Publications of the Academy of Finland 9/06

Methods for Evaluating the Impact of Basic Research Funding

An Analysis of Recent International Evaluation Activity

Sami Kanninen – Tarmo Lemola

Academy of Finland
Research funding and expertise

Publications of the Academy of Finland 9/06

Methods for Evaluating the Impact of Basic Research Funding

An Analysis of Recent International Evaluation Activity

Sami Kanninen – Tarmo Lemola

Academy of Finland in brief

The Academy's mission is to finance high-quality scientific research, act as a science and science policy expert and strengthen the position of science and research. The Academy's activities cover all scientific disciplines.

The main focus of the Academy's development activities is on improving opportunities for professional careers in research, providing resources and facilities for high-profile research environments and making the best possible use of international opportunities in all fields of research, research funding, and science policy.

The Academy has a number of funding instruments for various purposes. In its research funding, the Academy of Finland promotes gender equality and encourages in particular women researchers to apply for research posts and research grants from the Academy.

The Academy's annual research funding amounts to more than 240 million euros, which represents some 15 per cent of the government's total R&D spending.

Each year Academy-funded projects account for some 3,000 researcher-years at universities and research institutes.

The wide range of high-level basic research funded by the Academy generates new knowledge and new experts. The Academy of Finland operates within the administrative sector of the Ministry of Education and receives its funding through the state budget.

For more information on the Academy of Finland, go to www.aka.fi/eng.

Taitto PixPoint ky
ISBN 951-715-593-X (print)
ISBN 951-715-594-8 (pdf)
Paino: Edita Prima Oy, Helsinki 2006

Kuvailulehti

Julkaisija	Suomen Akatemia	Päivämäärä	18.5.2006
Tekijä(t)	Sami Kanninen – Tarmo Lemola		
Julkaisun nimi	Methods for Evaluating the Impact of Basic Research Funding: An Analysis of Recent International Evaluation Activity		
Tiivistelmä	Raportti on osa Suomen Akatemian käynnistämää Sight 2006 -hankekokonaisuutta, jonka tarkoituksena on tukea Suomen tutkimusjärjestelmän ja Akatemian vaikuttavuusarvioinnin kehittämistä. Raportti esittelee tulokset tutkimuksesta, jonka tavoitteena on tuottaa tietoa Akatemian kaltaisissa organisaatioissa hyödyn-netyistä vaikuttavuusarvioinnin menetelmistä. Tutkimuksen tarkoituksena on tuottaa tietoa menetelmistä ja indikaattoreista, joiden avulla perustutkimuksen sekä perustutkimuksen rahoituksen vaikutuksia ja vaikuttavuutta on arvioitu. Tutkimuksessa analysoidaan kahtatoista ulkomaisissa rahoitusorganisaatioissa toteutettua arviointihanketta. Analyysien tulosten pohjalta raportti esittää suosituksia Akatemian arviointitoiminnan kehittämiseksi kol-mella eri tasoilla: kansallisella tasolla, Akatemian toiminnan tasolla, sekä rahoitusinstrumenttien tasolla. Kansallisella tasolla arviointitoiminnan painopistettä tulisi suunnata aiempaa enemmän suomalaisen tutkimusjärjestelmän kehittämistarpeiden tunnistamiseen. Tällainen lähesty-mistapa tarjoaisi enemmän tietoa Akatemian tutkimusrahoituksen kohdistamisen tueksi. Akatemian tasolla arviointitoimintaan vaikuttavat tarpeet kehittää indikaattoreita, jotka mittaavat Akatemian toiminnan tuloksellisuutta ja tehokkuutta suhteessa toiminnalle asetettuihin tulostavoitteisiin. Kehittä-mistyössä tulisi painottaa indikaattorien validiteettia, relevanssia ja stabiliteettia. Instrumenttitasolla Akatemian arviointitoiminnan pääpaino tulisi olla laadullisessa vaikuttavuusarvioinnissa. Arviointitoiminnan kohteina olisivat sekä tutkimuksen vaikutukset että ne mekanismit, jotka edistävät vaikutusten syntyä.		
Asiasanat	arviointi, vaikuttavuusarviointi, vaikuttavuus, indikaattorit, perustutkimus, tutkimusrahoitus		
Julkaisusarjan nimi ja numero	Suomen Akatemian julkaisuja 9/06		
ISSN	0358-9153		
ISBN	Painetulle kirjalle annettu tunnus 951-715-593-X	Pdf-versiolle annettu tunnus	951-715-594-8
Sivumäärä	98		
Julkaisun jakaja	Suomen Akatemia, PL 99, 00501 Helsinki, viestinta@aka.fi		
Julkaisun kustantaja	Suomen Akatemia		
Painopaikka ja -aika	Edita Prima Oy, Helsinki 2006		
Muut tiedot	www.aka.fi/julkaisut		

Description

Publisher	Academy of Finland	Date 18 May, 2006
Author(s)	Sami Kanninen – Tarmo Lemola	
Title	Methods for Evaluating the Impact of Basic Research Funding: An Analysis of Recent International Evaluation Activity	
Abstract	The report is part of the Sight 2006 project initiated by the Academy of Finland the aim of which is to support the development of the Finnish research system and the impact assessment of the Academy. The report presents the results of a study carried out to generate information on the approaches of impact assessments adopted internationally in organisations similar to the Academy. The purpose of the study is to provide information on the methods and indicators with which the impacts of basic research, as well as the impacts of the operations of a research funding organisation, have been assessed. The study analyses twelve evaluation projects carried out in foreign funding organisations. Based on the results of these analyses, the report provides recommendations for developing impact assessment activities at the Academy of Finland on three different levels, i.e. at the national, Academy operations and instrument levels. At the national level, the accent of the evaluation activity should be shifted towards identification of development needs in the research system. This would provide more information for steering the Academy's funding allocation towards areas with more potential for impacts in relation to observed societal needs. At the level of Academy operations, the evaluation activity is influenced by the need for developing indicators that measure the Academy's productivity and performance in relation to its target outcomes. The development process should emphasise the validity, relevance and stability of indicators. Atthe instrument level, the main emphasis of the Academy's evaluation activity should be on the qualitative assessment of research impacts and the mechanisms that promote the generation of these impacts.	
Key words	evaluation, impact assessment, impacts, indicators, basic research, research funding	
Name and number of series	Publication of the Academy of Finland	
ISSN	0358-9153	
ISBN	Print 951-715-593-X	Pdf 951-715-594-8
Number of pages	98	
Distributed by	Academy of Finland, POB 99, FI-00501 Helsinki, viestinta@aka.fi	
Published by	Academy of Finland	
Place and date of printing	Edita Prima Oy, Helsinki 2006	
Other information	www.aka.fi/publications	

Contents

Foreword

The results and impacts of science and technology policy as well as research funding are intensively discussed in all developed countries. Knowledge-based strategies and increased funds invested in S&T emphasize the importance of impact evaluation. Societal impact is – according to the new University Law – the third mission of universities from September 2005 onwards also in Finland.

The Finnish Government finds impact evaluation as an important instrument of science and technology policy in its Resolution on the Structural Development of the Public Research System from 7 April 2005. Actually, the Academy of Finland and Tekes (Finnish Funding Agency for Technology and Innovation) will review the impacts of the Resolution as part of their general mission of developing impact evaluation in the Finnish system.

This report is part of a development and evaluation programme of the Academy of Finland called Sight 2006. The main objective of Sight 2006 is to review the state, structure and quality of Finnish scientific research. It includes bibliometric studies on the structure and quality of Finnish research in international comparison, studies on the impact of funding by the Academy's research councils, external evaluation of the impact of research funding of the Academy and development of indicators on the level and changes of Finnish expertise (in cooperation with Tekes). Its most extensive project is the first Finnish S&T foresight FinnSight 2015, which will study the future challenges of science and technology in the Finnish socio-economic context.

This study has been performed by Advansis Oy as a commissioned work. The study provides information on the approaches of impact evaluation that have been adopted internationally in organisations similar to the Academy. The purpose of the study has been to generate information on the methods and indicators with which the impacts of basic research, as well as the impacts of the operations of research funding organisations, have been assessed.

I convey my sincere thanks to Mr Sami Kanninen and Mr Tarmo Lemola for the strict and responsible ways of working and for flows of information, ideas and knowledge, which they have transmitted to us in numerous discussions during the work. My thanks also go to my colleague Ms Annamaija Lehvo who in many ways, with her good evaluation expertise, contributed to the formation of the study.

I hope that the study will be helpful for all those active in impact evaluation, in Finland and in other countries.

Paavo Löppönen
Director
Development and Evaluation

Executive Summary

Objectives and Methodological Approach of the Study

The study is part of the Sight 2006 project initiated by the Academy of Finland to inform strategic decision-making regarding the development of the Finnish research system. The study responds to the need for developing ex-post assessment of impacts of basic research funding in the Finnish context. Other themes addressed by Sight 2006 include the impact, challenges and opportunities associated with the structural development of the Finnish research system as well as the state of scientific research in Finland assessed by bibliometric studies.

The study provides information on the approaches of impact evaluation adopted internationally in organisations similar to the Academy. The purpose of the study is to generate information on the methods and indicators with which the impacts of basic research, as well as the impacts of the operations of a research funding organisation, have been assessed. The objectives of the study are:

1. To identify and select relevant evaluation projects for more detailed analysis based on a review of international evaluation activity;
2. To describe the evaluation practices and approaches employed by the selected projects;
3. To identify and compare the evaluation approaches as regards their advantages, disadvantages and areas of application;
4. To present suggestions regarding the use of the evaluation methods as well as the development of evaluation practices at the Academy of Finland.

The evaluation projects analysed in the study are:

1. Evaluation FWF: Impact Analysis, Joanneum Research, 2004;
2. The Impact of the EU Framework Programmes in the UK, Technopolis Ltd, 2004;
3. Funding Ranking 2003: Institutions – Regions – Networks, Deutsche Forschungsgemeinschaft, 2003;
4. A Wealth of Knowledge: The return on investment from ARC-funded research, The Allen Consulting Group, 2003;
5. Methods for Assessing the Economic Impacts of Government R&D, Gregory Tassey, National Institute of Standards and Technology, 2003;
6. Review of Wellcome Trust PhD Training: Career paths of a 1988–1990 Prize Student cohort, The Wellcome Trust, 2000;
7. Assessing research impact on non-academic audiences. Final report submitted to the Research Evaluation Committee, Economic and Social Research Council, SPRU, 1999;
8. A Composite Performance Rating System for ATP-Funded Completed Projects, Rosalie Ruegg, TIA Consulting, 2001;
9. Outcome Indicators 2003/04, Foundation for Research, Science and Technology, 2004;
10. Standard Evaluation Protocol 2003–2009 For Public Research Organisations, Vereniging van Universiteiten (VSNU), Netherlands Organisation for Scientific Research (NWO), Royal Netherlands Academy of Arts and

Sciences (KNAW), 2003;
11. PSA target metrics for the UK research base, Office of Science and Technology, 2004;
12. Taking Stock of R&D Across Three Sectors. Taking Stock of University & College Commercialization Efforts: Advising on The Way Forward. A Study for The Advisory Council on Science and Technology, Chris Riddle, 2003.

The selected evaluation projects can be roughly categorised into three groups: (a) individual projects that attempt to measure research impact (5 projects), (b) individual projects that attempt to review or develop evaluation methods (3 projects) and (c) reports that either describe a system of evaluation or provide an assessment of research performance at a national level with a set of indicators (4 projects).

Analysis of the Approaches Employed by the Evaluation Projects

The approaches with which the selected evaluation projects characterise the impacts of a funding organisation can be grouped into four conceptual themes. The themes reflect the stages in the process through which the impacts from research funding are realised: 1) allocation of funding, 2) research outputs, outcomes and impacts in the scientific community, 3) collaboration in research, and 4) exploitation of research and resulting impacts. The stages emphasise the conceptual distinction between the impacts of research on the scientific community and the impacts on external stakeholder groups.

With regard to *allocation of funding*, the evaluation projects characterise impacts from the operations of research funding organisations by assessing the inputs to the research base at the national level, by analysing the significance of the given type of funding for the applicants, by characterising the funding recipients in contrast to unsuccessful applicants, and by evaluating the additionality of the funding. The benefits of employing analyses of funding patterns as a method for assessing and communicating impacts from the activities of research funding organisations include the availability of relatively reliable, quantitative data that often enables international comparisons. Moreover, the information provided by the analyses of funding patterns is not diluted by external factors and thus provides a relatively accurate picture of the influences of the operations of a funding organisation. A disadvantage of the approach is that the link between research funding and the societal impacts accruing from research efforts can merely be assumed.

The evaluation of *direct outputs and outcomes* of research efforts has primary been addressed in the analysed material by assessing three issues. First, the assessments have focused on scientific outputs, by simply calculating the number of produced publications and citations, as well as by identifying factors that influence the output volume. Second, the outputs targeted to research users outside the scientific community have been assessed by calculating the numbers of publications, presentations and services provided for users. In the assessment of societal impacts at an aggregated level this approach has an important role: The numbers of user outputs serve as a proxy measure for research utilisation, but the data can be gathered from the recipients of the funding, which considerably simplifies the exercise. Third, the evaluations have

analysed the impacts of research efforts on the research performers. The main focus of the analysis is on researchers' increased opportunities brought about by research funding.

Since impacts on research users are complex to identify and quantify, *collaborations* are often used as a proxy measure for impacts. Furthermore, collaborative linkages are also viewed as an important impact of research funding in their own right, because they are assumed to have a positive impact on the performance of the innovation system in general. The analyses of impacts from basic research funding can be grouped into three categories. First, the evaluations have addressed the extent of scientific collaborations in research programmes. Second, the analysed evaluation projects exemplify methods for assessing international cooperation among research organisations. Third, the analyses have focused on knowledge transfer between the scientific community and the research users. The methods employed range from simple counting of collaborative contracts to visualisations of network structure and qualitative investigations of the formation of cooperative relationships.

Finally, the selected evaluation projects have employed analyses focusing directly on *the impacts on research users*. The approaches include assessments of the relevance of research for the users, indicators of the commercial exploitation of research results, investigations of the changes in the user organisations, as well as attempts to assess the impacts of research efforts on society at large. The assessment of changes in the activities of the research users and the evaluation of resulting aggregated impacts have proved to be the most challenging of the employed approaches. The impacts on research users are diverse and elude quantification. On the other hand, with qualitative methods it is difficult to achieve comprehensive coverage of the activities of a funding organisation. Evaluations focusing on narrowly defined research areas appear to be more conducive for successful assessments of societal impacts accruing from research funding.

Conclusions and Recommendations

The impact evaluation activity at the Academy of Finland can be conceptually structured by identifying three different levels as the objects of the evaluation activity: evaluation of the impacts of science at the national level, the impacts from the operations of the Academy, and the impacts from the instruments that the Academy employs in order to respond to societal needs. As a whole, the evaluation activity should be viewed as a system that provides strategic information and supports the long-term development of the organisation.

When the international evaluation projects are analysed from the perspective of these levels, thoughts and ideas regarding the development of the Academy's evaluation activity can be presented. The emerging ideas address both general directions towards which evaluation activity at the Academy of Finland could be steered as well as provide concrete suggestions regarding evaluation projects that could be initiated.

At the national level, the indicators employed by the review of the state and quality could be complemented with indicators of research commercialisation in universities and research institutes to reflect that aspect of research utilisation. As a whole, the accent of the evaluation activity should be shifted towards identification of development needs in the research system. This would provide

more information for steering the Academy's funding allocation towards areas with more potential for impacts in relation to observed societal needs.

At the level of Academy operations, the evaluation activity is influenced by the need for developing indicators that measure the Academy's productivity and performance in relation to its target outcomes. The development process should emphasise validity, relevance and stability of indicators. This implies that the focus of the development process should be on investigating the significance of the selected indicators. Particular attention should be paid on the ability of the outcome indicators to reflect utilisation of research and societal impacts more broadly. In order to complement the information provided by the selected primary indicators, a more comprehensive assessment of the Academy's outcomes can be conducted based on the projects' final reports.

At the instrument level, the main emphasis of the Academy's evaluation activity should be on the qualitative assessment of research impacts and the mechanisms that enhance the generation of these impacts. In research programme evaluations, impact assessment should focus on factors and conditions that influence the utilisation of research by users. As for centres of excellence, the impact evaluations should create data for assessing how the centres have achieved their goal of catalysing vibrant research environments. The evaluations of personal funding schemes, in turn, should investigate how the research funded by these instruments has contributed in the long term through the skills and knowledge acquired by the researchers to subsequent research efforts both in research organisations and in business enterprises.

1 Introduction

Fundamental scientific research is deeply embedded in society. There is an ongoing interplay between science and the political, techno-economic and cultural processes of society. The relationship between these is dual and dynamic: the societal processes both influence the directions that scientific research takes and are affected by the knowledge produced through scientific progress.

The results from basic research generate a broad range of influences to society. At the one end of the spectrum, research generates information for relatively well-specified aims and research problems. At the other end, research contributes to the general ways in which we view the world by providing new understanding of the phenomena around us. The information and theoretical models produced by scientific research permeate our thinking on many different levels.

Science being an essential part of the modern society, it is important to assess the many influences and impacts that it generates. For organisations whose primary mission is to serve society by providing better conditions for the conduct of science, the assessment of impacts from research plays a particularly central role: impact assessment can function as an instrument for creating information for strategic decision-making. Evidence of research impacts can provide a basis for valueing past actions as well as for informing of future decisions.

Due to the multiplicity of influences from science, combined with the various purposes for impact assessment, different types of evaluation methods are required. As new, societal needs for information on the impacts of basic research are created, the application of existing evaluation methods may not be appropriate and, thus, methodological development is needed. Similarly to the dynamic linkage between society and science in general, also impact assessment and societal processes are influenced by each other.

1.1 Context of the Study

During the recent years, the needs for the assessment of impacts from basic research have increased. Globalisation and increased competition among nations have created pressure for targeting the national inputs for basic research effectively and efficiently. The cost of supporting modern scientific infrastructures has increased, as basic research requires ever more capable and powerful instruments. As a result of the findings in modern innovation research, greater significance has been given to the impacts of basic research as a catalyst for economic progress. Basic research – and funding allocated for it – is seen as an integral part of a broader innovation system, in which basic research is linked to applied research, product development and culture in general, creating a complex web of influences. Together these trends have led to an increased need to assess the impacts from funding allocated for basic research (Arnold & Balázs 1998).

In Finland this pressure is apparent in the current policy discussion, which emphasises the importance of improving the productivity and effectiveness of conducted research activities. The current development efforts in the Finnish research system are marked by an effort to increase the societal benefits from research. More efficient utilisation of re-

search results has a key role in achieving this objective. The government resolution on the structural development of the public research system states (Government resolution 7 April 2005, p. 1):

"*The national objective for Finland is a sustainable and balanced societal and economic development. High employment, productivity and competitiveness are key factors in it. Focused measures to step up research and technological development and the utilisation of their results play a significant part in this.*"

In addition, the resolution stresses the significance of increasing the quality and relevance of research efforts, as well as prioritisation of development efforts at different levels of society.

As the expert organisation responsible for the allocation of competitive research funds for basic research, the Academy of Finland is one of the central actors implementing the government resolution. The primary objective of the Academy is to advance science and scientific research in Finland. The mission of the Academy is stated as follows (Academy of Finland Strategy 2003):

"*The Academy of Finland is a respected and attractive partner in cooperation that is committed to promoting a high standard of scientific research by means of*

- *long-term funding based on scientific quality,*
- *reliable evaluation,*
- *science-policy expertise and*
- *global cooperation.*

The Academy's aim is to raise public awareness of science and research as well as their social esteem. Furthermore, its work is geared to strengthening the diversity of research and its capacity for regeneration and to promoting the wide use of its results for the benefit of social welfare, culture, the economy and the environment."

The Academy's annual funding volume in 2003 was 185 million euros, which accounts for approximately 13 per cent of the Finnish government's research funding (Scientific Research in Finland 2003). The Academy's main funding instruments include the core project funding for individual projects (42% of total funding in 2003), research programmes (23%), support for researcher training (14%), Academy research posts (11%), international cooperation (8%) and the centre of excellence programmes. In 2003, no funding was allocated for the centre of excellence programme. In 2002, the share was 17% of total funding (Suomen Akatemian toimintakertomus 2004).

