

# The contribution of evaluation information to decision-making

- R&D evaluation in the field of science and technology policy in Finland

## Abstract

The paper discusses the importance of evaluation information for science and technology policy-makers in Finland. During the last two decades the volume and diversity of evaluations concerning the public research and development system have grown significantly. Today, they cover everything from funding agencies to research institutes, universities, and various R&D programmes. Over time, the interest behind decisions to launch an evaluation of research activities has shifted somewhat, but the main explicit purpose of the evaluations has been to provide knowledge on different facets of the national R&D system for strategic policy formation and decision-making processes.

The following questions will be raised against the above background: (i) what is the importance of evaluations for decision-makers when compared with other information sources at their disposal; (ii) how informative and usable are evaluations from the decision-makers' point of view; and (iii) how current practices should be developed so that more relevant information could be delivered in the future. The analysis is based on interviews conducted among actors of the public R&D system.

## Introduction

This paper will examine the role of evaluation in decision-making processes. To put it in another way, we are interested in knowing how evaluation and its results have been used by decision-makers. This can be considered a crucial question for the legitimacy of evaluation even in a broader sense: the rationale for evaluation as an intentional, goal-oriented activity is constituted by its capability to provide insights and information which is accurate and can be utilised in the development and planning of activities. The issue will be studied in connection with the evaluation of publicly funded research and development activities in Finland.

The potential contribution of evaluation to science and technology policy with regard to decision-making will be approached from the decision-makers' point of view. In order to get an idea of the potential uses of evaluation results, 40 decision-makers at three different levels of the national R&D system were interviewed:

- i. Policy-level - current/former members of the Science and Technology Policy Council of Finland, and officials involved in R&D related decision-making processes in five ministries.
- ii. Intermediary-level - higher management of research funding agencies, i.e. the Academy of Finland and the National Technology Agency, Tekes.
- iii. Operational level - higher management of selected government research institutes, and rectors of some major universities.

The following chapter will lay out the short history of R&D evaluation activities in Finland. The chapter will cover the development of the Finnish R&D evaluation scene from the beginning of the 1980's to the end of the 1990's. Forerunners and main players in the field will be identified and changes in the number of evaluations will be addressed. After presenting the background and context of the study, I will proceed to examine the use of evaluations in decision-making in more detail. The final chapter will raise issues which, according to the interviewed decision-makers, should be taken into account in the development of future evaluation practices.

## Diffusion of the R&D Evaluation Practice in Finland

The historical development of research evaluation activities in Finland has followed a path familiar from other industrial countries. Kuhlmann has discerned two heterogeneous development lines, which have left their mark on the present R&D evaluation practice in various OECD countries.<sup>1</sup> The historical core of evaluation activities is composed of peer review as an internal tool of the research community, which supports decision-making concerning the allocation of resources for research areas and disciplines. This activity has traditionally focused on research proposals laid out by individual researchers and research groups. Nowadays, an increasing share of peer review activities is linked with the evaluation of research projects in a competitive funding context. Peer review is also exploited more and more in evaluations which address the quality and performance of research programmes or groups in retrospect.

The core of peer review is encircled by a second group of evaluation studies, which Kuhlmann calls impact analyses of research and technology programmes. This approach is manifestly external vis-à-vis the research community. The launching of impact analyses is based particularly on the concerns of actors in the politico-administrative system: they are interested in finding out if and to what extent scientific, technological, economical or wider societal goals set up in political decision-making are achieved through publicly funded measures. The "outer shell" of evaluation activities consists of the evaluation of public institutions in which R&D takes place, i.e. government research institutes and universities. There are no independent approaches or methods for institutional evaluation; they are derived from models of peer review and impact analyses.

---

<sup>1</sup> Stefan Kuhlmann (1998) Politikmoderation. Evaluationsverfahren in der Forschungs- und Technologiepolitik. Nomos Verlagsgesellschaft: Baden-Baden

The model sketched by Kuhlmann applies well to the development in Finland, where the science and technology policy field has witnessed an emergence of a vital evaluation culture during the last two decades. The first efforts induced by international examples and internal considerations were carried out at the Academy of Finland and the Technical Research Centre of Finland, VTT, at the beginning of the 1980's - 1983 to be exact. The Academy of Finland was primarily interested in carrying out evaluations on scientific fields, whereas the evaluations at the VTT focused on the institute's research programmes. Some years later, the recently founded National Technology Agency of Finland, Tekes - which was at the time called the Technology Development Centre of Finland - adopted the evaluation model into its technology programme procedures.

