des Umwelttechnik-Programms eine Zielerreichung in Bezug auf die umweltpolitischen Zielsetzungen untersucht werden, da diese genau und konkret benannt wurden, eine ähnliche Analyse fehlt für die technologiepolitischen Zielsetzungen aufgrund deren ungenauer Formulierung. Eine zu allgemeine Zielformulierung (typischerweise in der Formulierung „Steigerung der Wettbewerbsfähigkeit“ zu finden) ist dabei ebensowenig handlungsanweisend wie eine überexakte Quantifizierung von Programmzielen (typischerweise in einigen EU-Programmen zu finden in Formulierungen wie: „Maßnahme x hat in 3 Jahren 2000 Arbeitsplätze zu schaffen“). In diesem Spannungsfeld das richtige Maß zu finden, bleibt eine Aufgabe für das Design künftiger Programme.
- Drittens sollen die **Programmziele in einem erkennbaren Bezug zu den Evaluierungsparametern** stehen und idealerweise beides bereits zum Zeitpunkt des Programmstarts formuliert sein. Dies erleichtert nicht nur die Arbeit der Evaluatoren sondern gibt auch zusätzlich hilfreiche Transparenz für die Programmteilnehmer. Darüber hinaus kann durch die Vorgabe von Evaluierungsparametern und bestimmter methodischer Minimalerfordernisse eine Basis dafür geschaffen werden, Evaluierungsergebnisse künftig besser vergleichbar zu machen.

Neben den oft mangelhaften Vorgaben von seiten des Auftraggebers gab es auch ein **Zuviel unterschiedlichster Wünsche**. Beispielsweise wurden sehr hohe inhaltliche und kommunikative Anforderungen an die ex-ante-Teile bei den Schwerpunkten Verkehrs- und Umwelttechnik gestellt. Bei der Umwelttechnik sollte die Evaluierung auch noch zusätzlich dazu beitragen, allgemeines Evaluierungs-Know-how nach Österreich zu bringen.

Insgesamt zeigte sich, daß Anstrengungen zur Verbesserung künftiger Evaluationen an verschiedenen Punkten gleichzeitig ansetzen müssen: Erstens an der Formulierung des Programms selbst, seiner Zielsetzung und Implementierung, zweitens an den Vorgaben für die Evaluatoren und drittens auch an der Verbesserung des allgemeinen Evaluierungs-Know-hows sowohl auf seiten der Programm-Verantwortlichen als auch auf seiten der Evaluatoren.

## **English Summary**

The article compares the evaluations of three Austrian technology programs, all of which were part of the Austrian innovation and technology fund (ITF): The programs targeted environmental technology, transport technology and energy technology, respectively. The comparison is mainly based on information made available by the evaluators themselves, who were asked to fill in a questionnaire including questions on evaluation goals, main issues to be examined by the evaluators and methodology applied.

Based on this comparison of the three evaluation studies, the following conclusions can be drawn together with some recommendations concerning the improvement of the Austrian evaluation practice:

All evaluations were negatively affected by serious problems with respect to data availability: Both access to and quality of the data material on participating firms was insufficient.

Serious problems for the evaluators arose due to the lack of accurately defined program goals: goals were either very long run or very short run oriented; operational goals or quantitative program targets, which are necessary to measure the success or failure of a program, were often missing. Ideally, objectives should be set by program planners such that the program effectiveness, i.e. the degree of goal achievement, can be measured accurately by the evaluators later on.

Evaluators had to cope with high demands by the institutions commissioning the evaluations, both with respect to the communication and reporting requirements as well as the number of issues to be addressed and goals achieved by the evaluation.

Overall, the benefit from evaluations could be increased if programs are designed in anticipation of their future evaluation; furthermore, both evaluators and commissioners of evaluations need to improve their evaluation know-how.

<p><b>Dorothea Sturn</b>, JOANNEUM RESEARCH <b>Oliver Fritz</b>, JOANNEUM RESEARCH <b>Gernot Hutschenreiter</b>, WIFO</p>
---



Für den Inhalt verantwortlich:

Bundesministerium für Wissenschaft und  
Verkehr - Dr. Michael Stampfer  
Renngasse 5, A-1010 Wien  
E-mail:michael.stampfer@bmv.gv.at

weitere Kontaktadressen:

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

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

Dr. Oliver Fritz  
Joanneum Research  
Wiedner Hauptstraße 76  
A-1040 Wien  
E-mail: oliver.fritz@joanneum.ac.at

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

Internet Adresse der „Plattform Technologieevaluierung“:  
[www.bmv.gv.at/Wirtschaft/wirt2.htm](http://www.bmv.gv.at/Wirtschaft/wirt2.htm)