One of the general objectives of the Academy of Finland is to evaluate the quality, relevance and impacts of research. The Academy of Finland has recently initiated several development projects that will generate information on the performance of the Finnish research system. Collectively, the project is called "Sight 2006". Sight 2006 addresses four types of issues: effectiveness of the research funded by the Academy, impacts of R&D at the level of the innovation system, quality of scientific research in Finland as well as foresight at the level of the innovation system.

1.2 Objectives of the Study

The current study is part of the Academy's Sight 2006 initiative. It responds to the need for developing ex-post assessment of impacts from basic research in the Finnish context. The study was initiated in January 2005 based on a call for tenders issued by the Academy. The study provides information on the approaches of impact evaluation that have been adopted internationally in organisations similar to the Academy.

The purpose of the study is to generate information on the methods and indicators with which the impacts from basic research, as well as the impacts from the operations of a research funding organisation, can be assessed. This includes an examination of the benefits, disadvantages and applicability of the methods in relation to the nature and context of the operations in the Academy of Finland.

The objectives of the study are:

1. To identify and select relevant evaluation projects for more detailed analysis based on a review of international evaluation activity;
2. To describe the evaluation practises and approaches employed by the selected projects;
3. To identify and compare the evaluation approaches as regards their advantages, disadvantages and areas of application;
4. To present suggestions regarding the use of the evaluation methods as well as the development of evaluation practises at the Academy of Finland.

1.3 Overview of Methods for Impact Assessment

There is no single best methodology for the evaluation of impacts of basic research supported by public funding. Furthermore, the perspectives and points of departure of the assessment exercise vary depending on the contextual factors. The methodologies for impact assessment comprise a heterogeneous group of research methods, partly because the concept of impact can signify many different types of results and influences stemming from the research process.

In the context of basic research, the realised impacts can be characterised, for instance, as scientific, cultural and societal. Of the areas of research impact, the techno-economic impacts are perhaps the most well-known. The following, partly overlapping types of economic influences from basic research have been identified (Martin et al. 1996):

- increasing the stock of useful knowledge;
- training skilled graduates;
- creating new scientific instrumentation and methodologies;
- forming networks and stimulating social interaction;
- increasing the capacity for scientific and technological problem solving;
- creating new firms.

For other domains of research impacts such as cultural, societal or political influences, similar categorisations are even more complex to make, because the processes involved operate at many different levels of society and are characterised by a multiplicity of objectives.

Of the overall impacts that materialise as a result of research efforts, only part is manifested as concrete outputs from research. The eventual impacts from the research process are realised as a result of complex chains of influences that may take years or even decades to materialise. This complicates the assessment of impacts from basic research. The main challenges in the assessment created by the nature of basic research include:

- Time lag – when should impacts from basic research be assessed?
- Attribution – what is the significance of the research efforts for the observed impact?
- Appropriability – who benefits from the research efforts?
- Complexities – what are the mechanisms through which impacts are generated?

Because the objects of impact evalu-

ation are ambiguous and diversified, no single methodological tool can be employed to address all the issues. Many complementing methods are needed.

Unfortunately, there are well-developed methodologies for few areas of research impact. Traditionally, impact assessment has focused on the scientific impact and quality of research. The evaluation of societal and cultural impacts has received less attention, particularly in evaluation research that attempts to generate information as the basis for political decision-making. Thus, the impacts from basic research are often assessed with methodologies borrowed from other areas of evaluation research, such as the evaluation of technological development projects. As such, these methods may not be appropriate for the evaluation of basic research funding. In the long term, they need systematic developing and adaptation that takes into account the specific characteristics of the context in which basic research is carried out.

The methods employed in the assessment of impacts of basic research range from quantitative – perhaps more exact methods on well-focused research areas – to the more qualitative methods that provide a richer, more detailed description of the results of research efforts. Moreover, the applicability of the methods on different levels of aggregation varies: at the one end there are methods which are able to generate information on the national and even international influences of research efforts; at the other there are methods that can be used to create a detailed understanding of the contribution of a single researcher or research project. It should be kept in mind that all methods used for the collection of information require an element of judgement by experts in order to qualify as evaluation methods. Strictly speaking, expert analysis in some form is always needed in evaluations. Below, methods used to generate information on impacts are briefly described:

The *peer-review method* is the primary methodology employed by many funding agencies to evaluate research. The method is based on the assessments of research by experts in the scientific field. Peer review techniques are mostly employed to evaluate the scientific quality of research, but the method is also used to conduct ex-post assessments of the socio-economic impacts of research, known as the modified peer review. In modified peer review, the review panel includes members whose expertise enables them to assess the relevance and value of the research for the potential users and society at large. When employed to assessment of research impacts outside the scientific domain, the peer review method is often complemented with other methods that provide background information on the basis of which the panellists are able to form an understanding of the impacts of research.

In the evaluation of impacts from basic research, *bibliometric methods* are the prevalent methodology, partially because of the heavy emphasis of evaluation research on the assessment of scientific impacts. Bibliometric methods refer to a group of methods that use data from publications to observe patterns of influence from research efforts. Bibliometric methods cover different types of methods, including publication counts, citation analyses, co-word analyses and patent analyses. Bibliometric studies have been performed at many different levels of aggregation. At the macroscale, bibliometric studies have been carried out to characterise scientific activity and impact at the national, international and scientific discipline levels. At the microscale, bibliometric studies have been

undertaken to assess, for instance, the relationships between research funding and scientific quality, the influences between science and technology and the dissemination of research knowledge (Kostoff 1998).

Particularly at the national and international levels, *science and technology (S&T) indicators* are employed in the assessment of research efforts. The development of indicators was initiated by the Organisation for Economic Co-operation and Development (OECD) and the National Science Foundation(NSF), US, in the 1960s and the 1970s (Godin 2003). Today, these organisations are still among the primary providers of science and technology data. The OECD publication on S&T indicators, "Main Science and Technology Indicators (MSTI)" publishes a broad set of indicators covering the OECD countries as well as nine non-member countries ranging from measures on R&D expenditure and personnel, patents, technology balance of payments and international trade in highly R&D-intensive industries (Main Science… 2005). Similarly, the National Science Board (NSB) at the National Science Foundation publishes science and technology indicators that measure the performance of the R&D in the United States (Science and Technology Indicators 2004). The success of science and technology indicators can partly be explained by the availability of internationally comparable data and the easily communicable nature of the information produced by them. However, these indicators remain criticised because they emphasise inputs to the research process rather than outputs, and because they fail to capture the cultural and environmental impacts from science.

The *survey method* is employed to gather information directly from a large group of informants. In the ex-post assessments of basic research, surveys are a common method for collecting data from organisations or individuals that have received public funding for research efforts, who have participated in research programmes, or who are potential end-users of of research results. The benefits of surveys include their adaptability to specific research problems and contexts. With surveys, evaluators are able to collect both qualitative and quantitative information from the informants, and address different stages in the process through which research impacts are materialised. The disadvantage of the survey method is that understanding of the influences that generate societal impacts may remain fragmented. Statistical research methods are sometimes employed as a tool to refine the survey data to shed more light on these influences. The use of statistical methods in evaluation research is largely context specific and dependent on the availability of appropriate data.

Another approach that can be employed to refine the information gathered with surveys or collected from existing databases is the use of *cost-benefit analyses*. The purpose of these methods is to arrive at a figure that indicates the value of research as a proportion of the benefits to the costs of the research effort. The value of the research is often indicated as a rate of return on the investment, but also indicators employed in corporate finance, such as the net present value, can be used. The group of methods includes both large-scale econometric studies as well as survey-based studies with a focus on the utilisation of research in different industries. However, both approaches have their limitations: Econometric approaches are based on simplifying assumptions on the nature of the innovations, whereas survey responses may be biased towards in-

ternal activities of the responding organisation and thus provide false information on the effects of research on the industrial sector (Salter & Martin 2001).

A common methodology for assessing the impacts of research based on more qualitative information is the *case study method*. With case studies, information on the outcomes and impacts of research results can be gathered and communicated in a detailed, process-oriented fashion. Case-study based research is particularly advantageous in situations and contexts where the understanding of the impacts of research is poor, and there is only some knowledge on the causalities involved. Selecting the case studies appropriately, the method can provide novel understanding of the significance of the phenomenon and influences involved. A disadvantage of the method is the often poor generalisability of findings and the lack of understanding of total impacts of the programme or funding scheme investigated.

In addition, there are a number of less typical methods for assessing research impacts, including *modelling of programme theory*, *historical tracing* and *social network analysis*. A common element for these methods is their focus on the processes and knowledge flows with which impacts are generated.

In practise, individual evaluations often employ many complementing methods, because they focus on different parts of the process through which impacts are realised. In addition, through the use of multiple methods, the assessment is able to generate different types of information on the same phenomena. This increases the use of validity and reliability in the evaluation. The approach is known as *triangulation*.

1.4 Methodological Approach Employed by the Study

The approach adopted in the present study can be described as data-based: The study generates information on impact assessment methods and indicators by identifying and analysing international evaluation projects and practises. The data-based approach is retained throughout the study; also suggestions to the Academy's activities have to a large extent emerged from the analysis of the selected evaluation projects.

The scope of the study has largely been defined through the selection of the international evaluation projects. The present study stresses methods which involve indicators that can be used to communicate the nature of the impacts from the research efforts. For its starting point, the study adopted a broad and versatile notion of research impact that involves different domains of impact from basic research: techno-economical, social, cultural and political, etc. However, because the study is focused on impact indicators, the evaluation projects selected for analysis emphasise techno-economic impacts while leaving the other domains of impact – which evade quantification – to much lesser notice. To our knowledge, there are very few evaluations which attempt to characterise, for instance, the cultural, social and political impacts of science with indicator measures. However, although these diversified impacts are difficult to subject to quantification, they are no less important. Other than indicator methods should be used and developed for their assessment.

In the context of basic research, the term "impact assessment" is conventionally used to refer to quantitative studies that attempt to measure the impact of scientific publications on the scientific or

technological development process by measuring citations in publications or patents. Because the impact assessment methods are analysed from the perspective of an organisation that funds basic research, the term impact is used here to denote both impacts that result directly from the allocation of funding as well as impacts that materialise as a result of the conducted research. Strong emphasis in the study has been placed on impacts that are realised by the utilisation of research results for other than scientific purposes.

The typical methods used to assess the impact of research, such as bibliometric assessments and the peer-review method, have not been emphasised in the study. Also, evaluation projects already well-known to the Academy have not been included in the present study. The selection of evaluation projects limits and directs both the issues addressed in the analysis as well as the suggestions made on the basis of the study.

The method used in the study involves four distinctive, but in practise partly overlapping stages. The report has been structured to reflect these stages. The purpose of the first stage was to identify relevant evaluation projects for analysis. Potentially informative evaluation projects were identified through a review of international evaluation activity. Then, evaluation projects were selected one by one to form a sample of evaluation projects for a more thorough analysis. A more detailed description of this approach can be found in Section 2.1.

Second, the selected evaluation projects were analysed individually. Systematic data collection was facilitated by the use of a template onto which data from all analysed evaluation projects was collected in similar form. The template is presented in Annex B. The second chapter provides an introduction to the evaluation projects analysed in the study. Each evaluation project is briefly presented.

Third, the material was analysed by examining thematic areas that cut across the evaluation projects. A conceptual framework was developed based on the thematic areas that reflect stages in the process through which research funding results in outputs, outcomes, and societal impacts. The thematic areas have been identified through the analysis of the approaches that evaluation projects employ in order to characterise impacts from the activities of a research funding organisation. The framework provided a structure for the comparisons and analyses of the evaluation projects. The conceptual framework and the thematic analyses are presented in Chapter 3.

Fourth, the study generated ideas and suggestions for the Academy of Finland based on the analysed material. The Academy's evaluation activity was shortly reviewed to identify potential areas for development. Adopting the Academy's needs for impact evaluation as a point of departure, the selected evaluation projects were examined in an attempt to generate suggestions as regards the use and appropriateness of the methods exemplified by the analysed material for the Academy's evaluation activity. The suggestions are presented in Chapter 4.

2 Evaluation Projects

This chapter presents the individual evaluation projects analysed in the study. First, the methodology used to select the studies is presented. Second, the sample of selected evaluation projects is described as a whole by characterising their purposes, methodological approaches and levels of analysis. Third, the evaluation projects are presented individually. The introductions include a brief description of the main objectives for the evaluation as well as an analysis of the methods employed in the project.

2.1 Selection of Evaluation Projects for Analysis

The selection of evaluation projects followed a two-staged process consisting of a review of impact evaluation activity in international organisations and an initial assessment of identified, potential evaluation projects. The purpose of the review of international organisations was to identify organisations that actively evaluate the impacts of their operations or, better yet, purposively develop methods for the assessment of research impacts. The review of international organisations was based on published material in the Internet and in the literature focusing on the evaluation of research. The search was started from international research funding organisations similar to the Academy of Finland, and then expanded to other organisations. The review covered most OECD countries and included different types of organisations potentially active in research evaluation activity: research funding organisations, research and advisory councils, government agencies, science policy research units and consulting agencies. A list of the reviewed organisations is provided in Annex A.

The review resulted in a list of organisations and their publications in the field of research evaluation. The evaluation projects identified initially as relevant were assessed one by one and selected based on their potential of providing new insights and information for the evaluation activity at the Academy of Finland. After selecting the most potential evaluation projects, complementary projects were added to the selection. Although some limitations were made prior to the selection of the evaluation projects, the material selected was allowed to be rather diverse. The criteria used to select the sample were partly explicit, partly implicit. Although exceptions were made, the selection was carried out according to the following guidelines:

- Focus on basic rather than applied research;
- Emphasis on projects carried out relatively recently;
- New methodological approaches, strong emphasis on outcome indicators;
- Emphasis on projects of which the Academy has little prior knowledge;
- Broad geographic coverage.

The process resulted in a selection of twelve evaluation projects for more thorough analysis. In practise, the analysis focuses on a single, central report produced in each of the analysed evaluation projects supplemented by other, complementary information when appropriate. The analysed evaluation projects are:

1. Evaluation FWF: Impact Analysis, Joanneum Research, 2004;
2. The Impact of the EU Framework Programmes in the UK, Technopolis Ltd, 2004;

3. Funding Ranking 2003: Institutions – Regions – Networks, Deutsche Forschunggemeinschaft, 2003;
4. A Wealth of Knowledge: The return on investment from ARC-funded research, The Allen Consulting Group, 2003;
5. Methods for Assessing the Economic Impacts of Government R&D, Gregory Tassey, National Institute of Standards and Technology, 2003;
6. Review of Wellcome Trust PhD Training: Career paths of a 1988–1990 Prize Student cohort, The Wellcome Trust, 2000;
7. Assessing research impact on non-academic audiences. Final report submitted to the Research Evaluation Committee, Economic and Social Research Council, SPRU, 1999;
8. A Composite Performance Rating System for ATP-Funded Completed Projects, Rosalie Ruegg, TIA Consulting, 2001;
9. Outcome Indicators 2003/04, Foundation for Research, Science and Technology, 2004;
10. Standard Evaluation Protocol 2003–2009 For Public Research Organisations, Vereniging van Universiteiten (VSNU), Netherlands Organisation for Scientific Research (NWO), Royal Netherlands Academy of Arts and Sciences (KNAW), 2003;
11. PSA target metrics for the UK research base, Office of Science and Technology, 2004;
12. Taking Stock of R&D Across Three Sectors. Taking Stock of University & College Commercialization Efforts: Advising on The Way Forward. A Study for The Advisory Council on Science and Technology, Chris Riddle, 2003.

2.2 Description of the Sample

Due to the broad scope of the present study, the selection forms a relatively heterogeneous group of evaluation projects and reports. The next paragraphs are intended to give a picture of this diversity inherent in the selection.

The analysed projects respond to different types of needs within organisations that fund research and development. A majority of the projects are ex-post assessments of research impact, initiated by research funding organisations. The motivations of these projects are related to the accountability of the activities, credibility of the institute's mission or improvement of existing funding instruments. For most projects it is the combination of these aspects that have led to the initiation of the assessment exercise. In addition to the motivations described above, there are three projects in which the primary purpose for undertaking the project is to develop methodologies for the assessment of certain facets of research impact.

The ad hoc nature of the evaluation projects usually implies that there is no corresponding data from earlier periods that could be used to assess trends. Among the twelve evaluation projects there are four that present approaches in which data is gathered regularly to provide an understanding of the changes involved. In two of the projects reporting is conducted annually; in two other projects data is gathered and assessed every three years.

The selected evaluation projects can be roughly categorised into three groups: (a) individual projects that attempt to measure research impact (5 projects), (b) individual projects that attempt to review or develop evaluation methods (3 projects) and (c) reports that either describe a system of evaluation or provide an assessment of research performance at a national level with a set of indicators (4 projects).

Because other aspects of the evaluation projects are to a large extent dictated by the purpose they serve, the variance in the selection is carried on to the scope, objectives and methods of the projects. Among the selection there are four projects in which impact assessment is only part of a wider selection of approaches that are employed to inform decision-making. In addition to impact assessment, the projects for example investigate the rationale of the funding mechanism, assess the delivery of the programme, or provide descriptions of best practises found in the programmes. In some cases impact assessment is only a minor part of the project. Other projects focus either on assessing research impact in various stages of the process or focus on a single measure of research impact, e.g. researcher mobility, return on investment etc.

Also the level of analysis varies according to the purpose of the project. In the selection, there are four evaluation projects which look at the impact of all funding allocated by a granting organisation, four which focus on individual programmes or funding schemes, two which focus on the national level, one that focuses on research institute evaluations and one that looks at individual projects. In each case, data is gathered from suitable units of analysis. For instance, in programme level evaluations data typically comes from projects provided by individual scientists. When the evaluation focuses on the operations of the funding organisation as a whole, data can also be gathered from institutions, depending on the nature of the assessment.

The appropriate time lag between assessment and research effort depends on how the measurement of research impact is carried out. More specifically, there are conceptually different stages in the process by which research influences society, and generally, the further one focuses in this process, the more time should have elapsed for one to observe sizeable impacts. In the analysed evaluation projects this shows in the following way: for surveys conducted at project level, data is usually gathered from projects that have ended one to three years ago. For evaluation projects estimating broad socio-economic impacts or gathering data on researcher mobility, the associated lag may be as long as ten years. Typically, the projects have gathered data from events that have taken place approximately three years ago, ranging to shorter or longer time periods depending on the selected approach.

As discussed above, the selection of evaluation projects emphasises methods that are being developed to fill existing gaps in the current toolbox for research impact assessment. However, more typical methods still dominate the projects analysed. In seven of the twelve projects, surveys have either been the primary method for impact assessment or had a secondary role in generating information for further refinement. Other common methods are document analyses and interviews. Explicit attempts to model programme theory are taken in four projects. Modelling of programme theory is a prominent method especially in projects that attempt to estimate the return on investment through measurement of

wider socio-economic impacts. To the more quantitative side, the projects also employ simple bibliometric indicators or statistical analyses based on data gathered with survey questionnaires.

Moving to the qualitative end of the method spectrum, it is apparent that the scope of the evaluation projects in the selection becomes more focused. Typically, the projects attempt to create a more profound understanding of the phenomenon investigated. In the more qualitative projects of the selection, also the element of method development is stronger. As for methods such as the case study method, expert judgment, historical tracing and analysis of the formation of social networks in research programmes, there are two to three projects that employ each of them. In addition to the methods already mentioned, the projects analysed employ economic case studies and composite indicators, both in two instances.

The selection of projects can be elaborated further on the basis of their focus on the process by which benefits from research are channelled to different actors in society. The indicators can roughly be categorised by the stage of the process they focus on.

All but three of the analysed projects attempt to measure the outputs of research. This is the most common focus of the analysis in the projects, which is not unexpected, as the generation of outputs can be relatively easily measured also in quantitative terms.

Influenced by the way in which sampling of the twelve projects was conducted, methods and indicators for assessing the impact on stakeholders outside the scientific community are also well represented. Altogether nine of the analysed projects include at least one indicator for assessing impacts on actors external to the scientific community. The methods used for impact assessment are not as robust as those for measuring research outputs.