The evolution and diffusion of the evaluation culture took place gradually without perceivable outside pressure: evaluation was perceived within administration as a management innovation and it was adopted by different organisations. After 17 years of development, evaluation is visibly anchored in the Finnish research and development system. Today, evaluation is included almost on a routine basis in numerous official documents dealing with strategy and development of the national innovation system. For example, in the science policy strategy of the Academy of Finland for 1998-2000 it was noted that "evaluating research and developing methods are among the Academy's main functions". This task was perceived to include a broad palette of subjects from general evaluation work and evaluation of scientific disciplines and research programmes to developing research indicators and the evaluation of funding.

During the 1990's the Science and Technology Policy Council of Finland, a body assisting the Government and its ministries in questions relating to science and technology, played an important role in nurturing the evaluation practice into the Finnish R&D scene. The Council actively promoted the diffusion of evaluation activities into different areas of the country's public R&D system. As part of the

development of sectoral research, the Council recommended in 1995 that all government research institutes should be evaluated by the end of 1999. This was included in the official action programme for the development of sectoral research, which was approved by the Council in 1996. Around the same time, a recommendation to evaluate all Finnish universities by the end of the decade was also made. These two proposals have together played a significant role in diffusing and establishing the institutional evaluation practice on Finnish soil. The evaluations of government research institutes are officially commissioned by the sectoral ministries, and university evaluations by the Ministry of Education. In this context, a new body under the administration of the Ministry of Education was also founded, namely the Higher Education Evaluation Council, whose main task is to assist and support universities and newly founded polytechnics in evaluation arrangements.

The diffusion of the evaluation practice among actors of Finnish R&D system has been profound: since the beginning of the 1980's the number of evaluations carried out has increased almost exponentially every fifth year - a raising curve is clearly visible. At the end of 1989, 25 evaluations were carried out; ten years later the number reached 175. Organisations whose core activities are situated at the technical end of the science continuum, namely Tekes and the VTT, have been especially keen to adopt the evaluation practice into their routines. These two organisations together account for a substantial share, around 50 per cent of the all R&D evaluations carried out in Finland. The Academy of Finland, which was previously an active commissioner of research evaluations, changed its policy in this respect in the middle of the 1990's: the evaluation of research fields and programmes was discontinued at least temporarily, and resources were instead concentrated on a broader review of the state and quality of scientific research in Finland. This assessment is scheduled to be presented once during the three year terms of the Research Councils. From Figure 1, it becomes clear that institutional evaluations increased very rapidly after the early nineties - thanks to the recommendations to evaluate both the governmental research institutes and the universities before the turn of the century. All institutional evaluations are formally commissioned by the

ministries responsible for their respective administrative domains, as mentioned above.