Next, there are seven projects which shed light on the ways in which network creation, cooperation and diffusion of research results can be studied. The methods range from simple indicators on the co-funding of projects to complex methods for mapping of network relationships.

There are six projects that provide indicators and descriptions of the pattern of funding allocation, including for example relative shares of funding sources and additionality impacts created by the funding. Four of the projects address the issue from the perspective of a granting organisation and two of the projects from a national perspective.

Finally, there are four projects which attempt to characterise the funding recipients to provide increased understanding of the aspects that influence the likelihood of receiving a grant or to explain the distribution of outputs and impacts. They serve as additional, supporting indicators but do not directly measure research results.

Only one of the projects addresses all of these issues. It is common that an evaluation project addresses only one "end" in the process: there are projects that focus on funding allocation (inputs) and outputs but not on the broader socio-economic impacts as well as projects that focus on network creation and impacts but not on the funding patterns. This also reflects the methods used; a single method cannot effectively capture all the aspects described above.

2.3 Approaches to Impact Assessment

After reading the previous section, the reader hopefully has a better understanding of the variation in objectives,

issues addressed and methods in the group of projects analysed. The following sections now present each of the projects in more detail. First, evaluation projects carried out from the perspective of a research funding organisation are presented. The order in which they are presented reflects a shift from quantitative indicators to the more qualitative methods. Simultaneously, the extent to which projects focus on method development also grows. Second, projects that develop or describe a system of impact assessment with a regular assessment cycle are summarised. Finally, reports that focus on the performance of the national research base are presented. After summarising the main content of each project, a discussion of the most interesting and relevant issues emerging from the project follows. A more detailed analysis of the themes addressed in the analysed evaluation projects is presented in Chapter 3.

2.3.1 Evaluation FWF – Impact Analysis

The Austrian Science Fund (FWF) is Austria's central body for the promotion of basic research. The FWF funds all branches of science and humanities. Its funding volume exceeded €100 million in 2004. The responsibilities of the Austrian Science Fund are the promotion of (FWF Austrian Science Fund 2005):

- *"High-quality scientific research, which represents a significant contribution to society, culture and the economy.*
- *Education and training through research, because support for young scientists represents one of the most important investments in the future.*
- *Knowledge transfer and the establishment of a science-friendly culture via an exchange between science and other areas of society."*

The impact evaluation of the Austrian Science Fund was carried out by Joanneum Research as part of a larger evaluation project the purpose of which was to evaluate the position and activities of the two most important Austrian funders of research and innovation, the Austrian Industrial Research Fund (FFF) and the Austrian Science Fund. The consortium that conducted the evaluation consisted of a large international group or experts and was led by Technopolis Ltd.

The impact analysis of the FWF was conducted as a separate assessment that focused only on the activities of the FWF. The objective was to generate an assessment of the significance of FWF funding in the Austrian context. As the majority of FWF funding (83% during the years 1998–2003) was allocated to Austrian universities and research institutions through project-based programmes, other types of funding instruments, e.g. personal grants, were excluded from the analysis. The ad hoc nature of the evaluation and its focus on project-level instruments made the survey method a logical choice for data collection. The objectives of the survey were to (Evaluation FWF… 2004, p. 30):

- *"Characterise the institute or research localisation (for both approved and rejected proposals)*
- *Characterise the research project/proposal (for both approved and rejected proposals)*
- *Characterise the results (only for approved proposals)*
- *Characterise the impacts on the individual researchers (only for approved proposals)*
- *Characterise the rejected proposal (only for rejected applications)"*

Since characterisation of rejected projects in relation to accepted projects was one of the points of departure, two

types of questionnaires for FWF applications were developed: one was administered to projects that received funding, another to project teams that had applied for funding unsuccessfully. The unit of analysis for both questionnaires was the project.

In addition to the surveys, the evaluation team analysed the patterns of funding allocation based on data supplied by the FWF. This data is based on submitted applications and could be used to describe funding allocation from different angles: funding volumes in different disciplines, approval and funding rates by solicited amount, application numbers and approval rates by institution, etc. In addition to the descriptive analysis, the evaluation team created a model of project approval. This model attempts to predict project approval based on the scientific field in question, characteristics of the application and characteristics of the coordinator.[1] The results show that there are no "natural winners or losers" among the applications, since changes in the measured characteristics of applications are able to explain very little of the variation in application acceptance. When the effect of the FWF's project rating system is included in the model, the predictive power is increased, but still remains low. The evaluation team concludes that "the actual decision to find the project is not based predominantly or even solely on its rating, but almost exclusively on the verbal assessments" (Evaluation FWF 2004… p. 22).

The results of this exercise are not unexpected. The value of this method lies mostly in the assurance that it gives; the model works as a way of checking that the funding system is not inherently biased towards some disciplines or applicants. However, the method is also quite laborious. But, if there is a reason to believe that some applicants are favoured over others, the method provides a quantitative way for examining this.

In addition, the evaluation team also conducted analyses of a database provided by the Austrian University Council. The database, ABIV, includes institute-level data both on publications and sources of external funding, enabling modelling of "output efficiency", where the number of publications is regressed on different external funding sources and fields of science. The results show that, in comparison to other sources, FWF funding is quite effective in generating scientific publications in a range of disciplines. Logically this can be expected, and thus the analysis does not provide many new insights into the effectiveness of FWF resource allocation.

The surveys addressed a wide range of topics reflecting the objectives presented above. Because all data is gathered from the same source, the project manager, even though the breadth of issues addressed varies, the reliability of the findings is not constant across the range of results. When data is gathered of institute characteristics, direct outputs and project impact on research team, the project manager or coordinator is the best informant. But when the survey questions address broader socio-economic impacts or the reasons why the evaluation panel rejected the proposal, the knowledge of the project coordinator of the investigated issues is limited. Thus, one should be cautious when interpreting the results of the survey.

These limitations notwithstanding, comparing the answers from the rejected projects with those from the accepted ones, the evaluation team has been able to extract some interesting findings: FWF funding seems to be directed towards internationally leading organisa-

1 The evaluation team used a probit model for estimation.

tions and the projects "are used more often to extend already existing main research activities and less to establish new main research activities at the research unit" (Evaluation FWF… 2004, p. 45).

As for project results, the survey questionnaires measure publication output, impact on performing organisation, network creation and exploitation of results. The treatment is mostly descriptive. The results provide evidence of relatively high impact of funding on a range of indicators, but without points of reference it is difficult to establish the level or magnitude of these impacts. Comparisons between scientific disciplines alleviate this problem a bit, but international referents or time series analyses would be needed to establish the significance of findings.

Results from some of the questions indicate low impact. For example, interpreting the relationship between two of the indicators, relevance of project results for industry and network creation, the evaluation team states that "though the relevance of research results for industry is rated relatively high, researchers apparently do not see the need, or do not have the possibilities to forge links with industry" (Evaluation FWF… 2004, p. 46).

Because the organisational mission of the FWF is very similar to that of the Academy of Finland, the methodologies used are all potentially relevant to the Academy. Both organisations focus on funding basic research and thus the challenges of impact assessment are similar. The FWF exemplifies the types of findings that can be acquired by using a cross-sectional survey. The statistical methods used are also relevant to the Academy, but they need to be modified based on the data that is available from the reporting systems of the Academy of Finland.

The methods used in the assessment are most appropriate in a large evaluation which covers all of the activity of a funding organisation. For evaluations which concern individual programmes or research institutes, the methods should not be transferred without modifications. Nevertheless, parts of the survey instrument also provide elements for research programme evaluations. In the FWF assessment, the measures of impacts are in a very general form, and thus when applied to programme level they ought to be customised to the needs of the individual programme. In comparison to interview-based methods, the ones used by the FWF evaluation are quite inexpensive. In an evaluation with such a wide scope, it is economic to use existing databases and gather complementary information with surveys.

2.3.2 The Impact of the EU Framework Programmes in the UK

The Office of Science and Technology (OST) in the Department of Trade and Industry (DTI), UK is responsible for developing and coordinating UK Government policy on science and technology both nationally and internationally. OST allocates the Government Science Budget into research via eight research councils focusing on different scientific fields. In 2004, the Science Budget amounted to nearly £2.4 billion (About the Office of Science and Technology 2005).

In 2003, the structure of the Seventh EU Framework Programme (FP) was being planned and policy makers in the UK needed information to guide negotiations on the priorities of the FP. As part of their contribution, the OST commissioned an evaluation to assess the impact of the EU Framework Programmes in the UK. Based on a call for tenders,

Technopolis Ltd. was selected to conduct the assessment. The objectives of the evaluation were manifold (The Impact… 2004, p. 1):

1. *"To obtain well-found evidence on the impact on UK interests of participating in Framework and the added value that European funding brings compared with other options.*
2. *To obtain well-found evidence on the drivers and barriers to organisations participating and how the Framework instruments, funding levels and contractual processes impact on this, including by comparison with national collaborative programmes.*
3. *To support decisions on policy in relation to future Frameworks and any necessary national support measures to improve performance in participation levels, technical excellence, access to technology and exploitation.*
4. *To identify lessons for potential participants on the keys to successful projects."*

The themes addressed by the evaluation included considerations of the rationale and added value of the FP in relation to other programmes, its impacts in the UK, the performance of UK participants in the FP relative to other countries, and programme delivery. For the current study only the second and parts of the first topic are relevant.

At the time of the evaluation, the Sixth EU Framework Programme was already ongoing, but its impacts could not yet be assessed. The evaluation team gathered information on impacts from the Fourth and Fifth EU Framework Programmes. Reflecting the broad scope of the assessment, the evaluation team employed a range of methods, including interviews with various stakeholder groups, questionnaire surveys, document analyses and case studies. Information for impact assessment was generated with a questionnaire survey and a few additional case studies.

The survey method used to assess the impact of the FP in the UK is similar to the one employed in the FWF evalution. However, there are a few important differences, partly because the perspective adopted in the former evaluation is that of a national policy maker that views the FP in relation to other policy instruments available. First, the descriptive analysis of funding patterns is not included in the analysis, although this could be seen as a way to characterise the impact of the FP. Instead, there is a related analysis of the "added value" of the FP. This is assessed through the motivations of participants for applying for FP funding. The findings of the evaluation team indicate that participants apply mostly to access research funding, which then allows them to "tackle problems of international relevance" (The Impact… 2004, p. 14). The objective of cost sharing is not as significant a reason to participate. The results indicate that FP fills an important gap in national UK funding for research. However, more than one fifth of the projects could also have been supported at a national level.

Second, when the scope of the evaluation is an individual funding scheme with specific objectives rather than the impact of a funding organisation in its totality, it is possible to measure impact of basic research against the objectives. This provides a complementary, less general measure for assessing impacts. This can be done in two ways, either by asking respondents how their project has contributed to programme objectives, or by first asking respondents to indicate (project-level) goals they consider important and then measuring success against these objectives. Both methods were used in the FP evaluation. Because the objectives of the FP address the de-

velopment of research capabilities at the level of the European Community, the objectives are stated at a very general level. For this reason, it is difficult for individual projects to assess their impact in relation to the objectives and, thus, reliability of their responses remains low.

When the objectives are stated in a general fashion, the second approach may provide more relevant information on the ability of the programme to correspond to the expectations of the various participant groups (The Impact… 2004, p. 27):

"The findings suggest that Framework is better at delivering the 'softer' knowledge and networking benefits sought most by the universities and public research institutes, and is least able to deliver the more concrete, commercially oriented outputs and benefits sought by companies…"

As an additional benefit, the analysis provides information on the service offering of the funding scheme as experienced by participants.

Third, to deepen the understanding provided by descriptive analyses of survey data, the evaluation team analysed associations between project success and other measured variables. As a proxy variable for success the evaluation team used a measure of costs and benefits experienced by the participant. Such an analysis complements descriptive statistics by providing information on potential causal relationships and characterising the differences between successful and unsuccessful projects. Some notable correlations were found. For instance, the findings indicate that learning takes place among the participants: earlier experience with FP is associated with higher benefit-cost ratio.

In addition to the survey, the evaluation team investigated wider impacts of the FP with case studies. As in the case of the survey, the broad scope and general nature of FP objectives made it difficult to assess impacts based on project-level data. The evaluation team noted that project-based assessment of wider impacts may not be a feasible solution, because significant and observable changes in society hardly ever result from individual projects. Overall, it appears that aggregation of project level data to assess the impact of a large international programme is a complex task which, at times, does not lead to significant results.

2.3.3 Funding Ranking 2003: Institutions – Regions – Networks

The German Research Foundation (Deutsche Forschungsgemeinschaft, DFG) funds research in German universities and research institutions. It supports "individual projects and research collaboration, awards for outstanding research achievements, and funding for scientific infrastructure and scientific cooperation" (DFG Mission 2005). Its annual funding volume is approximately €1.5 billion which amounts to 34 per cent of total third party funding of German universities. DFG has reported on the distribution of its funding allocation on three occasions, in 1997, 2000 and 2003. The present study focuses on the report published in 2003, which reviews the performance of German universities and non-university research institutions during the years from 1999 through to 2001. The report is entitled Funding Ranking 2003: Institutions – Regions – Networks, and it was conducted by DFG in cooperation with the German Federal Statistical Office, the German Academic Exchange Service (DAAD), the European Liaison Office of the German Research Organisation (KOWI) and the Alexander von Humboldt Foundation (AvH).

The assessment takes the form of an evaluation that ranks German universities based on different indicators. The purpose of the evaluation is not so much to assess the impact of research conducted at universities, but to identify locations of academic excellence, i.e. universities and research institutes that, in comparison to other research organisations, perform well in academic activity both in absolute and relative terms.

The evaluation builds on earlier work extending the time series available on the distribution of DFG funding, but also broadens the scope of the evaluation activity. The evaluation develops new methods for rating academic institutions: it "undertakes to expand the base data and to include new, revealing aspects of the world of research – internationalism, network building and the international resonance of academic publications – in the analysis. This is not intended to bring the discussion of the evaluation of research, of the best-suited methods and indicators, to a close, but rather to place it on a broader basis" (Funding Ranking 2003, p. 5). In addition to these new indicators of performance, the evaluation expands the scope of the earlier reports by viewing academic excellence not only at the level of institutions, but also by aggregating the findings to the level of German regions. Furthermore, the evaluation also analyses networks that have formed within the main scientific disciplines.

The ratings are based on five different types of measures indicating academic excellence: DFG funding approvals, centrality in research networks formed by DFG cooperative programmes, distribution of DFG reviewers, internationality of research and publication output. The distribution of DFG reviewers refers to the number of reviewers working for the DFG per institution and research area. These include both peer reviewers of the DFG that decide on proposal acceptance and special reviewers who serve as advisors in their field of specialisation. Internationality of research is operationalised with several ratings: number of visiting AvH prize-winner scientists, number of DAAD scientists and academics, number of DAAD students and graduates and, finally, participation in the Fifth EU Framework Programme. For each of these indicator classes, the evaluation provides a rating for the universities in absolute terms and relative to the number of scientists and academics working in the given organisation. The results are then disaggregated to the level of the main scientific disciplines and regions.

The purpose of the DFG rating differs from a typical impact assessment and, accordingly, the underlying logic of the evaluation is different. Impact assessments at least implicitly attempt to create evidence of causalities between the funding provided and the impacts from research projects. In contrast, the point of departure in the DFG evaluation can be characterised as follows: Throughout the evaluation the associations between DFG funding approvals to universities and other indicators are analysed, but the purpose of this exercise is also to show that DFG funding itself serves as an indicator of academic excellence. The use of DFG funding approvals as an indicator of academic excellence is legitimised by arguing that DFG approvals correlate strongly with total third party funding, and that high levels of third party funding indicate strong research activity with high relevance for non-academic stakeholders. This may not be a justified approach to impact assessment, but the methods used to study associations between DFG funding approvals and the other performance indicators

can still be useful. Namely, the analysis of associations between DFG approval rankings and other performance indicators show very high correlations, even when corrected relative to the number of academic staff in the organisations. However, rankings based on absolute figures are much stronger than the relative figures. The conclusion made by the DFG evaluation is that the volume of DFG funding approvals per research organisation is a good indicator of academic performance (Funding Ranking… 2003, p. 130):

"As a whole, the various findings first and foremost confirm the assumption that was already a central theme of the first so-called 'DFG ranking': DFG approvals are a good performance indicator. They display a high correlation to third party funding overall, with the appeal to visiting researchers from abroad, with the participation in international programmes and with the level of publication activity as well as acceptance success in international journals".

As a method of impact assessment this type of an evaluation that takes the form of ranking list is limited to assessing the performance of research organisations. Because the main instrument with which granting organisations fund research is an individual project, ranking lists may not be the best choice for assessing the impact of such an organisation. However, there may be exceptions to this conclusion. If a large proportion of the total funding of research institutes comes from a single funding organisation being assessed, assessment carried out at the institute level may be feasible. Moreover, if an evaluation project focuses on an individual funding scheme the purpose of which is to fund research units instead of research projects, then such a comparative method that rates individual units may be called for.

2.3.4 A Wealth of Knowledge: The Return on Investment from ARC-Funded Research

The following evaluation project is in some aspects much more ambitious than the three projects presented so far. In the assessment of research impact it goes a step further, and attempts to estimate the return on investment in research funded by a granting organisation, the Australian Research Council (ARC). The ARC funds primarily basic research and research capabilities in the Australian universities and attempts to enhance the links between university research and end-users. The purpose of the ARC evaluation project was to generate information on the earlier returns on investment to inform political decision-making regarding future levels of research funding. The assessment was conducted by The Allen Consulting Group and it is called "A Wealth of Knowledge: The Return on Investment from ARC-Funded Research."

The ARC evaluation presents two methods of estimation, one based on estimating the impact of ARC funding on Australian multifactor productivity (MFP) growth and another based on estimating different types of benefits from publicly funded R&D. The first method is less detailed and serves to provide a rough estimate of the return on investment. The estimate is based on a literature review. The structure of the logic is as follows: first, the evaluation team estimates that the contribution of MFP growth to total output growth is approximately 40 per cent. Second, the evaluators assess that approximately half of MFP growth is due to R&D. Third, they assume that 25 per cent of the contribution of R&D to MFP growth in Australia can be attributed to publicly funded R&D. Then, based on the share of ARC's funding of total public R&D

funding in Australia (5.2%) they arrive at an estimate of the share of MFP growth attributable to ARC funding: 0.65 per cent. However, they argue that ARC funding is more "effective" in generating higher returns than average public funding because of its competitive and targeted process of funding allocation. Therefore, they estimate twice as high a figure for the ARC's contribution: 1.3 per cent. Finally, the evaluators calculate the social rate of return of ARC funding. Applying the above logic, they calculate the permanent increase in GDP due to ARC activities during the years 1990–2000 and divide this with the investments made by the ARC. As a result, they arrive at a figure of 51.5 per cent.

The second method is more detailed and provides the actual figure of return on investment concluded by the assessment. First, the evaluation team has reviewed the extant literature in order to identify and categorise different types of benefits from basic research. Based on the review they arrive at a model with six categories of benefits shown on the left-hand side of Table 1.

Second, they operationalise each of the benefit channels, and calculate the return on investment in each area. In order to do this, many simplifying and limiting assumptions have to be made. For instance, the social rate of return calculated via the first category, benefits from building the basic knowledge stock, is estimated in the following way: Based on patent citation analysis, the evaluators assume that approximately 10 per cent of the ARC-funded research "finds a route to use", i.e. research is conducted collaboratively with industry or research findings are published and subsequently used by others for commercial purposes. In order to estimate the social rate of return of such research, the evaluators analyse a number of case studies they have conducted on Australian high tech companies that have been funded (indirectly) by ARC, and arrive at a figure of 10 per cent. A direct multiplication of the two above numbers provides them with an estimate of the overall rate of return from ARC research funding (regarding the benefits from the first category): 10 per cent.

To provide another example, the social rate or return calculation regarding the third category, direct benefits from improving the skills base, is carried out as follows: First, five categories of students who benefit to some degree from ARC funding are identified:

1. post-graduate students directly funded by the ARC (75%);
2. post-graduate students working in projects or facilities funded by the ARC (20%);
3. post-graduate students that are not directly funded but benefit from high-guality research infrastructure funded by the ARC (10%);
4. non-research post-graduates whose educational experience is improved by the strong research component in Australian higher Education sector (2%);
5. under-graduate students (0.5%).

Second, the evaluators make two types of assumptions. The first regards the extent to which the skills formation of the students can be attributed to ARC funding and the second the productivity premiums that are associated with post-graduate and under-graduate degrees. For the former assumption, the estimated figures are given above in parenthesis after the category in question. For the latter, the productivity premium for post-graduate degree holders versus under-graduate degree holders is estimated at $13,200 per annum while that for under-graduate degree holders versus year 12 completers is estimated at $38,600. To

correct for the fact that the productivity premium associated with different qualifications may result partially from students' personal attributes, the evaluators assume that only two thirds of the productivity premium can be attributed to ARC funding. The social rate of return is finally calculated by multiplying (1) the productivity premium per annum, (2) the ARC contribution to the formation of the skills of the student and (3) the number of students per year for each category of students. These are summated to form the total output premium of ARC funded students. The analysis yields results such as this (The Allen Consulting Group 2003, p. 53):

"*The total permanent output premium for the 2002 student cohort is… $34.5 million. Therefore, given that total* ARC *funding was $272 million for 2002, the direct social rate of return from skills formation due to ARC funding can be estimated at approximately 12.5 per cent.*"

Because of space limitations, the remaining rate of return calculations are not summarised here. The interested reader may turn to the report by the Allen Consulting Group for a complete discussion on the estimation.

Finally, the total return on investment is calculated by summating the val-

Category of benefits	Measured Benefits
Benefits from building the basic knowledge stock	In this area a 10% social rate of return on average from ARC funding (clearly some ARC funding generates higher returns in this channel and other types of ARC funding generate lower returns through this channel) with a ten-year time lag was identified. This is based on the relative success of ARC-funded research finding a route to use and the high pay-offs achieved when this occurs.
Benefits from generation of commercialisable intellectual property	In this area an average 3% social rate of return from ARC funding with a ten-year time lag was identified. This estimate is based on observed impacts from commercialisation of ARC-funded research over the past 25 years.
Direct benefits from improving the skills base	In this area an average 12.5% social rate of return from ARC funding with a four-year time lag was identified. This estimate is based on observed output impacts from the skills formation to which the ARC contributes.
Benefits from improved access to international research	In this area an average 7.5% social rate of return from ARC funding with an eight-year time lag was identified. This estimate is based on the level of access to international research funding that the ARC enables Australia to gain and the returns we obtain from this international research.
Benefits from better informed policy making	In this area an average 6% social rate of return from ARC funding with an eight-year time lag was identified. This estimate is based only on the observed impacts associated with microeconomic reform in Australia and the assumption that the ARC contributed only 0.25% to microeconomic reform policy.
Health, environmental and cultural benefits	While the benefits in this area are likely to be significant, it has not been possible within the time constraints of this study to estimate returns in monetary terms.

Table 1. (Source: The Allen Consulting Group 2003, p. 70)

ues from the individual benefit categories. The rates of return on the six individual benefit channels which make up the total ARC social rate of return are presented on the right side of Table 1. The estimate produced by this process is a total social rate of return on ARC investment in Australia of 39 per cent.

The above examples illustrate the speculative nature of the assessment. In both methodologies there are a number of very gross assumptions which need to be taken in order to arrive at a concrete figure of return on investment. Moreover, the assumptions appear to be based on a number of different methods ranging from literature reviews to case studies to mere educated guesses. Thus, the figures to which the evaluation team has arrived at are not very reliable but, rather, depend on the definition of benefit categories, the ways of operationalisation and the assumptions made. Moreover, as the evaluation team notes, the approach used is highly specific to the Australian environment, and therefore not amenable to copying to other environments as such. However, being conducted by a granting organisation, the methods employed may provide ideas as how to conduct or develop impact evaluation activities also at the Academy of Finland. The evaluation project exemplifies many of the challenges involved in the estimation of the economic impact of basic research.

Despite its limitations, the ARC assessment is an interesting contribution to the impact assessment literature. It is important to keep in mind that the objective of the evaluation is very challenging and, thus, allowances regarding the methodological purity have to be made. The assessment of research impact is difficult even at project-level, let alone attempting to estimate the return on basic research funding at the level of the whole organisation! Due to the sensitivity of the estimated figures to the assumptions made, it may not be very sensible to base political decisions on future funding levels on the figures provided by this estimation. Instead, the value of an exercise such as this may be in the understanding it brings of the benefit channels and the mechanisms through which the impacts of research funding materialise. In order to implement a method such as the one used by the Allen Consulting Group, one has to create an explicit model of how the funding of research influences society and understand the contribution of funding in different areas, which is an important contribution in its own right. Although the reliability of the method and the use of its results for political decision-making can be questioned, the ARC assessment complements evaluation approaches that are based on assessments of impacts at the level of individual research projects.

2.3.5 Economic Impact Assessment at the National Institute of Standards and Technology

The National Institute of Standards and Technology (NIST) is a non-regulatory federal agency operating within the U.S. Commerce Department's Technology Administration. The NIST develops and promotes measurement, standards, and technology to enhance productivity, facilitate trade, and improve the quality of life in general. Its budget is approximately $860 million per annum. The NIST implements its mission via four cooperative programmes (NIST: General Information 2005):

– *"the NIST Laboratories, conducting research that advances the nation's technology infrastructure and is needed by U.S. industry to continually improve products and services;*

– *the Baldrige National Quality Program, which promotes performance excellence among U.S. manufacturers, service companies, educational institutions, and health care providers; conducts outreach programs and manages the annual Malcolm Baldrige National Quality Award which recognizes performance excellence and quality achievement;*
– *the Manufacturing Extension Partnership, a nationwide network of local centers offering technical and business assistance to smaller manufacturers; and*
– *the Advanced Technology Program, which accelerates the development of innovative technologies for broad national benefit by co-funding R&D partnerships with the private sector."*

The economic impact assessment methodology developed by the National Institute of Standards and Technology is in many respects similar to the method used by the Allen Consulting Group. They both attempt to inform policy: at NIST, retrospective impact analyses "were initiated not only to inform management on relative performance across projects, but also to develop a database of quantitative and qualitative impact information that (1) educated the policy and budget processes on the types and magnitudes of economic impacts and (2) imparted credibility to the agency's overall mission." (Tassey 2003, p. 45) The NIST approach is developed to assess the impacts of government funded R&D measured with a set of financial indicators. Thus, similar challenges permeate the assessment: how should potential impacts be defined? How should they be measured? What should be included in the calculation of costs? The main question is shared by two methods, if impacts of R&D are observed, how are they translated to a numerical value that has significance to stakeholders?

The methodology developed at NIST explicitly focuses only on research programmes with potentially high impacts (Tassey 2003, p. 45):

"…the subjects of economic impact studies are not chosen randomly. Rather topics are selected based on the R&D agency's perceptions that significant economic impact had occurred, thereby affording the opportunity to document both the types and magnitude of the economic contributions being realized from laboratory research programs and projects".

Moreover, the analysis focuses only on those industries that directly benefit from the development of a new technology.

The assessment process of the method developed by NIST can be described as follows. After screening for areas in which substantial market impact is likely to have occurred, the major steps of the assessment are (Tassey 2003, p. 61):

1. *"Construct a taxonomy of the relevant technology, which disaggregates the technology into its major elements.*
2. *Map this taxonomy onto the industry structure and competitive dynamics associated with development and delivery of the technology.*
3. *Develop testable hypotheses that represent the relationships among technology, strategy, and economic trends.*
4. *Utilize this framework to select a set of qualitative and quantitative output and outcome metrics.*
5. *Select measures that summarize the metrics and are intelligible to stakeholders and other target audiences.*
6. *Develop and implement a data collection plan, emphasizing primary data collection from impacted industries.*
7. *In analyzing results, make careful determination of degree, if any, to which results can be extrapolated to other economic sectors.*

8. *Write a full report and then prepare several summaries of varying lengths for each target audience."*

The first four stages of the assessment process indicate that modelling programme theory, i.e. mapping the mechanisms with which the research programme is supposed to generate impacts, is a necessity for conducting the exercise. The evaluation team must understand both the logic of the research programme as well as the dynamics of the related industries that benefit from the government research effort.

"As an initial step, the analyst must identify the objectives of the research program and then carefully describe the technical outputs. These outputs constitute the technological infrastructure that government contributes to an industry's overall technology. Thus, they are the drivers behind the eventual economic impacts.… Understanding how this infrastructure complements the industry-developed proprietary technology elements is essential to the eventual selection of metrics, impact hypotheses, and construction of the survey instrument." (Tassey 2003, p. 38)

The results of the economic impact estimation are given in the form of financial measures. The use of three complementary measures is recommended by the author. These are the net present value (NPV), the benefit-cost ratio (BCR) and the internal rate of return (IRR). Each of them provides somewhat different information on the returns of the research programme. The assessments conducted by the NIST display great variance among the results, for example, the social rate of return varies from 32 per cent to 877 per cent.

It is clear that the methods developed by the NIST cannot be applied to the context of basic research as such. However, they may provide ideas for measuring impacts of research funding organisations in the applied end of the research spectrum, where the range of potential users is limited. Such areas may be found for example in narrowly focused research programmes within the natural and medical sciences. Organisations that primarily fund basic research are increasingly aware of the need for research that is relevant for other stakeholders and the results of which can be exploited. In the future, potential users may well be easier to identify also in more basic research-oriented research programmes. Therefore, the use of methods similar to the ones developed by the NIST may become more feasible in the future.

2.3.6 Career Paths of 1988–1990 Prize Student Cohort

From quantitative estimation of returns on investment the present study moves on to a quite different approach for assessing the impacts of research funding, namely analysis of researcher mobility. Researcher mobility refers to the movement of academics from one organisation to another. Typically, researcher mobility is assessed to provide information on knowledge diffusion from research organisations to industry or on the internationality of research. Methods for assessing researcher mobility are exemplified in the next two cases.

The Wellcome Trust is a UK-based independent foundation that funds biomedical research. The trust promotes research by providing biomedical research funding, organising research funding programmes in the medical humanities, providing technology transfer funding to encourage commercial application of research and by funding activities for promoting public engagement with science (Wellcome Trust… 2005). A sizable share of its funding is allocated through

personal award schemes which support individuals and provide increased possibilities for an academic career. The Trust considers the assessment of the outcomes of these award schemes of high importance, both from the perspective of the Trust as well as the awardees.

In 1999, the Wellcome Trust undertook an assessment of the outcomes of its investment in postgraduate training. The assessment focused on the 136 Wellcome Trust-funded PhD students that received their Prize Studentship between 1988 and 1990. The time frame of approximately ten years allowed sufficient time for the Prize Students to submit their thesis and gather work experience (five to seven years) after the examination. Interesting information on their career progress could thus be gathered by the assessment.

The method adopted by the Wellcome Trust included both analyses of publication databases as well as interviews with the awardees. Initial information on the subsequent career progress was provided by the Wellcome Trust's own grant records. A preliminary publication search was conducted to identify individuals that were still active in research. After locating the individual, the primary method of data collection was through telephone interviews.

The analyses conducted in this assessment can be grouped into two categories: publication output and researcher mobility. As for publication output, distributions of publications among the cohort members generated some interesting findings: the publication output was significantly higher for men than women and only 14 individuals contributed over a third of all publications. Overall, the analysis indicated considerable research productivity of the cohort.

To provide a description of the process by which the careers of the researchers progressed, researcher mobility was reported at three points in time: At the first position after the submission of a thesis, position after three years, and after approximately 10 years (current positions at the time of the assessment). The results showed that just under half (46%) had remained in academic research, but 73 per cent of those who had left were still working in science, medicine, or health-related fields. As for international mobility, 33 out of 58 researchers that had remained in academic research had obtained experience outside the UK, and almost half of them had returned to the UK.

The primary purpose of the Wellcome Trust evaluation was not to assess impact of research. Rather, it attempted to measure outcomes from the perspective of the personal award scheme's objectives. In other words, the assessment does not touch upon the subject of how the research conducted by the Prize Students influenced their subsequent work or the organisations they were employed in. If such an evaluation study were conducted, the elements of the Wellcome Trust assessment (analyses of career paths, publication output) would function as a skeleton for such a study. Assessing also the research substance would require a more case study oriented approach. An example of such an approach is provided by the next evaluation project.

In general, research impact assessment based on tracking research careers is well-founded, because a significant share of the knowledge diffusion between research organisations and other actors can be expected to take place through individuals. This is especially true for the diffusion of research methods and problem-solving skills based on a thorough understanding of a given field. Moreover, when the evaluation as-

sesses impacts that take place long after the project has ended, analysis based on individuals has important advantages over project-based evaluations: whereas project organisations usually dissolve shortly after the project has ended, individuals possess the skills acquired in the project throughout their life and thus facilitate tracking of research project's impacts.

2.3.7 Assessing Research Impact on Non-Academic Audiences

The following evaluation project discusses impact assessment in the context of social sciences. Impact assessment of basic research is not a straightforward endeavour in any scientific field due to the fluid definition of impact, the many forms that it takes, and the problems associated with its measurement. In the case of social sciences, however, the above mentioned issues complicate the assessment effort perhaps even more than in other scientific disciplines.

One of the actors that have been active in the improvement of impact evaluation methods for social sciences is the Economic and Social Research Council (ESRC) in the UK. The ESRC funds research and researcher training in the social and economic sciences with over £100 million per annum (About ESRC 2005). The ESRC employs a system of evaluation that systematically assesses the performance of research centres and programmes. One of the aspects being analysed in the research programme evaluations is interaction with the users of the research and the impact on these groups. To this end, the ESRC has pursued the development of new methodologies for assessing the impact of research on other than scientific actors.

In 1999, the ESRC commissioned a study that piloted new approaches for assessing the impact of research programmes on audiences outside the scientific community. It was conducted by the Science and Technology Research Unit (SPRU). The objective of the study was to "identify practical approaches to test systematically the impact of ESRC-funded research on non-academic audiences" and to determine for each approach (Molas-Gallart et al. 1999, p. 4):

- *"their appropriate level of application (Programme, Centres, themes, projects)*
- *their coverage required*
- *their feasibility and cost*
- *their applicability across different areas of research*
- *which approach or combination of approaches appears most fruitful"*

The evaluation team conducted three pilot studies labelled the 'user panel' assessment, 'networks and flows' and 'tracing post-research activity'. The approach adopted by the study team was based on a conceptual framework developed through a literature review of studies on research impacts. The evaluation team took special care to take into consideration the idiosyncrasies of social sciences as regards the assessment of research impact. The conceptual framework develops an all-encompassing model of impact mechanisms based on three categories: research outputs, diffusion channels and forms of impact. The pilot studies attempt to take into account all potential combinations of these three categories, thus covering all the possible mechanisms with which benefits from research are channelled to users outside the scientific community. The authors note that

"*When compared with the physical, engineering and medical sciences, the contribution of social science is likely to be more indirect, and more difficult to observe. Our approach emphasises the variety of ways, many of them indirect,*

in which socio-economic research affects economic performance and society at large, and studies their impact mainly in terms of the processes by which such impact takes place." (Molas-Gallart et al. 1999, p. 1)

Their notion of research impact is a logical continuation to this approach. The authors define research impact as occurring "whenever a research effort results in identifiable influences on current social, policy, and management practices" (Molas-Gallart et al. 1999, p. 4). Throughout the study there is an emphasis on the processes through which research influences users, partly because it is difficult to observe these two, process and the impact, in isolation.

The method labelled 'tracing post-research activity' is similar to that used by the Wellcome Trust. The tracing focused on the non-academic work carried out by the researchers in the ESRC's AIDS Programme after the research projects had ended. Scientists that had been active in the research projects were contacted and they were inquired whether they had taken employment, or provided consulting services to business enterprises or other stakeholder groups. The results showed that less than one third had moved to positions outside research organisations. The findings of the study indicated that the extent to which research skills, knowledge and methods developed in the projects had been applied in non-academic contexts was considerable.

The 'networks and flows' assessment was conducted simultaneously with the tracing of researchers, as information for both methods could be gathered in the same interviews. The 'networks and flows' method aimed at "mapping networks of researchers and relevant non-academic users/beneficiaries, and tracing the impact of research" (Molas Gallart 1999, p. 1). The approach attempted to distinguish between social networks existing already before the project, and networks that had been created during the research project. The identification of actors in the social networks was carried out through a 'snowball approach' consisting of three stages: first, researchers who were active in the AIDS programme were identified. Second, researchers indicated potential end-users of research results. Third, users identified in the second stage identified other users. The third step was included because the first set of users is likely to consist of people that follow research projects but do not actively use the generated results. Rather, they distribute the results in their organisations to people that may need such information. This methodology also provides information on the flows of knowledge that begin from research projects.

The 'user panel' study was conducted on another programme, namely the ESRC Innovation Programme. The panel consisted of individuals to whom the research was relevant and who were thus expected users of the results. Initially it was planned that the users would be contacted once every six months. This would have provided a way to investigate the patterns with which impacts are realised. However, it was quickly noticed that the strength of interaction between the users and the researchers was not intensive enough to warrant semi-annual interviews. It was decided that only one follow-on interview would be conducted at the end of the programme.

The findings of the study were mixed. Although there was little evidence of the direct application of tools developed in the research projects, the research had still influenced planning practices in the user organisations, especially by contributing to the develop-

ment of new strategies. Many users also claimed to have established new contacts through involvement in the programme, but the study team was unable to locate the source of this network creation. They note that "there was little evidence of new strong networks being formed as a consequence of the Programme." (Molas-Gallart et al. 1999, p. 57)

Based on the experiences from these three pilot studies, the evaluation team came to recommend a combination of 'networks and flows' and 'tracing post-research activity' approaches. Apparently, the user panel approach did not provide as much interesting information as the other methods, partly because the interaction between the users and researchers was not very strong. The implementation of the two other methods was also more economic because data collection could be conducted simultaneously. The methods provide a way to assess the diffusion of research results both through researcher mobility and through social networks.

There are also complicating factors associated with these methods. Notably, the attribution regarding the accumulation of researcher skills to a specific research project proved to be difficult. Nevertheless, as the study team notes, there are no better methods for evaluating such individual knowledge pools than assessments made by the researchers themselves.

The authors discuss the applicability of these methods to other levels of evaluation, and conclude that it may be possible to apply the methods to the levels of research centres, but not to individual projects initiated through a 'bottom-up' process. They base these conclusions on contextual factors that create boundaries for evaluation and thus facilitate the implementation of impact assessments. Such factors include focusing on a narrow research area, relevance for current policy problems, and a determined period of time for the research efforts. These aspects also support the formation of easily identifiable user communities.

2.3.8 A Composite Performance Rating System for the Advanced Technology Program

The Advanced Technology Program (ATP) of the National Institute of Standards and Technology fosters the development of innovative technologies by supporting collaborative R&D partnerships between industrial companies and research institutions. In 2004, total ATP project expenditures amounted to $270 million, of which the ATP's share was $155 million (ATP Awards… 2005).

The Composite Performance Rating System (CPRS) was developed as part of the evaluation programme of the NIST's Advanced Technology Programme. The evaluation programme was initiated to inform policy makers of the benefits and costs of the ATP, but later it assumed many additional roles, one of which was to improve existing methodologies for R&D evaluation. The evaluation programme consists of a diverse set of evaluation projects that assess various issues related to the ATP and employ a multitude of methods (see Ruegg & Feller 2003). Since the evaluation literature generated by the programme is quite expansive, it is impossible to present the evaluation projects in a comprehensive fashion within the scope of the current study. Thus, here we will focus on a single method being developed within the programme.