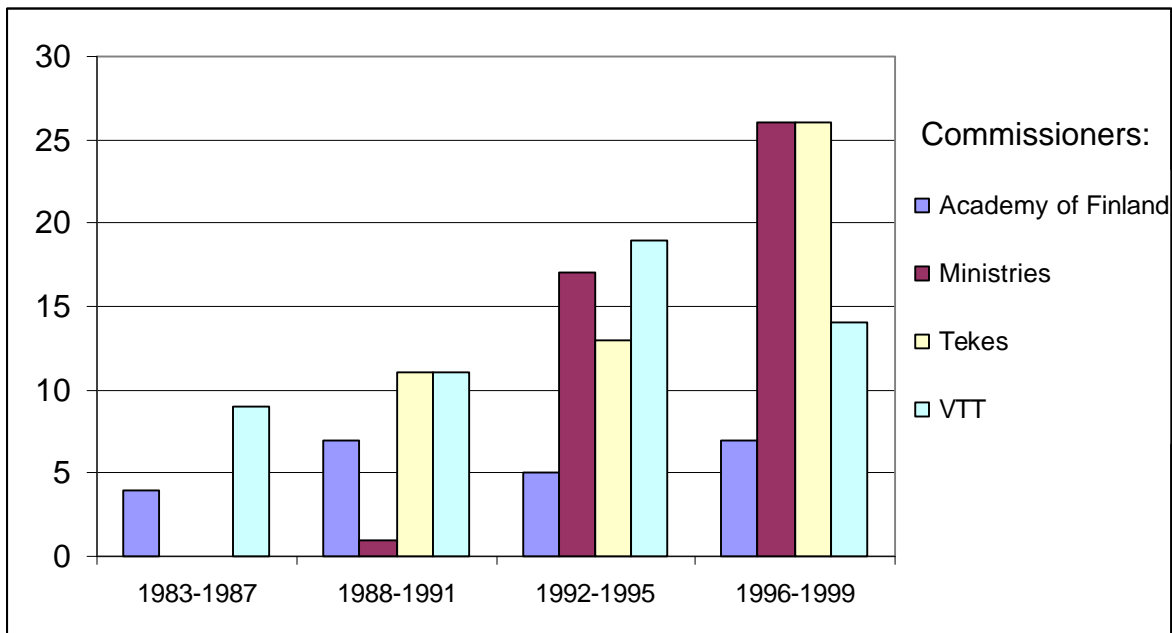


Figure 1. Number of R&D evaluations carried out in Finland between 1983-1999.

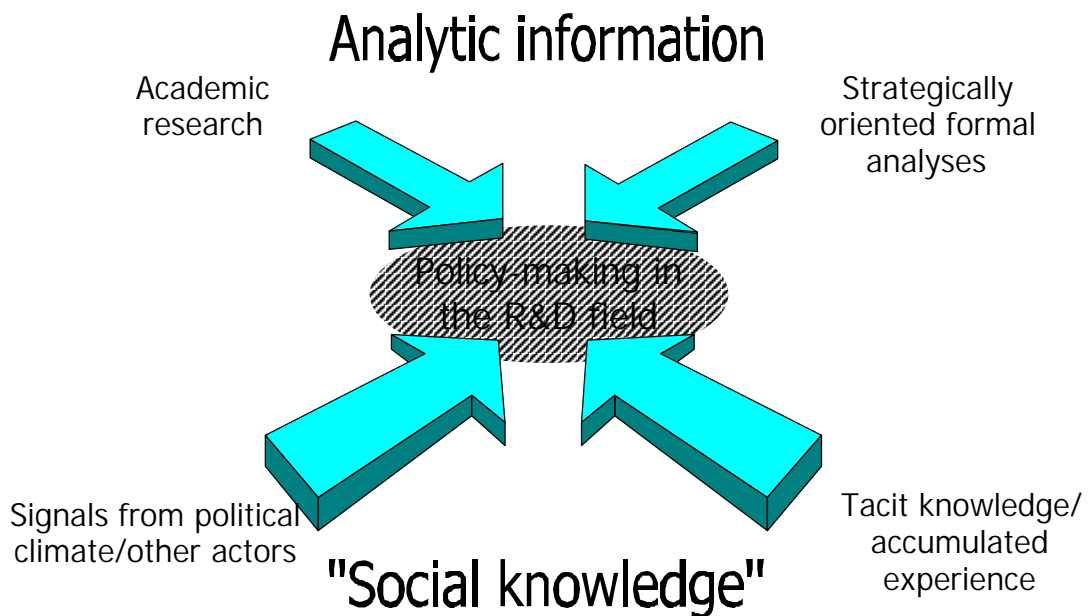
## Does evaluation contribute to decision-making?

The general attitude towards the evaluation of R&D activities among the interviewed decision-makers was both positive and pragmatic. Evaluation was seen to include many elements which make it useful in decision-making. It was reminded, however, that evaluation is not a magic cure for all situations; rather, the launch of an evaluation has to be well-planned and it must be based on clear goals.