The Composite Performance Rating System was developed by Rosalie Ruegg from TIA Consulting Ltd. It addresses a need of programme managers to quickly assess the performance of projects at the portfolio level. The background of the

project as follows: In 1997, a system of assessment of finished R&D projects was implemented. It involves a short case description of each project approximately two years after the project has ended, called the 'status report'. The case description contains the following summary elements for all projects (Ruegg 2001, p. 13):

1. "a succinct project statement describing the technology and its applications;
2. project duration;
3. ATP project number;
4. project funding, with a breakout of ATP and industry's cost shares;
5. a broad accomplishment statement, followed by a list of major items of accomplishmen t– technical and business – including, where applicable, a listing of patents; commercial products or processes identified by name; technical and business awards received together with supporting information; new alliances formed and the purpose; capital raised; and an approximate number of publications and presentations;
6. an assessment of the current commercialisation status of the technology;
7. the analyst's brief assessment of outlook for continued pursuit of the technology;
8. development and commercialisation by the awardee(s) or its close collaborator(s);
9. company address and contact information for the project leader;
10. a list of joint venture participants and subcontractors;
11. for small companies, the number of employees at the project start and at the time the case study was performed;"

The status reports provided ATP programme managers both quantitative indicators and qualitative information on the performance of projects. However, they still could not easily assess what the overall performance of the portfolio of projects was like, because

Input	Knowledge creation	Knowledge dissemination	Commercialisation progress
Awards by outside organisations to recognise technical and scientific achievements	✓	✓	
Awards by outside organisations			✓
Patents filed by project participants, granted and not yet granted	✓	✓	
Publications and presentations by project participants	✓	✓	
Collaborative activity of project participants		✓	
Attraction of additional capital to take the technology further			✓
Employment gains by small-company award recipients			✓
Products and processes in the market or expected soon	✓	✓	✓
Analysts' outlook assessment for the technology in the hands of the project participants and their collaborators			✓

Table 2. (Adapted from Ruegg 2001, p. 16)

the indicators were not mutually comparable, their significance to the general performance of the project was not clear, and much of the information was in a qualitative form. This motivated the development on a composite indicator, which would summarise the information generated with the case studies.

The conceptual framework of the composite performance rating system draws on the general objectives of the ATP. The framework is based on three categories of indicators that match the general goals of the ATP and are consistent with innovation research. The indicator categories are (1) knowledge creation, (2) knowledge dissemination, and (3) commercialisation progress. These are shown in the columns in Table 2.

The inputs to these indicators, shown in the rows in Table 2, are weighed based on the importance of the indicator to each of the three categories. For instance, the contribution of the technical awards to the knowledge dissemination composite score is calculated by the following formula:

Weight contributed = 0.25 * square root (number of technical awards).

Similar calculations are made for each of the inputs and indicators. The check marks in Table 2 indicate which inputs contribute to each of the composite scores.

Summating and normalising the composite scores to provide a number between 0 and 5 yields the overall composite performance rating which is provided in the form of 0–4 stars. The approach has been tested on the first 50 projects of the ATP and the results show that the rating effectively separates high-performing projects (as defined by the method) from low-performing projects.

The merits of the method lie in its ability to summarise complex, qualitative information into an easily communicable rating. The example of the ATP composite performance ratings brings out a wider issue related to the use of composite indicators in general. The process of simplifying and compressing information into a composite indicator intentionally loses a lot of the information. For what purposes can these composite indicators be used? If the indicator is normalised somehow as in the case of the CPRS, the indicator can only provide information relative to something else, and therefore needs a frame of reference. Thus, the use of such indicators is limited to the portfolio level and trend analyses. It does not provide information on the performance of individual projects in isolation nor does it serve as a tool to communicate the performance of the programme to other stakeholders. It can, however, illustrate the distribution of projects as regards their performance. In the case of the CPRS, the rating system indicates that impacts of the R&D projects are quite unevenly distributed across the portfolio.

The indicators and their weighing are context specific, so if the method is applied to another setting, it needs to be "recalibrated". Probably changes in the composition of the indicators are also needed. This makes comparisons between different contexts, for instance scientific disciplines, impossible. The composite indicator methodology also requires instruments for gathering detailed information on different aspects of the project. Thus the method is somewhat laborious, especially if the monitoring systems are not already in place. Due to these limitations, the composite performance rating system is best applied to larger research programmes that have a long duration and sizeable projects.

2.3.9 Outcome Indicators at the Foundation for Research, Science and Technology (FRST)

The next evaluation project presents a performance measurement system implemented by the Foundation for Research, Science and Technology (FRST) for assessing outcomes of research. The FRST is a New Zealander funding organisation that invests public funds in both fundamental and applied research. The funds invested by FRST amount to approximately €260 million per annum. The foundation allocates its funding through a number of schemes that have been grouped according to the output classes specified in a New Zealand government's funding package called Vote: RS&T. The government's investment in Vote: RS&T is divided into 14 output classes. The performance measurement system focuses on six of these:

- Output Class 6: New Economy Research Fund (NERF)
- Output Class 7: Research for Industry
- Output Class 11: Maori Knowledge and Development Research
- Output Class 12: Health Research
- Output Class 13: Social Research
- Output Class 14: Environmental Research

The monitoring system consists of six indicators:

- Tangible benefits: number of new or improved products, processes and services for users;
- Third party revenue;
- Number of intellectual property;
- Number of reports, presentations and publications for users;
- Number of peer reviewed publications, awards and keynote presentations;
- Number of partnerships, contracts and linkages.

The indicators reflect the key objectives of the Foundation and thus provide information on how the Foundation is performing. They were developed in a consultative process together with the research organisations that provide the information. The information is gathered annually using standard forms. The FRST primarily funds portfolios of research projects, which raises the question of how outcomes from research efforts should be attributed to FRST investments. In this case, the outcomes are advised to be reported if FRST funding "significantly contributed" to outcomes. In order to determine a reasonable scope for this exercise, the attribution to FRST investment is limited to the previous ten years.

The information on outcome indicators was first collected in 2001, after which the system was refined in response to feedback from research organisations. Since 2002, the system has remained relatively stable.

Where possible, the indicator data is disaggregated to allow comparisons among different output classes and among groups of research users. The results indicate how the distribution of users in different types of stakeholder groups differs among programmes and, thus, provide indirect information on the types of benefits generated in the Foundation programmes.

Since the data from multiple years has been gathered in the same form, it is possible to investigate how the outcomes of FRST programmes have developed over time. At this time, however, the outcome measurement system has been operational for only four years, so not much information has yet been provided by the time series analysis. The value of stable indicators will grow in the future, as trends will be easier to identify.

The limitations of this measurement system are similar to those of the CPRS.

Indicating the outcomes of research efforts as a summated number, the indicators conceal a world of variety. The approach is oblivious to the variety of processes through which research influences society, because only the final results of the process are abstracted and counted.

To alleviate this problem, the FRST has conducted case studies on two of these indicators. The purpose of the case studies was to increase the understanding of the significance and relevance of the indicators by investigating how they are perceived in the research organisations that provide the information. The findings of these studies demonstrated that the meaning of the indicators to the research organisations may vary significantly depending on the research programme (and thus the research area) (Report on… 2004, Webber 2003). This indicates that care needs to be taken when comparisons among the research programmes are conducted.

2.3.10 Standard Evaluation Protocol 2003– 009 For Public Research Organisations

During recent years, discussion on the social relevance of public sector research has been particularly heightened in the Netherlands. In 2003, the organisations participating in the evaluation of public sector research, Dutch universities (Vereniging van Universiteiten VSNU), the Netherlands Organisation for Scientific Research (NWO) and the Royal Netherlands Academy of Arts and Sciences (KNAW), defined a new protocol for research evaluation. 'Standard Evaluation Protocol 2003–2009 For Public Research Organisations' is a follow-up report of the working group 'Kwaliteitszorg Wetenschappelijk Onderzoek' (Quality Assurance Scientific Research). The purpose of the protocol is to improve and unify national evaluation activities. The process resulted in a new, national system for evaluation.

The objectives of the evaluation system address a much broader set of issues than mere impacts of research. The objectives of the evaluation system are (Standard… 2003, p. 5):

- *"Improvement of the quality of research through an assessment carried out according to international standards of quality and relevance;*
- *Improvement of research management and leadership;*
- *Accountability to higher levels of the research organisations and funding agencies, government, and society at large."*

The evaluation system addresses both past performance and future plans. The evaluations attempt to answer the following questions (Standard… 2003, p. 11):

"For past performance:

1. *What are the quality and relevance of the institute?*
2. *What is the quality of the leadership, management, strategy and research programmes of the institute, its (human) resources, organisation and infrastructure and how can they be improved?*
3. *To what extent has the institute/research programme achieved its mission and goals formulated for the period under review?*

For future plans:

1. *Is the mission of the institute well chosen and phrased in view of the actual developments in the relevant research field(s)?*
2. *How do you assess the institute's research plans and is there sufficient coherence in the research portfolio of the institute?*
3. *What is the quality of the leadership, management and strategy of the institute, its (human) resources, organisati-*

on and infrastructure and how can they be improved?

4. *Which of these aspects has room for improvement and how could that be accomplished?"*

Notably, the ex-post assessment of research impact does not play a very significant role in the overall evaluation. Investigation of research impact as it is understood in the current study provides an answer to the first of the questions, and generates background information to support the analysis regarding the other evaluation questions.

One of the guiding principles of the evaluation system is its ease to the respondents. Consistent with the protocol, self-evaluations are carried out once every three years and an external evaluation once every six years. A yearly monitoring system based on a national information system is being devised to store all relevant data for these evaluations. The external evaluation is based on the self-evaluation and conducted through site visits during which the evaluation committee interviews key staff at the institute.

The assessment addresses both the level of research institutes and the level of research groups/programmes within institutes. The extents to which research impacts are addressed differ between the two levels (Standard… 2003, p. 9.):

"The assessment criteria for an institute as a whole and those for the research programmes are similar, but differ in scope and depth. The institute assessment puts emphasis on strategy and organisational aspects, whereas the programme assessments focus on the results and quality of the scientific research and on the future."

Similarly to the FRST outcome indicators, the data for impact assessment is gathered with standard templates from research organisations. However, despite its explicit focus on social relevance, the Dutch system focuses mostly on direct outputs from research, and assesses exploitation of the research results only in a qualitative way. The data reported by the public research organisations include:

- Researchers and other personnel in the previous six years
- Resources, funding and facilities in the previous six years
- Academic reputation (bibliometric analyses, prizes, etc.)
- Aggregated results of the institute in the previous six years (publications, monographs, PhD theses etc.)
- Other commendable results (patents, awards, etc.)
- Overview of the results (selection of three to five publications that represent the quality and impact of the research, a numerical overview of the results, a full list of publications)

The fourth and the fifth of these indicators are reported only at the institute level, others are reported for both the institute and each individual research programme. In addition to these indicators, the respondents submit qualitative self-evaluations on research strategy processes, internal and external collaboration as well as on effects of collaboration and dissemination of research results outside the scientific community.

Congruent with the philosophy of the evaluation system, the data is relatively easy to provide, as few subjective assessments need to be made. The disadvantage of this system is its cursory treatment of the influence on business enterprises, policy makers and other stakeholder groups which, in fact, is the central issue as regards the assessment of societal relevance.

As a response to this, the Council for Medical Sciences of the Royal Netherlands Academy of Arts and Sciences

undertook an exercise for complementing the protocol presented above with methods that focus on the societal impacts. The report produced by the project is called "The Societal Impact of Applied Health Research: Towards a Quality Assessment System". The Council notes that evaluation of societal impact is particularly important for applied health research since research institutes in the field also have a strong societal mission in addition to a scientific one. The council members argue that (The Societal… 2002 p. 10):

"Applied health research differs from 'fundamental' (bio)medical research in its dual mission, which is both scientific and societal: it is explicitly concerned not only with the acquisition of scientific knowledge as such but also with the usefulness and implementation of scientific achievements. So, relevance to society is an important explicit objective of applied health research and evaluation should, therefore, not be restricted to scientific quality. A formal evaluation would acknowledge the importance of societal impact of the research at issue."

The Council suggests that evaluation activities focusing on societal impact should be included in the new national evaluation system and provides suggestions for complementary methods as to how this is to be achieved. While some of the suggested methods involve general quantitative indicators, the Council notes that the mission of the research institute needs to be taken into consideration when evaluating the societal impact. The mission of the institute influences the extent to which societal quality should be evaluated and serves as a basis against which the impact can be assessed. This implies that mere quantitative indicators are not sufficient for the analysis, and that qualitative investigations into the ways in which research is used needs to be conducted. The methods suggested include:

- Studies on how research results are incorporated in medical guidelines and protocols
- Assessment of contributions to initial and post-initial education
- Studies on improvement of health care technologies
- Evaluating the process by which research results are implemented to solve specific societal problems

The response of the Council for medical sciences highlights the need for qualitative impact assessments, and that the need may vary depending on the research area. Although assessment of societal quality – which the council sees as the main element of the evaluation system outlined by the national evaluation protocol – is important in its own right, it is not sufficient to judge whether the research efforts of an organisation have been effective.

2.3.11 PSA Target Metrics 2004

The remaining two reports discuss the measurement of outcomes of research from the perspective of the national innovation system and its performance, rather than from that of a research funding organisation assessing the impacts of its investments. However, the indicators presented are very similar across these two levels and, thus, indicators used for both purposes are relevant to the current study.

The first set of indicators is employed by the Office of Science and Technology (OST) to demonstrate the effectiveness of research funding and to investigate the performance of the UK research system in relation to other nations. The indicator system was developed to "support a system for assessing outputs, outcomes and impacts related to the Public Service Agreement (PSA) target to improve the relative inter-

national performance of the UK research base" (PSA Target Metrics 2004[2]). The selection of indicators was developed by Evidence Ltd. The indicators were selected so that they would illustrate dynamism and change in the research system, an element which was partly overlooked by the previous set of indicators.

The indicators are based on data from two sources, the OECD and Thomson ISI. The majority of the OECD data is gathered from "Main Science and Technology Indicators" (MSTI, 2003–2 second edition) and "Research and Development Statistics" (RDS, 2003 edition). There are altogether 37 indicators which are grouped into seven categories. The categories are:

- Theme 1 – Inputs (including expenditure on research);
- Theme 2 – Outputs (including people and publications);
- Theme 3 – Outcomes (research recognition, citations; training and research quality);
- Theme 4 – Productivity – financial (outputs and outcomes related to inputs);
- Theme 5 – Productivity – labour (outputs and outcomes related to other measures);
- Theme 6 – People;
- Theme 7 – Business Expenditure;

To provide an understanding of the performance of the UK research base, the relative and absolute values provided by these indicators are set in context by comparing them to the values of competing countries. For each country the values are typically provided for the years 1994–2002, so that changes in the relative positions of countries can be seen. In addition, for many indicators the data is disaggregated to reveal differences at the level of scientific disciplines.

Because the multiplicity of indicators makes it difficult to assess the performance of the research base in each country, Evidence Ltd has selected six indicators to form a graph they call the "Research Footprint®". The indicators selected to form this graphical presentation are:

- Publicly performed R&D as proportion of GDP (theme 1)
- Number and share of OECD PhD awards (theme 2)
- Number and share of world publications (theme 2)
- Number and share of world citations (theme 2)
- Rank of share of world citations by nine main research fields – frequency in top 3 rankings (theme 3)
- Researchers per thousand workforce (theme 6)

It is apparent that the indicator selection emphasises theme 2 – scientific outputs.

The main finding produced by the system of PSA target metrics is that in 2004, the UK research base is performing relatively strongly in international comparison (PSA target… 2004, p. 10):

"*The UK is strongest overall in the natural sciences, and on many indicators is second only to the USA. Although the UK has been overtaken by other nations in some areas, it sustains a more consistent performance across fields than those countries. The UK's strong international performance has been achieved with lower average investment compared to its competitors and with a relatively lower availability of people with research training and skills.*"

2 Source: Office of Science and Technology, PSA target metrics 2004. Crown copyright material is reproduced with the permission of the Controller of HMSO.

2.3.12 The Advisory Council on Science and Technology, Taking Stock of University & College Commercialization Efforts

The Advisory Council on Science and Technology, Canada (ACST) is a federal advisory council with a mandate to provide the Prime Minister with policy advice regarding the development of science, technology and innovation in Canada. "The Council's role is to review Canada's performance in research and innovation, identify emerging issues of national concern, and advise on a forward-looking agenda with a view to positioning Canada in an international context" (ACST… 2005).

In a recent policy stocktaking, the ACST assessed Canadian policies for supporting commercialisation of research conducted in universities and federal laboratories. The purpose of the exercise was to increase the productivity of public funding spent on research by enhancing the utilisation of research. As part of this exercise, the ACST reviewed the current commercialisation performance of Canadian universities. A background report addressing this issue, called "Taking stock of R&D across three sectors. Taking stock of university & commercialisation efforts: Advising on the way forward" was prepared by Dr. Chris Riddle (Taking… 2003).

The report addresses two distinct issues. First, the Canadian investments into R&D in the university, industry and government sectors are reviewed. Second, research commercialisation performance in universities and colleges is discussed. For both themes, national-level indicators are presented. The report also provides an overview of the main policy levers employed by the federal government to influence both R&D funding as well as research commercialisation. Best practices, success stories and issues for further development are highlighted.

The review is based on a number of documented reports on Canada's performance. The majority of data used by the report is provided by Statistics Canada and the Association of University Technology Managers, Inc. No primary data collection has been undertaken as part of the review.

The R&D investment indicators examined in the review include various measures of gross domestic expenditures on R&D (GERD), higher education expenditure on R&D (HERD) and business expenditure on R&D (BERD). The sources of R&D funding as well as R&D expenditures in different sectors are analysed. The commercialisation indicators, in turn, include following types of indicators:

- Number of different type of IPR
- Income from IPR licensing
- Number of start-up companies
- Equity value of start-up companies
- Efficiency measures (above indicators in relation to the research expenditures)

In most of the analyses, the indicator scores are interpreted in international context, and with respect to development trends (Taking… 2003, p. 12):

"Taking stock of Canada's R&D performance cannot be done in isolation. Absolute increases in R&D activity do not translate into competitive advantage if other countries are increasing their performance at an even greater rate. Furthermore interdependencies are growing in the increasingly global economy and it is important to take stock of Canada's R&D performance in a global context. There is evidence to suggest that Canada has some particular strengths at the international level and these may provide opportunities for enhancing Canada's competitive advantage if better understood and exploited."

3 Themes in the Assessment of Research Impacts

This chapter discusses the assessment of research impacts in relation to specific themes found in the analysed material. The thematic areas are structured to reflect different stages in the process through which societal impacts from research funding are realised. To a certain extent, the themes also reflect current priority areas in innovation policy. For each of the thematic areas identified, methodological approaches employed in the analysed material are discussed. The chapter begins by developing a conceptual framework, which sets the course for subsequent discussion.

3.1 Conceptual Framework

The purpose of the conceptual framework is to facilitate the analysis of evaluation projects by providing a basis for structured analysis. The framework was not developed prior to the analysis of individual evaluation projects but, rather, the framework emerged during the data collection from individual projects, based on the thematic areas that could be identified in the analysed material. As a result, the framework is data-based, and it reflects the themes which have been addressed in the evaluation projects in order to characterise the impacts of research funding organisations. The issues addressed, or "themes of impact assessment" as they are labelled here, can be grouped into four broad stages:

1. Impact of research funding allocation on the operations of research organisations;
2. Impact of funding through conducted research activities on science, researchers and the scientific community at large;
3. Impact of funding on the interaction and knowledge transfer between research organisations and other types of organisations such as business enterprises, government agencies, etc.;
4. Impact of funding through utilisation of research on other than research organisations.

This framework is illustrated in Figure 1. Three types of actors are identified: research funders, research performers and research users. Impacts from research are considered to materialise through the activities and interactions of these actors.

Because utilisation of research findings in other than research organisations is one of the major focus areas in the present study, the framework emphasises the conceptual distinction between research providers and research users. Below in the analysis, this distinction is also referred to by categorising research impacts as belonging either to the domain of science and scientific community or to the domain external to the scientific community. Furthermore, the same idea is represented by distinguishing between academic audiences and non-academic audiences for research. In reality, categorising research utilisation as an interplay of research organisations and external users is an oversimplification. The majority of knowledge produced by basic research is targeted to, and utilised by, other researchers within the scientific discipline. However, the distinction is useful for the purposes of the present study and thus adopted as one of the points of departure.

The four stages attempt to cover all the themes that have been addressed in

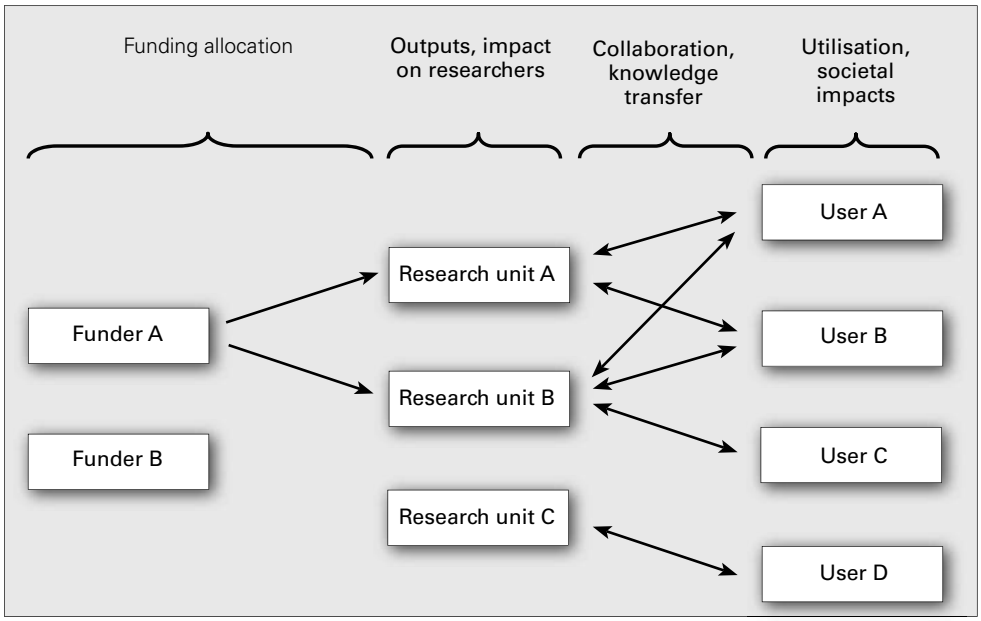

Figure 1. Conceptual framework.

the analysed evaluation projects to characterise the impacts from research funding. The first stage, allocation of funding, covers those areas of the analysed evaluation projects that attempt to characterise the impacts of a funding organisation by describing how it allocates its funding. This includes descriptions of the volumes and patterns in funding allocation, as well as characterisations of the types of research projects and organisations that are likely to receive funding.

The second stage, research outputs and impacts in the scientific community, covers impacts from research funding to the activities of research performers (e.g. resources), scientific outputs from research (e.g. publications) as well as the impacts of research efforts on the research performers themselves (e.g. changes in the future opportunities of researchers). The second stage also covers the activities and outputs "produced" by the researcher that transfer knowledge from the researcher to the user (e.g. publications directed for users, patents, consulting, and services for users). Here these are collectively referred to as "user outputs".

The third stage, collaboration and knowledge transfer, focuses on the methods used to assess collaboration between researchers and end-users of research. The analyses of collaboration between research organisations have also been included in this category, because the methods used to analyse the collaborations are similar to the ones used in the assessment of researcher-user collaboration.

The fourth stage, utilisation of research and resulting impacts, covers methods that have been employed to characterise the influences and impacts that basic research has on the end-users of research, and on society at large.

The domains addressed by the stages overlap, so strict distinctions between the themes cannot be made. Notably, the distinction between researcher-user collaboration and impacts on research users can be difficult to make and even complicate the attempts to understand the impacts from research, as it generates a

mechanistic and simplified view of the processes involved. Because indicators focusing on individual aspects of research impact play a relatively large role in the present study, this distinction has, nevertheless, been made.

The next sections discuss the methods and indicators used to address each of these stages in more detail. For each section the discussion follows a similar structure. First, the significance of the topic in impact assessment is discussed. Second, the main themes discussed in the analysed evaluation projects are presented. The emphasis is on the presentation of methods employed to investigate the issues at hand.

3.2 Allocation of Funding

Although portfolio analyses of funding patterns do not provide evidence of research impacts as such, from the perspective of a funding organisation they provide important information on how funding is allocated and, thus, on the direct impacts the activities of the funding organisation have on the research community. There are a number of advantages to characterising impacts of the funding organisation based on the way it provides grants.

Because the indicators provide information directly on the operations of the funding organisation, the availability of data is usually not a problem. Data describing funding patterns can be gathered from the databases of the funding organisation and supplemented with statistical data from other sources. Moreover, in the context of impact assessment, this data is relatively reliable. The information is not diluted by the influence of other, external factors.

Since the information provided by the indicators addresses issues which result from direct actions of the funding organisation, the organisation can also change and configure its operations based on the findings. In contrast, it would be difficult to base decisions on the operations of a funding organisation solely on information gathered from assessments of research exploitation, because the links between the actions of a funding organisation and the research impact are not clear. Of all the four stages described by the conceptual framework, allocation of funding is the stage on which the funding organisation has most influence, and through which it can configure its operations.

The other side of the coin is that portfolio analyses on funding patterns do not tell us much about the actual impacts realised. They are unable to provide answers to the question on whether the allocated funding has supported the attainment of any of the societal goals it has been assigned to. The approach does not measure research impact but, rather, assumes there is a link between the funding allocated and the goals which it attempts to achieve. However, analyses on funding patterns can tell us about the relevance of conducted research and, perhaps more importantly, they serve as a basis for the monitoring of programmes' activities.

In the analysed material, indicators on the funding allocation and the information provided by them have mainly six types of functions. The first is related to the issue discussed above. The indicators provide a description of the inputs to the research system, whether it is at the national, programme, or project level. The second has to do with providing a context for assessing the significance of a given type of funding: the volume of a given type of funding is viewed in comparison to other funding sources. Third, the descriptions of funding allocation provide indirect measures of the rele-

vance of research (This is discussed in Section 3.5.1). Fourth, the indicators are used to characterise the organisations, projects and grant applications for which the funding is allocated. This serves to describe how the funding is "directed" within the population of potential recipients. Fifth, the indicators are used to provide descriptive statistics as the basis for more detailed analyses. For instance, analyses of approved proposals are usually preceded by more general descriptions of how funding is allocated among scientific disciplines. (This is not discussed further as it does not provide insights for impact assessment.) Sixth, there are methods which attempt to assess the additionality of funding, i.e. the increase and changes in research activity brought about by the received grant.

3.2.1 Inputs to the Research Base

Two of the analysed evaluation projects, the Office of Science and Technology's (UK) Public Service Agreement (PSA) Target Metrics 2004 and the Advisory Council of Science and Technology's (Canada) Taking Stock of University and College Commercialization Efforts, discuss inputs to the research base on the national level. The indicators are based on OECD data and thus provide few new insights as regards the development of the methodology of impact assessment. The indicators are based on gross expenditure on R&D (GERD), R&D performed in the higher education sector (HERD) or a similar all-encompassing indicator of research intensity in the economy. To provide a more detailed picture of the funding patterns at the level of the research system, these primary indicators are broken down into smaller constituents. The indicators provide either information on the funding sources which comprise the total expenditures or on the sectors which perform the research. The results are disaggregated to the level of different sectors (universities, business, government).

The results of the indicators are presented in comparison to a comparator group of countries with similar economies. Time series are presented to provide an understanding of the change in positions of each country.

3.2.2 Significance of a Granting Organisation's Funding

Two of the projects, the impact evaluation of the Austrian Science Fund (FWF) and the Funding Ranking of the German Research Foundation (DFG), provide a description of the funding sources at the level of individual funding organisations. The results of these indicators have more relevance to a granting organisation, because it can use the information to demonstrate its significance in the activity of its clients. Also, the indicators serve as a proxy measure for the extent to which research results produced by the clients can be attributed to the given funding organisation.

The indicator may take many forms depending on its use and the available data. In the analysed projects, the main approach is to calculate the share of external funding of total funding in universities, and then to demonstrate the share of the funding organisation of the external funding. The DFG funding ranking provides an additional, and a somewhat more elaborated, way to assess how funding allocated by a single funding organisation and the total external funding are related. DFG approvals and third party funding income are measured for each university and then plotted in a scatter diagram. The diagram shows that there is a close correlation between the total external funding and grants provided by the DFG.

3.2.3 Characterisation of Funding Recipients

Characterisation of funding recipients has a much more prominent role in the analysed evaluation projects than the two purposes discussed earlier. Characterisation of funding recipients is done with three purposes in mind. The first of these has to do with describing how the organisation allocates its funds among the population of potential recipients. The second has to do with explaining what influences the generation of outputs and impacts. This is mainly done at the project level. The third is the least significant of the three, and concerns the calculation of inputs to projects to assess the efficiency (or performance) of research in the organisation.

The characterisation of grant recipients can be accomplished with various means. The attributes of a potential funding recipient can be based on personal characteristics, the attributes of the application, the attributes of the research project or the attributes of the research organisation. In the analysed material, there are examples of all these approaches. The analysis is carried out at the level of grant applications, individual projects or organisations applying for funding.

Beginning from individuals applying for funding, the applicants have been characterised on the basis of their age, earlier experience on the application process, type of profession and sex. The data has usually been gathered from the database of the granting organisation.

As for the applications, the characterisation has addressed issues such as their level of elaboration, resources available for generating the application and the main field of science indicated. The understanding of the application process as well as of the projects' attributes has been formed based on data from surveys of applicants.

The projects, in turn, have been characterised in a more qualitative fashion. The measured attributes of projects have to do with motivations for undertaking the project, project objectives, project content, interdisciplinarity and the significance of the project for the applying institution. The measured attributes are always specified in a way which reflects the specific challenges of the given evaluation.

Finally, organisations are characterised based on their type, research area, geographic location, their performance in comparison to other organisations, etc.

The analysed material provides few examples of methods which can be used to analyse the different attributes listed above in order to characterise what types of actors are funded by granting organisations. The simplest approach is to ask the applicants directly to characterise their project or institution. To provide information on the differences of approved and rejected project proposals, the same questions regarding the attributes of the projects should be targeted at both, applications that were successful as well as applications that failed to receive a grant. The FWF impact evaluation provides an example of such an approach.

Another approach used by the evaluation team for the assessment of the FWF's impacts was to create a regression model of project approval. The model was used to investigate whether the decision to approve was influenced by the characteristics of the coordinator and the application in question. The analysis indicated that the approval process did not favour some applicants over another and that the decisions to approve a project were based mostly on verbal assessments of applications.

The DFG funding ranking in its entirety can be interpreted as a large-scale attempt to characterise the types of organisations funded by the organisation. From this perspective, the main finding of the evaluation is that DFG funding is granted to organisations of high academic performance, whether performance be measured in terms of outputs, centrality in research networks, participation in international programmes or number of DFG proposal reviewers.

3.2.4 Additionality

There are only few attempts to measure additionality among the issues addressed by the analysed evaluation projects. A simple way exemplified by the evaluation of the impacts of the EU Framework Programmes to the UK is to request the funding recipients to assess what would have happened in the absence of funding. The results of this question indicated that additionality impacts did take place, but as this is the opinion of the funding recipients, it may well be biased. Another, complementing way is to ask the individuals who did not receive funding whether the forms in which the project was carried out were affected by the rejection. This method was employed in the analysis of the impacts from the operations of the Austrian Science Fund. The findings indicate that the majority of the proposals were not carried out at all, and most of the remaining projects were carried out in a slightly modified fashion.

3.3 Research Outputs, Outcomes and Impacts in the Scientific Community

This section reviews the methods employed by the twelve evaluation projects to assess the impact of research as demonstrated by the outcomes of basic research. As described in Section 3.1, this category covers indicators which provide information on the impacts on the scientific community itself and on the direct outputs targeted to research users. The methods discussed either provide information on the outcomes of research activity in general, or attempt to relate the outcomes to a specific type of research funding.

The assessment of publication patterns generated by scientific research is the prevalent method in the evaluation of impacts from basic research. The reason for this has to do with the central role of publications in the organisation of scientific activity: publications are the main channel through which research results are disseminated in the scientific community, and they are also an important element in the scientific reward system. In addition, publication output can be measured in a quantified form and the attribution of publications to individuals, research projects or organisations is relatively straightforward. However, typical bibliometric methods are not well represented in the analysed material since they were not emphasised in the selection. Instead, the emphasis is on methods that attempt to measure the influence of research funding on the generation of outputs or on methods that focus on more qualitative aspects of research impact. Examples include publications targeted to non-academic audiences, services offered by researchers, and impacts on the capabilities and skills of researchers.

However, even with this shift towards qualitative assessment of impacts on users, the indicators covered here do not directly measure the influence of basic research on other actors in society. If conclusions about the impact of research on non-academic audiences are drawn, they rest on the assumption that the out-

comes from research are relevant to, and that they are exploited by, end-users.

3.3.1 Scientific Outputs

The starting point for most of the analyses employed in the analysed evaluation projects is a mere calculation of research outputs. These are employed at all levels of assessment from the level of individuals to the level of national research systems. The indicators mainly focus on publication output, citations, and the number of new, qualified research personnel. Depending on the needs of the analysis, these are disaggregated to yield distributions among scientific disciplines, funding schemes, organisations or individuals. Also secondary indicators measuring efficiency in the production of outputs are possible, as demonstrated by the PSA Target Metrics 2004. Some of the indicators used employ standardisation techniques, but this is usually limited to normalisation by the size of the analysed unit measured as a number of researchers, and by the size of the inputs, i.e. investments in the research efforts.

More interesting methods for the assessment of outputs from research include those that attempt to relate the outputs to the patterns of research funding. The impact assessment of the FWF involved a method to assess the influence of the source of funding on publication activity. The method was based on regression models that used the number of different types of outputs generated by university research institutes as the dependent variables. The independent variables consisted of the research areas of these institutes and the sources of the institutes' funding. The analysis allows comparisons to be made as regards the influence of different funding types on the outputs generated by these institutes. The results indicate that, in comparison to other sources of funding, funding channelled through the FWF results in relatively many outputs. Unfortunately, this method is limited by the availability of data. In the case of the FWF, a unique database with both the publications and sources of funding for approximately 1,000 university institutes was available.

Another approach used by the FWF evaluation team focuses on the project level. Its main purpose is to assess the differences in output generation among scientific disciplines. The average number of different types of publications, presentations and theses is calculated and they are presented for each of the six main scientific disciplines. An analysis of variance indicates that for most of the output types there are statistically significant differences among the disciplines. For example, natural sciences are strong in scientific publications whereas social sciences have generated the most lectures held in conferences. This is also an indirect indication of the types of impact generated by research in different disciplines.

The DFG funding ranking used two existing bibliometric studies to demonstrate that the grants were allocated to high-performing organisations. Reversing this logic, a similar methodology could be used to assess the impact of DFG grants on the generation of scientific outputs.

Instead of a quantitative assessment of research outputs, also more qualitative approaches can be employed. One of the elements required by the new evaluation protocol in the Netherlands is a list of three to five publications that represent the quality and impact of the research. A qualitative assessment of the most important publications conducted by a peer review committee emphasises quality of impact more than quantity.

Another approach with a more qualitative orientation is the one conducted

by the Wellcome Trust. The Trust conducted an analysis of the careers of the most productive members of the Prize Studentship cohort. The objective of this exercise was to understand better who the productive scientists are and how their career has progressed. The results show that of fourteen individuals selected for further analysis, twelve were men and two were women. They contributed just above one third of all publications produced by the cohort. At the time of the assessment, five of the individuals were working outside the UK. Six of them had received grants from the Trust later in their career.

3.3.2 User Outputs

Moving from scientific outputs to research outputs that are primarily targeted to non-academic users, the assessment or research results begins to involve increasing challenges as regards the quantification of the results. The challenges are created by the ambiguous definition of the observables and the associated attribution problems. Here the term "user output" is used broadly to designate the activities and outputs generated or produced by the researcher who transfers knowledge from the researcher to the user. User outputs include:

- Publications written for non-academic audiences
- Presentations, lectures and discussion meetings regarding research results
- Patents
- Products, services and consulting

Some of these output types are quite commensurable, but for others, e.g. the generation of products, it is difficult to measure how many products have been produced. What counts as a single product? What is the effect of a research project in the generation of this product? Has funding from a granting organisation contributed to the development of the product? These are all relevant questions when assessing the impact of research with measurements of user outputs.

Despite these challenges, the measurement of user outputs is an important indicator of the relevance of research and of the knowledge transfer between academia and society at large. Its benefit is the position on the boundary between academic research projects and end-users of research: user outputs serve as a proxy measure for impact on actors outside academia, but yet the data for its measurement can be collected from researchers who have received a grant. Particularly in assessments with a broad coverage, this approach is considerably simpler to implement than an assessment focusing on potential end-users of research: requesting potential end-users of research to indicate the number of products or services they have exploited would not be a feasible approach unless the scope of the exercise is very narrow and there are only few potential users.

In the analysed material, the assessments of user outputs focus either on individual projects or research organisations. The data is gathered either with surveys or by means of annual reporting systems. The outcome indicator system developed by the FRST is an example of a well-designed system for collecting data on user outputs. The data is collected from research organisations annually. The FRST collects information regarding the user outputs with three indicators:

- Tangible benefits: New or improved products, processes and services for users;
- Intellectual property;
- Number of reports, presentations and publications for users.

For each of the outputs recorded, the funding scheme which contributed

to the output is also indicated. For tangible benefits as well as reports, presentations and publications, the user type is indicated. Potential user types are (Outcome… 2004, p. 34):

- *"New Zealander government agency*
- *New Zealander local authority and/or regional council*
- *Maori organisation*
- *New Zealander community and/or voluntary group*
- *Private New Zealander business*
- *Overseas organisation (excluding research organisations)*
- *Researcher or research organisation"*

This additional information provides the FRST the possibility of comparing the impact of funding schemes on the user groups and provides indication of the types of impact mechanisms that the funding schemes have.

Perhaps more importantly, the FRST also conducts case studies to increase the understanding of the significance of these indicators. The case studies serve to validate the indicators, generate information for improving the indicator system, and create a more detailed picture of the variance behind these indicators (Webber 2003, Report on… 2004). Thus, the case study approach attempts to mitigate the challenges in user output measurement described earlier.

3.3.3 Impacts on the Research Performer

Tangible outputs capture only one facet of the resulting benefits from basic research. Assessment of outputs emphasises the benefits of research on a short-term basis, overlooking the impact of research projects on the capabilities and skills of researchers that increase the capacity for conducting research in the long run. For these reasons, some of the analysed evaluations have addressed the impacts that research activities have had on the research organisations or individuals performing the research.

The methods used to address this question are limited. They all involve self-assessment by the researchers themselves, carried out either with questionnaire surveys or interviews.

At the level of individual researchers, the assessment focuses on the influence of research projects on researchers' position in the organisation, their contact network, their skills, or their capacity to use new techniques and methods. At the level of research organisations, the assessments have addressed issues such as the research projects influence on generating follow-up projects or the influence on research organisations capabilities and competitive position.

To better understand the sources of these benefits, it may be possible to analyse the associations between benefits and characteristics of the projects. This is an approach exemplified by the evaluation of the impacts of the EU Framework Programmes in the UK. The evaluation team requested the respondent to assess whether benefits associated with the participation in the programme outweighed costs involved, and then used this as a proxy measure for project success. Next, exploratory correlation and regression analyses were carried out to find out what aspects influenced the success. The analysis suggested that projects are most successful when they are aligned with the participant organisations' internal objectives. Contrary to expectations, the results of the analysis indicated that very large projects are as likely to succeed as smaller project despite the potential coordination problems among their multiple. An interesting finding is also the fact that cost-benefit ratios did not correlate significantly with project outputs but, rather, with the impact of the project on the performing

organisation. The evaluation team concludes that "These correlations suggest, reassuringly, that participants value qualitative impact more than they value numbers of project outputs" (The Impact… 2004, p. 32). While this method may provide new insights as regards the determinants of project success in a particular programme, the significance of the findings is decreased by the exploratory nature of the method and by the unreliability associated with the measurement of the employed variables.