The whole question of the utilisation of evaluation information turned out to be intricate for the decision-makers. They found it often difficult, if not impossible altogether, to discern afterwards to way in which certain evaluation effort had contributed to decision-making in a particular situation. Evaluation information was seen to be closely intertwined with other relevant sources at the decision-makers' disposal. The figure below presents the different information sources that were

deemed as important by the interviewees. The emphasis and significance of the different sources varies not only depending on the situation, but also on individual working methods and organisational cultures. Some decision-makers prefer to read documents and reports whereas others find it more natural to gather background information through face-to-face communication.

## Information Inputs into Decision-Making Process



The opinions concerning the usability of evaluation diverged somewhat depending on organisational affiliation and, of course, on the interviewees' familiarity with the issue. It seems that the most positive attitudes are found among people from ministries. The middle ground was occupied by university rectors and directors of government research institutes. In the funding organisations, thoughts on evaluation were varied: evaluation was understood to be an integral measure in the work of the R&D community, but at the same time there were some concerns about its present capability to support strategy work.

These slight differences can be explained by people's varied experience with different kinds of evaluations. The practical contribution of an institutional evaluation is often easier to detect compared to the potential effects that an R&D programme evaluation can have on the evaluated activity. People at ministries and research institutes have witnessed that evaluation of a research institute has, for example, led to extensive organisational reforms. Whereas in the context of R&D programme evaluations, the potential contribution of evaluation is often more indirect and takes place within a longer time span; for example, the effects may surface, if at all, after many years in the planning of new programmes, and then it will be much more difficult to discern their origin.

## Evaluation as a management tool

The majority of the interviewees considered evaluation to be a management tool which can be used in many different ways. A recurring view associated evaluation with cost-awareness both at the level of an individual research organisation/programme, and at the level of the national R&D-system as a whole. Many comments emphasised that the public financier, and finally the taxpayer, has the right to know whether public resources are used efficiently and for relevant purposes. This emphasis on accountability might be interpreted as one dimension of the public sector reform which was made during the past two decades across the OECD countries. The evaluative attitude is an integral part of this so-called "new public management", and management by result is probably its most obvious manifestation.

However, evaluation was not only seen as a tool of ensuring the accountability and relevance of public R&D activities, but also as a procedure closely linked with internal development and learning. According to many interviewees, the evaluation process had offered the staff an opportunity to pause for a moment and to reconsider what they are doing, why they are doing it, and how they are doing it. This way they have a chance to identify weak points in the functioning of their organisation and

ambiguities in their goals. This kind of an internal learning process has led to new opportunities and started reforms even before the completion of external evaluation process. The learning aspect in the evaluation process was deemed particularly important among university rectors and directors of government research institutes: some of them even noticed that, from the organisation's standpoint, the internal evaluation preceding external evaluation was the more productive and useful part of the evaluation process.

A somewhat more critical viewpoint concerning the internal learning aspect was brought up in some interviews, too: the interviewees referred to the possibility that the current "evaluation boom" in Finland can also be understood as "governing from a distance". Administration induced evaluation activities and various indicators used in the R&D field affect the self-understanding of research organisations and research teams by directing attention to certain issues, whereas other potentially relevant questions are left aside. In this way, evaluation may trigger a desired self-steering process in the organisation or team.

Apart from the opportunity for internal learning, evaluation helps to clarify the mutual expectations of partners and other interested parties. Thus, an evaluation of an individual institution can turn out to be a learning process for a wider institutional environment, including agencies higher in the hierarchy, such as ministries and other partners. For example, an evaluation of a research institute has offered the ministry concerned an opportunity to review the institute more thoroughly; its functioning, its position within the administrative field, and its role in the development of the sector in question. In addition, some interviewees at the ministry level felt that institutional evaluations have succeeded in pointing out weaknesses which would demand internal changes even within ministries.

Recent evaluations of the government research institutes have increased connections between ministries and subordinate research institutes. The evaluation has given ministries a legitimate means to get the institutes' attention regarding issues deemed important from the ministry's point of view. At the same time,

however, horizontal co-operation across administrative sectors has not developed to the same extent. This last point has raised some concern at the Science and Technology Policy Council. According to the Council, it is clear that evaluations have proved to be beneficial for the research institutes, and for the inter-relationships between the evaluated institutes and the ministries concerned. The Council is, however, concerned that evaluations do not answer - not even when reviewed together - to the questions such as how well the arrangements concerning research activity compare to the overall needs of sectoral research in Finland.