3.4 Collaboration in Research

In the evaluation projects analysed, collaboration between research organisations as well as between research organisations and end-users of research seems to be assessed primarily for three reasons. First, research collaboration can generate externalities in the form of knowledge spillovers from one cooperation partner to another, and are thus important in their own right. Second, adopting a long-term perspective, research collaborations can be viewed as links formed in the innovation system. The networks created by cooperation are seen as a societal impact (social capital) that increases the capacity of the research system and influences its long-term performance. Third, the assessment of research collaborations may be easier to conduct than assessments of actual impacts of research. Thus, research collaborations are used as a proxy measure of the impact that basic research has on the society. Fourth, it is acknowledged that impacts are abstractions which are artificially separated from the research process, a process that may include cooperation as an essential element of the research effort. From this perspective, research collaborations are an inseparable part of the impact.

Below, the issues addressed by the analysed literature are grouped into three themes. The first of the themes is associated with research cooperation within the scientific community. The second concerns the formation of international links in the research system. Finally, collaborations bridging the "boundary" between research organisations and non-academic actors (industry, policy makers, etc.) are discussed.

3.4.1 Scientific Collaborations

Of the evaluation projects analysed, only one, i.e. the DFG funding ranking, addresses the formation of networks between research institutions in a detailed fashion. The method employed by the German Research Foundation is based on an analysis of four types of DFG-coordinated research programmes (Research Units, Research Training Groups, Collaborative Research Centres and Priority Programmes) with participation by at least two research institutions. In total, there are 489 such coordinated programmes involving multiple institutions.

The analysis relies on two qualitatively different approaches. The first investigates the extent to which research institutes collaborate with other institutes. This is achieved by simply counting the number of participations and the number of partner institutions in DFG-coordinated programmes. The outcome of this exercise is a list of institutes ranked on the basis of both indicators. The results are disaggregated to shed light on the differences between types of DFG-coordinated programmes and between different scientific disciplines. Examples of the findings provided by the analysis are (Funding ranking… 2003, p. 59, 62):

"The individual lists display a relatively high level of agreement: Six of the institutions with the highest participation

rates in Collaborative Research Centres are also amongst the "top ten" for participation in Priority Programmes, with the two universities in Munich leading both of these lists."

"Coordinated programmes involving multiple institutions are of notably above-average importance in research areas belonging to the engineering sciences as well as in 'biology' and in 'medicine'."

The second approach provides an overview of the cooperation structures among research institutes by visualisation techniques. Visualisations of network structure are provided for the complete network and for main fields of science. The method provides an intuitive visualisation of complex data which can be used as a basis for qualitative assessments on the role and significance of individual institutes within a research area. Also the strength of relationships between individual institutes is easily observed. In addition, the method provides insights into regional cooperation structures.

However, it is difficult to analyse the extent to which collaborative programmes have contributed to the strengthening of research networks. Nevertheless, a conclusion regarding causal influences is drawn in the DFG funding ranking (Funding ranking… 2003, p. 72):

"This look at a total of five core network structures based upon the various cooperative relationships between universities and non-university institutions arising through the coordinated programmes operated by the DFG has made clear that a particular goal of these programmes, the establishment of cooperative structures, is indeed achieved. Relationships between institutions arise according to different criteria for each scientific discipline. As well as various topical and subject-specific aspects which have not been looked at in depth here, factors of location play a not insignificant role. The potential of regional structures, for instance neighbouring universities and non-university research institutes, are utilized to a varying degree in each scientific discipline."

3.4.2 International Cooperation

The formation of international contacts in national research systems has become an increasingly prominent theme in innovation policy. It is a theme also addressed by two of the analysed evaluation projects, the DFG funding ranking and the evaluation carried out by the Wellcome Trust (Despite its focus on an international programme, the purpose of the assessment of impacts from the EU Framework Programmes is to assess the benefits to UK-based participants rather than to measure the formation of international links). Due to its focus on indicators of academic performance, the emphasis of the DFG evaluation is on "inbound" internationalisation, i.e. international mobility of researchers from abroad to German research institutes. In contrast, the Wellcome Trust evaluation emphasises international mobility of UK researchers.

As part of the indicator selection, the DFG funding ranking evaluates the performance of German universities and research institutes on the basis of their attractiveness to foreign researchers. The analysis employs data on the distribution of foreign researchers between German host universities and other institutes. The data is provided by the Alexander von Humboldt (AvH) Foundation and the German Academic Exchange Service (DAAD). In addition to descriptive statistics regarding distribution, the DFG assessment provides a ranking of universities based on the total number of

visiting researchers by scientific discipline. The link to the DFG funding is established in the following manner (Funding Ranking… 2003, p. 96):

"AvH researchers predominantly chose universities which have also developed a reputation as being especially active in research through the total volume of DFG approvals they attract".

In addition to the international mobility of individual researchers, the DFG funding ranking investigates the internationality of research by analysing the participation of research institutes in the Fifth EU Framework Programme. The findings show that German institutes collaborated most frequently with research institutes in the UK (15.6%), Italy (7.1%) and France (5.9%). Research collaboration with the UK is particularly strong. A visualisation of the network formed by joint projects provides information on the centrality of, and strength of linkages between, the countries participating in the Framework Programme.

"A core structure clearly emerges, consisting primarily of EU member states. The central position is taken by the UK, which took advantage of its strong participation in the Fifth FP to establish cooperative links to researchers in almost all countries involved in the programme". "Germany's position also reflects its strong participation in the EU programme, whereby… the partnership with the UK is particularly robust." "Regional concentrations are found, if at-all, towards the edges of the core structure" (Funding ranking… 2003, p. 107).

The link with DFG grants is again established by analysing the relationship between DFG approvals and the number of projects in the Framework Programme by university.

"In general it can thus be noted that universities which are strong in terms of DFG approvals also perform above average in relation to the number of funding applications to the EU granted per professor." "It can be noted that it is primarily universities with an engineering sciences focus that take advantage of the programmes offered by the EU comparatively often." (Funding ranking… 2003, p. 109).

In its evaluation of a personal award scheme, the Wellcome Trust has adopted a perspective that views international movement of researchers as one aspect of their career development. The Wellcome Trust's interest in internationalisation is limited to the extent that information on international work is able to characterise the career progress of the individuals. It does not attempt to explain how the Wellcome Trust grants have contributed to this internationalisation.

The Wellcome Trust investigates where the Prize Students who received awards between 1988 and 1990 have been working during their subsequent careers. The findings imply that the awardees have forged links between the UK research base and international research organisations:

"Substantial numbers of those who have remained in academic research (33 of 58) obtained postdoctoral academic research experience outside the UK…" (Review of… 2000, p. 3.) "At the time of the survey, almost half had returned to the UK and a significant number of those who remain in academic research in the UK today have obtained research group leader status or command long-term research grant funding" (Review of… 2000, p. 29).

With more detailed investigation into the motivations to move abroad it may have been possible to highlight contributions that Wellcome Trust funding had on their international mobility. For instance, during their doctoral research, the researcher may have acquired contacts

abroad that influenced later decisions to move abroad. However, the analysis does not investigate these types of influences that the grant may have had.

3.4.3 Knowledge Transfer between Scientists and Research Users

In the assessment of collaborations between research organisations and research users, there is an implicit assumption that collaboration involves knowledge transfer between academia and other actors in society. The evaluation projects analysed appear to be primarily interested in investigating two issues: to what extent knowledge diffusion between research organisations and other actors takes place during research projects, and whether research projects have helped to establish cooperative networks among the participating actors. Because the selected evaluation projects focus on the assessment of research impact, the emphasis is on knowledge transfer from academia to industry, or to policy makers, rather than from non-academic stakeholder groups to the scientific community.

The outcome indicator system developed by the FRST measures cooperation similarly to user outputs: number of partnerships, contracts and linkages that the research organisation has had with other stakeholders are reported annually. In addition, for each indicated cooperative contract, the user category and FRST programme type are indicated.

While the FRST system serves to provide information on the whole range of funding schemes, it does not help us to understand how the links between academia and other stakeholder groups form. The ESRC investigates this in great detail with three different methods described in Section 2.3.7. The 'user panel' method addresses the extent to which the members of the user panel were in contact with the researchers prior to and after the research programmes. It also investigates user involvement in different programme activities, such as attendance in programme meetings and advisory roles for projects or the programme. The evidence of network creation generated by the method was mixed (Molas-Gallart et al. 1999, p. 48):

"A majority of users… claimed to have benefited from new networks and contacts since becoming involved in the programme and most (six) expected to use these new links in the future. We were not able to detect the specific source of these new links."

The second technique, the 'networks and flows' method is most suitable for mapping the channels through which research knowledge flows to research users. The main finding produced by the study was that the research projects primarily strengthened existing networks. The authors note that (Molas-Gallart et al. 1999, p. 65):

"If the user-academic networks were not strong before the start of the Programme, they remained weak after its end. In general, networking rated low among researchers as a benefit derived from participation in the Programme." "However, there is a type of academic/ user link that did strengthen during the Programme. Many of the users who were known to the researchers before the Programme started, and who maintained an interest in their studies during the Programme, offered for the first time consultancy and research contracts to project researchers."

The third method used by the ESRC, historical tracing of post-research activity, is quite similar to the method used by the Wellcome Trust. Both methods generate evidence of the extent to which individual researchers have moved from academia to non-academic pos-

itions. There is an important difference, however. The Wellcome Trust assessment does not attempt to track the impacts of research projects on later work of these individuals. Rather, the assessment provides only descriptive statistics on the career progress of the individuals, and supplements those with qualitative assessments of the motivations of the researchers to leave academic work. The ESRC study, in turn, attempts to investigate whether the individuals have applied the skills learned during research work on problems in their later work outside academia.

The nature of the findings provided by these two methods differs. The main findings from the Wellcome Trust evaluation include the following (Review of… 2000, p. 3):

"A substantial majority of the cohort (81 per cent) took a first postdoctoral position in academic research, although the proportion which remained fell to just under half (46 per cent) at the time of the survey. A significant proportion (18 per cent) is currently employed in the pharmaceutical or biotechnology industries. The majority (73 per cent) of those who left academic research are still working in science, medicine, or health-related fields. More women than men left academic research, usually within the first three years. None of the cohort has been involuntarily out of work for any significant period of time."

"Members of the cohort who elected to leave academic research consistently identified three reasons for doing so: Many cited the lack of job security inherent in short-term academic contracts and the need to apply for research funding continually. Another reason often cited was the lack of a defined career path or career structure in academia. The third, and almost universal, reason was that academic research was underpaid when compared to the salary opportunities available elsewhere, for example in the commercial sector." (Review of… 2000, p. 13.)

The ESRC evaluation has resulted in similar findings on the extent to which researchers take up positions outside academia, but the evaluation also provides information on how research benefits non-academic actors. In this case, there is evidence of strong impacts through the application of skills in non-academic work (Molas-Gallart et al. 2003 p. 69):

"…most researchers pointed to a variety of tasks undertaken outside academia where they had applied skills, theories and methods developed during the Programme. Almost half have since the end of their projects played advisory roles, and about one third have been involved in contract research and commissioned work, reflecting substantial non-academic activity, even among those remaining in academia. More important, they overwhelmingly pointed out that they had applied to this work skills and knowledge gained from the project… and new methodologies and concepts developed in the project…"

The approach employed by the Wellcome Trust is perhaps more appropriate when the purpose of the evaluation is to generate information to be used in the improvement of existing research funding schemes. The 'tracing of post-research activity' approach employed by the ESRC study, in turn, is appropriate in evaluations that attempt to assess the long-term impacts of the research. While the use of the Wellcome Trust method for impact assessment would have to rely on an assumption according to which research mobility automatically leads to the use of the skills and knowledge developed in the research projects, the ESRC method explicitly attempts to assess whether such

impacts take place and what are the channels that lead to these benefits.

3.5 Utilisation of Research and Resulting Impacts

As the last of the categories of assessment, the study now presents the methods that were used to assess the extent to which the various influences from basic research result in impacts at stakeholder organisations outside the scientific community.

Here the concept of impact is understood very broadly: all the changes in organisations brought about or influenced by the research activities comprise the impact of research. As discussed in Section 1.3, the various impacts can be characterised, for instance, as economic, societal or cultural, and thus the spectrum of potential impacts is wide. An ambiguous concept such as this leads to challenges as regards its measurement. Impacts are difficult to measure, because there is a broad variety of potential changes that can take place and a similarly broad range of potential users. This implies that a decision has to be made as to the level of detail and coverage desired in an impact assessment. A pattern emerges from the analysed evaluation projects: Typically, the approach chosen involves either general assessments by a large group of researchers, not end-users of research, of the realised impacts, or evaluations by a very precisely identified group of users that are able to provide more detailed assessments of the impacts. An aggregation or generalisation of findings, respectively, then provides an understanding of broader impacts. An exception to this is the method employed in the assessment of the Australian Research Council. Rather than assessing wider impacts based on aggregation of project level findings, the ARC evaluation employed a 'top down' approach to evaluate the total impacts of ARC funding by estimating the value of the various types of benefits.

In the following sections, the approaches used to evaluate research impacts on non-academic stakeholder groups are elaborated in more detail. First, indicators used to imply relevance of research are discussed. Second, the study presents indicators and methods associated with commercialisation of research results, including both commercialisation by researchers themselves and commercialisation by other companies. Third, methods used to assess changes in user organisations are presented. Finally, the methods employed to assess broader socio-economic impacts are discussed.

3.5.1 Relevance of Research

When the scope of impact assessment is very broad (analysis is carried out for example at the national level or at the level of the granting organisation), it is difficult to gather information on the impacts of basic research, particularly when the impacts need to be attributed to specific funding instruments. Partly because of this challenge and partly because there is plenty of data available, third party inputs to research process are used as a proxy measure for the relevance of research. From the perspective of impact measurement, an assumption can be made that research funded by third parties such as business enterprises and non-private organisations also have a greater impact on the non-academic stakeholder groups.

Of the analysed evaluation material, the DFG funding ranking, the review commissioned by the Advisory Council on Science and Technology and the Office of Science and Technology's PSA Target Metrics 2004 measure third party funding with the interest of drawing

conclusions about the relevance of research. The underlying logic is presented by the DFG funding ranking as follows (Funding Ranking… 2003, p. 31):

"The evaluation presented here leads to the conclusion that – both in absolute terms and in relation to the number of scientific staff – the significance of third party funding varies greatly between research areas. The simple formula "high third party funding levels = high research activity" is basically valid. However, it only describes the true situation in the so-called 'third party funding subjects', for which the amounts sourced in this way constitute a significant portion of the costs".

The PSA Target Agreement 2004 report presents a comprehensive set of indicators for the assessment of business expenditure:

- Business R&D investment (BE) in publicly performed R&D (BE-PUBERD as a proportion of PUBERD)
- Business R&D investment in R&D performed in the government sector (BE-GOVERD as a proportion of GOVERD)
- Business R&D investment in R&D performed in the private non-profit sector (BE-PNPERD as a proportion of total PNPERD)
- BE-HERD as a proportion of total HERD
- BE-HERD as a proportion of total HERD in four main research areas

The purpose of these measures is similar to that in the DFG funding ranking report (PSA target… 2004, p. 120):

"Business expenditure via investment in other sectors may reflect confidence in the research and relevance to business objectives."

Likewise, the ACST's report analyses research expenditure by source of funding. The emphasis is on trends and comparisons between countries to increase understanding of the progress (Taking… 2003, p. 1):

"Despite the limitations of using [input measures such as GERD], there are significant benefits. These relate to the availability of good data and the ability to compare data over several years and between different jurisdictions. Few other indicators are as robust."

The advantages of these indicators are that data is readily available, it is relatively accurate, and comparisons with other countries are simple to carry out. However, based solely on these input measures it is impossible to know whether exploitation of research actually takes place, let alone attempt to understand how basic research influences society.

3.5.2 Commercialisation of Research Results

Commercialisation of research results is an outcome of basic research that can be observed relatively easily, because the process results in identifiable products and services offered by companies. An additional advantage associated with impact assessment focusing on commercialisation is that the value of the benefits from research is reflected by the revenues generated through the sales of products. These reasons explain why impacts of research are quite often estimated by valuing commercialised research outcomes.

In the analysed material, commercialisation is assessed by analysing statistics describing commercialisation patterns at the national level, by requesting researchers to assess the potential that research outcomes are commercialised, and by conducting case studies on successful commercialisations. Based on various sources of data, the ACST's report provides multiple indicators of commercialisation performance in Cana-

da, including measures for the actual and proportionate performance of Canadian institutions in relation to the US:

- Total sponsored research expenditures
- Invention disclosures
- US patents filed
- Licenses and options executed
- License income
- US patents issued
- Start-up companies formed
- Income/$ expended
- $ expended/number of start-ups formed
- Number of licenses yielding income
- Equity value of university spin-off companies.

As well as commercialisation efficiency measures:

- Invention disclosures per $ research expenditures
- Licenses yielding income per $ research expenditures
- Gross license income per $ million research expenditures
- Number of start-ups per $ research expenditures.

For most of the indicators, the results are compared with the United States, which provides an appropriate frame of comparison. The analysis provides, for example, the following findings (Taking… 2003, p. 1):

"For the majority of the key indicators, Canada contributes some 6% of the continental total. When normalized against research expenditures (as reported by AUTM) Canadian institutions appear to do slightly better than average in terms of innovation disclosures and licenses executed, but less well in terms of patents and license income. Canadian institutions appear to be better at the activities that the outcomes when compared with the U.S. Normalized against total research expenditures, license income for Canadian institutions is only 70% of the continental rate. However, start-up company formation is 250% of the continental average."

Similarly to the third party funding indicators discussed above, the data usually does not enable attribution of commercial benefits to specific research activities. From this angle, the question can be approached by inquiring researchers about the commercial potential of the research. Both the assessment of the impacts of the EU Framework Programmes in the UK and the FWF impact assessment have employed questions regarding commercial exploitation of research results in the questionnaires. The questions address both the potential for commercialisation by researchers themselves and commercialisation by other organisations.

The evaluation system of the NIST's Advanced Technology Programme also addresses these issues. The context is somewhat different, however, as the research is partly conducted by companies and, thus, the commercialisation potential can often be evaluated directly by assessing the information characterised in Table 2, Section 2.3.8.

Another typical method employed to assess commercialisation benefits is the case study method. In the analysed literature, case studies are primarily used as a means to generate data for further refinement. The CPRS method is based on case study type assessments of projects in the ATP. Similarly, the case studies in the ARC assessment are used to estimate the rate of return of two of the benefit channels identified in the assessment, namely 'benefits from building the basic knowledge stock' and 'benefits from generation of directly commercialisable intellectual property'. The case studies focus on twelve ARC funding-based technologies, technology companies, research institutes or other organisations. Key elements in the case studies

include the market capitalisation and sales revenue of companies as well as the benefits accruing from ARC funding-based research institutions and technologies to society.

3.5.3 Changes in User Organisations

In the analysed material, the changes in user organisations are usually conceptually structured by asking who the users of the research are, and how the research results are used. Typically, the user groups are broadly categorised as business enterprises and government agencies. More elaborate categorisations can be made depending on the programme and its target groups.

The conceptual framework developed in the ESRC evaluation project provides an example of the ways to categorise purposes for which the research is used in user organisations. In the framework, the uses of research include influence on judgement (justification, confirmation or development of policies) and use as a problem-solving tool (direct, indirect). The uses of research can also be identified based on more specific needs in a given research field. For applied health care research, examples of such include the use of clinical guidelines and implementation of programmes to change clinical practise.

The methods used to evaluate impact on user organisations rely either on indirect assessments by researchers or on direct assessments by users. Because researchers are often unable to assess the purposes for which research has been used in the user organisation, surveys targeted to researchers typically do not attempt to collect information on the types of research impact. Instead, with surveys one typically collects information on whether different types of potential user groups have utilised the research in any form.

In studies where the primary informant is the user, the purposes of use and impacts of research can be investigated in more detail. Moreover, in order to define the group of users, the evaluation requires a narrow focus. This enables the assessment to focus on certain types of impacts from research, generating a more elaborate picture of the uses of research. The 'networks and flows' and 'user panel' methods serve as an example of such approaches. These methods enabled evaluators to make detailed interpretations of the impact processes (Molas-Gallart et al. 1999, p. 57):

"Users confirmed that the research results had had some impact on their planning practices, mainly by providing contributions to the development of new strategies"

"… in the absence of involvement in the formal framework of the Innovation Programme, the control group members gathered information in the field, and made contact with academics, in a less structured, and possibly more time-consuming way." (Molas-Gallart et al. 1999, p. 53)

By addressing the processes which lead to impacts on the user organisations, these assessments also enable conclusions to be drawn on the factors that influence utilisation. In comparison to surveys, the potential of these methods for providing ideas for the improvement of funding instruments is likely to be greater, but their focus and coverage are typically much narrower.