## Evaluation as persuasion

It is clear that the evaluation of publicly funded research and development activities also involves a strong political dimension - when politics is understood as an effort to convince other people about the soundness and necessity of certain plans and decisions. This dimension of the evaluation activity is in itself a "delicate" issue, as one of the interviewees put it.

Justifying and convincing takes place at different levels. For the organisation's management, evaluation offers general support and assurance in decision-making situations. The support that evaluation gives for them is clearly valued among decision-makers. Evaluation is seen as an important external "second opinion", in relation to which decision-makers can reflect their own ideas. This "second opinion" gives management the opportunity to assure itself that planned actions and strategic choices are also reasonable from an external point of view.

In addition to the "second opinion" effect, evaluation is valuable for justifying and convincing others about the necessity of proposed decisions. Within the organisation, evaluation provides support to the decision-makers in relation to the staff: decision-makers do not have to stand alone when they can point out that someone from the outside, who has 'objectively' assessed the situation, has also come to the conclusion that certain kinds of decisions are inevitable. Evaluation is also understood to be a

handy tool for showing the relevance of activities for actors with whom the organisation has close links, such as ministries, funding agencies and other stakeholders. If, however, evaluative conclusions turn out to be negative, the evaluation process itself can function as an indicator of the organisation's active willingness to improve its performance.

This aspect of justifying and convincing is particularly important in a time of difficult decisions when evaluation results may be used to legitimise decisions and activities which would otherwise be difficult to carry out; these include organisational reforms and the reallocation of resources, or the prioritisation of R&D areas. This came out clearly in most of the interviews. For instance, from the ministries' point of view, the evaluation of research institutes has had a clear pull-the-trigger effect. It was felt that it had been easier to launch larger organisational restructuring processes or a notable readjustment of strategies through external evaluation. Some interviewees doubted that without evaluation, the ministry in question would not have had enough power to get the research institute to carry out reforms that were deemed as necessary, or reforms would have been postponed.

Overall, conclusions and recommendations in an evaluation report can be utilised in many ways, and by different actors. The same conclusions and recommendations - especially if they are positive - can be used as arguments for additional resources and other demands, too. Evaluation results have also been used to point out the relevance of the work for a broader public. The strategic use of evaluation information seems to be a part of business, so to say. This is understandable because after publication the evaluation report, like any other published document, is open to different interpretations. Individuals and groups who feel that they have been 'mistreated' in the evaluation and presented in the wrong light may try to downplay the relevance of the whole endeavour. On the other hand, people who think that the evaluation presented them in a favourable light will try to gain from it.

## Strengths

Overall, in light of the interviews, the value of evaluation as a management tool does not lie in its special capability to identify and raise totally new issues. Rather, it is more the rule that the evaluated institution and its environment are to some extent aware of the questions that are brought up by the evaluation group. The reason why the evaluation is still regarded as worth carrying out is based on the idea that the evaluation will succeed in highlighting and emphasising issues and aspects which are easily forgotten or lost in everyday business. One of the interviewees expressed this by noting that "external opinion is always welcome, because it is human that people tend to become blind to their daily routines; then someone who comes from outside the organisation sees things that the staff may not notice". Evaluation also constructs a more coherent picture about the situation and draws attention to certain questions; it clarifies and gives more weight to them.

## Weaknesses

The interviewees also brought up some weaknesses in current evaluation practices, which may diminish the usability or trustworthiness of the conclusions and recommendations. The most often mentioned weak points were related to following two questions: (1) lack of time for evaluations which may lead to inaccuracies and superficiality in results; (2) evaluators' competence and in particular, their insufficient knowledge of local circumstances.

Lack of time does not allow evaluators to familiarise themselves thoroughly with the object of the evaluation. As a result, conclusions and recommendations may be biased, even inaccurate. Reasons for the lack of time are related both to financial issues and full-booked schedules of the evaluators. Reputed and well-known evaluators are involved in many different activities at the same time, and they only have limited time to concentrate on an individual evaluation case. One of the

interviewees noted that "we do not have any means of securing that the evaluators go through all the documents that were delivered to them in advance".

The problems caused by the lack of time was also commented on by the evaluators of the Finnish Institute of Occupational Health in a follow-up report: "The group was not satisfied with the amount of time that it had to do the evaluation".<sup>2</sup> They identified a number of reasons why the time reserved for the evaluation was limited: the time that the evaluators could give to the review; the increased costs of a more extensive evaluation; the timetable of the commissioner, and the need for answers in a reasonable time period. On the other hand, the visible lack of time for the evaluation was not considered to be such a serious issue by everyone: commissioners know the limitations, and thus they can be taken into account when the conclusions and recommendations are later reviewed.

Another weakness pointed out by the interviewees was related to role of international evaluators. The contribution of international experts was highly valued by many decision-makers. At the same time, however, potential shortcomings related to this common practice were mentioned surprisingly often, bearing in mind that the use of international evaluators is well established in Finland today. There is a good reason for this practice, namely the need to ensure the neutrality of an evaluation, which would otherwise be difficult because of the limited pool of native experts without stakes in a particular evaluation. International experts are also seen to offer a valuable outside perspective on national R&D activities; evaluators from abroad may identify strengths and weaknesses which could go unnoticed by a native. Experienced and well-reputed evaluators are also in a position to compare local R&D arrangements to international standards and best practices.

The preceding comments were widely shared by the interviewees. However, they also saw drawbacks in the use of international evaluators. The interviewees maintained that a lack of knowledge of local conditions had several times led the evaluators to draw inaccurate conclusions and recommend measures that could not

be applied in the Finnish context. These kinds of shortcomings have adverse impacts on the usefulness of evaluation information in decision-making. It is also noteworthy that still today, the significant part of official documentation about research activities and the publication of research results are in Finnish. This natural barrier is hard to overcome - even if evaluated research groups and organisations, as well as commissioners use time and human resources to translate the principal documents into English. Evaluations usually try to resolve this problem by ensuring that at least one Finnish member is appointed to the evaluation team.

### Decision-makers' views on the future of research evaluation in Finland

In general, the continuation of research evaluation was seen as indispensable part of managing the national R&D system at its different levels. At the same time, however, the interviewed decision-makers widely concurred that excessive evaluation of R&D activities should be avoided. This applied especially to institute-wide evaluations, which should not be conducted too often, because of the intensive documentation needed to prepare for the evaluation and the large amount of staff time that the whole process takes. As a one interviewee remarked "there should also be time to concentrate on core business". The majority of the interviewees estimated that a suitable period for broader institutional evaluation efforts is 5 to 10 years. Single R&D programmes and parts of research areas can be evaluated more often.

Interviewees also brought up another aspect related to the improvement of evaluation practices so that they would be relevant in decision-making situations. This aspect concerns the development of measures and tools that combine elements and perspectives from future-oriented strategic studies to more retrospectively-oriented evaluation activities. For instance, ex-post evaluations of programmes produce valuable information about the performance of past policies and decisions, but they are not informative enough for actors such as funding organisations, which

---

<sup>2</sup> Investment in Health. The Follow-up Report of the International Evaluation of the Finnish Institute of Occupational Health. Ministry of Social Affairs and Health Finland. Publications 1998:2

are continuously scanning the terrain of R&D activities in order to identify signs on which to base their resource allocation decisions.

The third point regarding the improvement of evaluation practices deals with the need to assure a smooth flow of information between evaluators, decision-makers and other interested parties from the beginning of the evaluation process. From the decision-makers' standpoint, the final evaluation report is often an insufficient medium to ensure that the insights and information produced in the evaluation are put into use in decision-making. Instead, many interviewees emphasised the significance of continuous interaction during the evaluation process as a means to link evaluation more closely with decision-making. They also mentioned that there was a need to connect the design of evaluation already to the planning phase of a programme or policy to be evaluated. This would enable more systematic data gathering and better co-ordination between evaluation effort and programme management.

To conclude, there is also need to ensure that evaluators have at least some basic knowledge of the local context in which the subject of the evaluation operates. This becomes even more important in the future, if R&D evaluation continues to emphasise the social relevance of publicly funded research and development activities.