3.5.4 Impacts on Society at Large

The ultimate goal of any research funding instrument is to generate broader societal impacts that effectively respond to the observed needs in society. The impacts are usually not a result of any single activity, but are aggregated results of research efforts that materialise in the

long term as changes in the activities of various groups of actors.

The analysed evaluation projects provide various examples of methods which can be employed to assess this. Similarly to the assessment of research impact on user organisations, broader societal impacts from research projects can also be measured as subjective assessments by researchers. The impacts are typically categorised into thematic groups that are either based on the goals of the addressed programme or organisation, or on conceptual frameworks of different potential research accruing from research efforts.

An example of the former approach is provided by the assessment of the impacts from the EU Framework Programmes in the UK. The impacts are grouped into the following categories reflecting the various European Community objectives:

- European scientific and technological capabilities;
- Social cohesion across the member states;
- European industrial competitiveness;
- Quality of life and health of European citizens;
- Preservation or protection of the environment;
- Employment situation across Europe.

An example of the latter approach is provided by the FWF impact assessment. In this case the impacts are measured in relation to:

- Perception of the field of science;
- Diffusion of knowledge;
- Enhancement of knowledge base;
- Impact on public discussion;
- Contribution to social problem solving.

The reliability of this method can be questioned, because it is probable that the way in which individual research projects contribute to these goals is not known to researchers – or anyone else for that matter. However, although the reliability is poor, comparisons among the various measured impacts in relation to different funding programmes and fields of science may provide an overall estimate of the types of impacts these instruments have, as perceived by the participants.

Case studies of research projects, in turn, provide a richer description of the impacts that a project has had. However, reaching a conclusion regarding impacts at a more aggregated level – for example at the level of the programme or at the national level – based on project-level data is also challenging in case-based research. There are at least two approaches. First, the case examples can be selected as to include projects with exceptionally large impacts (the approach taken in the evaluation of the ARC's contribution to the Australian economy). This provides evidence that at least the described impacts have materialised, but is not able to tell what additional impacts might have been generated. Second, one can select the cases more comprehensively and attempt to interpret the findings to reflect impacts at the programme level or the national level. This, however, is complicated by the broad and long-term nature of the impacts. The evaluators of the impacts from the EU Framework Programmes note that (The Impact… 2004, p. 37):

"*Our efforts to trace wider impacts, in policy and industry, using case study methods were inconclusive with the contributions of even the most successful projects tending to be downgraded when scrutinised and weighed against the many other contributory factors… and rather modest and gradual pace of exploitation and diffusion. This does not mean that framework projects are not worthwhile, but rather that single pro-*

jects, and even small collections of projects… are unlikely to be decisive in the emergence of new products, processes or markets."

In the Advanced Technology Programme, the difficulty of assessing the programme in its totality has been addressed by the development of a composite performance rating system. This, together with individual case descriptions on project impacts would perhaps provide a better understanding of the distribution of impacts within a portfolio of projects. However, the CPRS method was not developed as a tool for impact assessment and, thus, the feasibility of a similar method in the context of impact assessment would have to be investigated separately.

Other methods for estimating broader socio-economic impacts include the approaches employed by the National Institute of Standards and Technology and the Allen Consulting Group for estimating an economic value for research activities. Between these methods there are important differences that stem from their differing coverage. The evaluation conducted by the Allen Consulting Group attempts to estimate the total impacts from all funding provided by the Australian Research Council. Basing the estimation on user assessments is not feasible, because potential users are too numerous to be identified. As a result, the assessment was carried out with a 'top-down' rather than 'bottom-up' approach.

In contrast, the economic estimations conducted by the National Institute of Standards and Technology focus only on the actors that benefit from the activities in a specific R&D programme. This sharper focus is made possible by the nature of the development efforts. The economic impact analyses at the National Institute of Standards and Technology focus on technological development, e.g. development of standards that facilitate economic interactions within value chains. In comparison to the method used by the Allen Consulting Group to estimate the returns on investment from the Australian Research Council, this considerably simplifies the estimation exercise, since it can be based on a calculation of benefits to individual actors (i.e. a 'bottom-up' method). In addition, the link between the value of research and research effort is easier to establish. In the case of basic research, the user of the research results is typically unknown.

In regard to the assessment of basic research impacts, these methods entail different challenges. In the method used by the Allen Consulting Group, the challenge is related to the estimation of total impacts based on a relatively few case examples of funded research with large impacts. In the method employed by the National Institute of Standards and Technology, the challenge is related to identifying the cost savings and increased profits attributable to the funded research.

Finally, the results provided by the methods are different. The estimates of the rates of return for the Australian Research Council are very speculative, but they attempt to cover all ARC funding. In contrast, the estimates for the National Institute of Standards and Technology are likely to be more accurate, but they address only a specific R&D programme, and may not be feasible in the context of basic research.

3.6 Summary

The approaches with which the analysed evaluation projects characterise the impacts of a funding organisation can be grouped into four conceptual themes.

The themes reflect the stages in the process through which research funding results in impacts: 1) allocation of funding, 2) research outputs, outcomes and impacts in the scientific community, 3) collaboration in research and 4) exploitation of research and resulting impacts. The stages emphasise the distinction between research impacts on the scientific community and the impacts on other stakeholder groups.

With regard to allocation of funding, the evaluation projects characterise impacts from the operations of research funders by assessing the inputs to the research base at the national level, by analysing the significance of the given type of funding for the applicants, by characterising the funding recipients in contrast to unsuccessful applicants, and by evaluating the additionality of the funding. The two former approaches are usually based on existing databases, while the data for the two latter approaches is often collected with questionnaire surveys. The benefits of employing analyses of funding patterns as a method for assessing and communicating impacts include the availability of reliable, quantitative data that enables international comparison. Moreover, the information provided by the analyses of funding patterns is not diluted by external factors and thus provides a relatively accurate picture of the influences from the operations of a funding organisation. The disadvantage of the approach is that the link between research funding and impacts accruing from research efforts can only be assumed. Assessments of funding allocations are unable to provide direct information on the societal impacts from the research funded by the organisation.

The evaluation of direct impacts from research has primary been addressed in the analysed material by assessing three issues. First, the assessments have focused on scientific outputs, both by simply calculating the produced publications and citations, as well as by identifying factors that influence the output volume. Second, the outputs targeted to research users outside the scientific community have been assessed by calculating the numbers of publications, presentations and services provided for users. In the assessment of societal impacts at an aggregated level this approach has an important role: The numbers of user outputs serve as a proxy measure for research utilisation, but the data can be gathered directly from the recipients of the funding, which considerably simplifies the exercise. Third, the evaluations have analysed the impacts of research on the research performers. The main focus of the analysis is on the increased opportunities of the researchers that have materialised as a result of the funded research.

Since impacts on research users are complex to identify and quantify, collaborations are often used as a measure for impacts. Furthermore, collaborative linkages are also viewed as an important impact of research funding in their own right, because they are assumed to have a positive impact on the performance of the innovation system in general. The analyses of impacts from basic research funding can be grouped into three categories. First, the evaluations have addressed the extent of scientific collaborations in research programmes. Second, the analysed evaluation projects exemplify methods for assessing international cooperation among research organisations. Third, the analyses have focused on knowledge transfer between the scientific community and the research users. The methods employed range from simple calculations of collaborative contracts to visualisations of network

structure and qualitative investigations of the formation of cooperative relationships.

Finally, the selected evaluation projects have employed analyses focusing directly on the impacts on research users. The approaches include assessments of the relevance of research for the users, indicators of the commercial exploitation of research results, investigations to changes in the user organisations, as well as attempts to assess the impacts of research efforts on society at large. The assessment of changes by the research users and the evaluation of resulting aggregated impacts have proved to be the most challenging of the analysed issues. The impacts are diverse and elude quantification. On the other hand, with qualitative methods it is difficult to achieve comprehensive coverage when the scope of the evaluation is broad. Evaluations focusing on narrowly defined research areas appear to be more conducive for successful assessments of societal impacts accruing from research funding.

As regards the differences of the four stages, the granting organisation's ability to influence the generation of the analysed impacts decreases as the focus of the analysis shifts from the assessment of funding allocation towards the assessment of impacts on research users. On the other hand, the more the focus of the analysis moves further along the chain, the more relevant the results of the associated indicators become to the assessment of the societal goals of research funding. The scope with which it is possible to conduct evaluations becomes narrower. There is thus a trade-off between the relevance of results and the amount of influence that can be exerted to the analysed issues by the funding organisation. A similar trade-off is apparent between the coverage of impact assessments and the extent of particularity in the descriptions of research impacts that can be achieved with the assessments.

4 Conclusions and Recommendations

The purpose of this chapter is to generate ideas for the impact evaluation activities in the Academy of Finland. Due to the nature of the present study, the recommendations given should be viewed as suggestions and possible new directions for developing and complementing current activities, rather than as strong instructions of how the activities should be organised. The suggestions are based on the ideas emerging from the analysis of the international evaluation projects.

4.1 Evaluation Activities in the Academy of Finland

Evaluation activity in the Academy of Finland branches out to three directions. First, the Academy of Finland operates as the central body for advancing science and scientific research in Finland. This is reflected in the Academy's main objective: "Scientific research in Finland is characterised by high quality and prominence". In order to provide information as regards the achievement of this goal, the Academy conducts investigations to assess the state of science at the national level.

Second, evaluation is used to generate information on the performance of the Academy of Finland, assessed on the basis of its agreement on target outcome with the Ministry of Education. The Academy's objectives are related to four target outcome areas (Opetusministeriön… 2003):

1. The development of research capabilities and research environments;
2. Advancing research careers;
3. Expertise in science policy;
4. Interaction between science and society.

All of these are objectives that can, at least to a certain extent, be informed by evaluations of research and scientific activity. Ex-post assessments of research are part of the tool box that can be used to generate information for indicating how well the Academy has performed in relation to its objectives.

Third, evaluations are used to generate evidence on the impacts of the Academy's funding instruments. Funding instruments are the main tools with which the Academy of Finland generates societal impacts through scientific activity. Evaluations at this level are thus central to the development of the Academy's operations.

In addition, there are investigations that are carried out on current, specific themes in science policy, the scope of which can be outlined according to the need at hand. This category consists of a heterogeneous group of reviews, assessments and surveys on a range of issues. Due to its heterogeneity, the fourth category is not further discussed in the present study.

From the perspective of ex-post assessment of research impact, the three conceptual levels of evaluation activity can be viewed as hierarchical: Evaluations at the national level focus on the state of science in Finland irrespective of the activities of the Academy. Evaluation at the national level provides information on the context in which the Academy operates. The second level focuses on the operations of the Academy in this context. The funding organisation steers its operations in relation to the understanding of the national research system and the sectors connected to it. The cen-

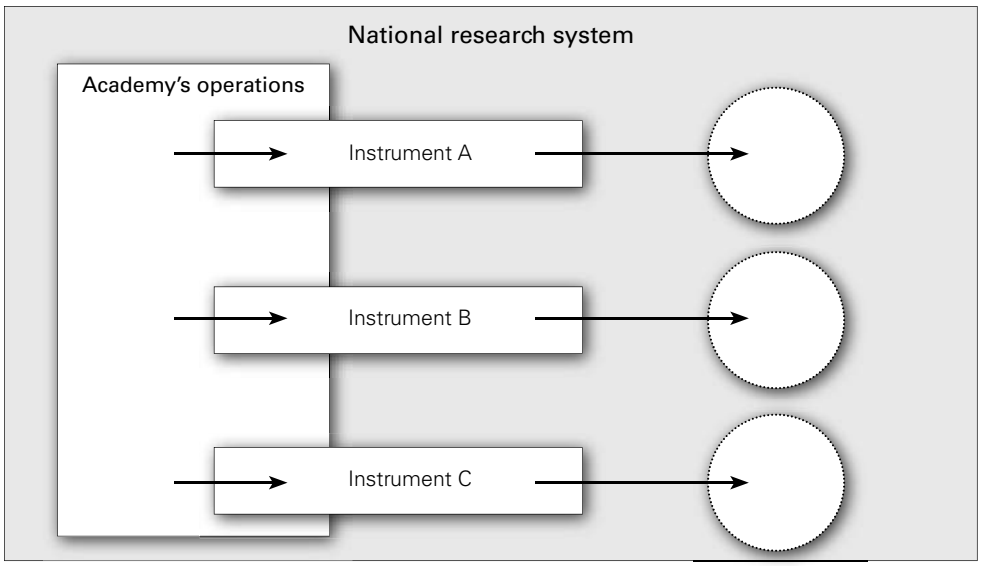

Figure 2. Three levels of the Academy's evaluation activity: national, organisation, and instrument levels.

tral question for impact assessment at this level is "how is the Academy of Finland contributing to the Finnish research system and society at large?" The third level addresses the instruments that the Academy employs to respond to the needs of scientific community. Impact evaluation at the level of instruments can provide information on the actual operations through which the Academy executes its strategy and objectives. Figure 2 provides a schematic representation of this conceptual framework.

This is also the order in which the ideas emerging from the analysis of international evaluation projects are presented. First, general perspectives for impact assessment are discussed. Second, the study briefly reviews evaluation activities that address the national level. Then it raises ideas from the analysis of the twelve evaluation projects and discusses the conditions in which the new approaches would be appropriate for the evaluation activity in the Academy. The discussion on the next two activity areas, evaluations of impacts from the Academy's operations and ex-post evaluations of funding instruments, is similarly structured.

It should be kept in mind that the evaluation projects analysed in the study emphasise impact indicators, which as a method of evaluation is more suitable for assessing techno-economic impacts than other, perhaps more subtle and less goal-oriented, outcomes of research such as impacts on culture. Thus, also the suggestions made in this chapter are coloured by an emphasis on the assessment of techno-economic impacts.

4.2 Advancing Scientific Research in Finland

4.2.1 Perspectives to Impact Assessment at the National Level

In the context consisting of the three impact evaluation activity areas, the activities generating information on the bearing of science on society can be viewed as having two different purposes. The first seeks to shed light on the general question of how science influences the

national economy, culture and society at large. The emphasis is on the generation of evidence on the impacts that science has on society.

Because the paths through which science influences different aspects of society are multiple and in many cases difficult to identify, the assessment of the impact of science at the national level is a very challenging task. There are few indicators that would reflect the impact of science on the users of scientific information because the forms through which this information is used are numerous. Assessment of the impacts of science relies mostly on nation-level assessments of the relation between research intensity and the development of technology intensive sectors, indicators of scientific output, and estimations based on collections of more focused qualitative analyses on the impact of scientific research.

The second purpose is related to identifying areas in the national research base that are underperforming, and that may have an influence on the extent to which scientific research generates the different outcomes that are needed by society. Such areas may be related, for instance, to the international mobility of researchers or to an inadequate level of exploitation of scientific knowledge. Thus, the main purpose of the assessment activity is no longer that of creating evidence of impacts. Rather, it seeks to identify potential areas of improvement in order to steer the activities for the structural development of the research system so that the performance of the national research base increases in the long term. The scope of the assessment becomes broader, because one is no longer interested only in the impacts of science, but also in the inputs, mechanisms and conditions that influence how societal impacts are realised.

As for organisations funding scientific research, this type of analysis may provide ideas for managing their operations and for steering their funding allocation. Thus, analysis of underperforming elements in the national research base may be closely linked to the impacts that a funding organisation has on national scientific activity and through that on society at large. There is a more general issue bearing on this, namely that the effectiveness of the funding organisations' activities is dependent on the need for these activities. The extent to which impacts can be said to materialise through the funding organisations' activities depends on the relationship between the need and the activities conducted relevant to this need. In short, identification of areas of improvement in the national research base may influence the impacts of a funding organisation and its instruments.

However, it should be noted that ex-post impact assessments are perhaps not the best way to identify and prioritise areas for development in the national research system. There is a need to complement methods that assess past activities with methods that focus on future possibilities and challenges, such as foresight techniques. As part of the Sight 2006 initiative, the Academy of Finland, together with Tekes, the Finnish Funding Agency for Technology and Innovation, investigates the future challenges and strategic priorities of the Finnish research system. As a complementary technique, impact assessment methodologies can make a contribution to this process of prioritisation.

4.2.2 Academy's Evaluation Instruments

The Academy of Finland has two major evaluation instruments that generate in-

formation on the performance of the Finnish science base, the review on the state and quality of scientific research in Finland, and evaluations carried out of research fields. The review on the state and quality of scientific research discusses the research base – its organisation, relevant policies, quality and impacts – at a national level. The scope of the review includes all scientific disciplines. The assessment has been carried out once every three years. The purpose of the review is stated as follows (Scientific… 2003):

"*The primary aim of the review is to serve the needs of national and international bodies and organisations responsible for science and technology policy as well as research funding. It also provides useful information for researchers and research organisations around the world.*"

The main areas of interest in the 2003 review include descriptions of the R&D expenditure and sources of research funding in Finland, a discussion on the human resources in research, a description of science policy, and an assessment of the scientific and social impact of research. These main themes are complemented with discussions on issues currently important in the policy discussion and with reports by the Academy's four research councils. All these aspects are addressed at the national level, independent of any single funding or research organisation.

In the future, the Academy will continue assessing the state of scientific research in Finland, but the scope of the assessment will be broadened. In comparison to the 2003 review on the state and quality of scientific research in Finland, the Sight 2006 project includes a foresight project, a more extensive bibliometric study on the performance of Finnish science as well as a stronger emphasis on the societal and structural impacts of the operations of the research funding organisations.

In contrast to the review on the state and quality of scientific research, the second evaluation instrument has a narrow scope. These evaluations focus on the assessment of research fields. The evaluations are carried out as peer-reviews, supported with additional methods that generate information for the basis of the assessment. The most recent evaluations on research fields include evaluations on Finnish business disciplines, nursing and caring sciences and geosciences.

Typically, the evaluations address issues such as the historical development of the scientific field in the Finnish context, the resources and funding allocated for it, the conditions for researcher training, and scholarly performance. In addition, the evaluations discuss how the research is organised, how intense international mobility is in the field being evaluated and often also include descriptions of the main research units in the field. The Finnish research system is compared with international counterparts throughout the evaluation (Finnish Geosciences… 2003, Nursing… 2003, Research in Business… 2005).

The assessment of the societal impacts and the relevance of research has not been strongly emphasised, although this issue is also addressed. The main themes related to impact assessment include assessment of publication activity, identification of research areas in collaboration with external actors and the organisation of the academia-industry interface in general.

Reflecting on these two evaluation instruments in relation to the two perspectives discussed above, it is possible to further characterise the contribution of the instruments for impact assess-

ment. The review on the state and quality of scientific research in Finland appears to have a stronger role in generating information on the role of science in society in general than serving as a basis for decision-making by assessing strengths and challenges of the Finnish research system from an international perspective. Though there are indicators of Finland's performance in comparison to other OECD countries, these appear to be relatively few. They address issues such as research expenditure, number of research personnel, numbers of publications and citations.

As for creating evidence of research impact, the review on the state and quality of scientific research in Finland provides national-level measures of the output related to the qualified research personnel, of publication and citation outputs as well as a collection of qualitative assessments on the societal impacts of science. The task of reviewing the impacts of research at the national level is very challenging, due to the multiplicity of ways in which science may have an impact and the number of actors influenced by it. There are no general measures for assessing the extent to which science influences society. Thus, the review provides descriptions and case studies of the various ways in which science influences actors external to the scientific community.

The purpose of the evaluations on research fields, in turn, is geared towards informing decision-making rather than evaluating the impacts of research conducted in the field. The evaluations create recommendations for both funding organisations as well as research performers for improving performance in the field. Thus, from the perspective of the Academy of Finland, they can be viewed as a potential source of information on how to increase the effectiveness and impacts of the Academy's funding instruments.

4.2.3 Applicability of Methods Employed in the Evaluation Projects

Among the twelve evaluation projects analysed there are two projects which address research outcomes at the level of a national research system, the Office of Science and Technology's (UK) Public Service Agreement (PSA) Target Metrics 2004, and the Advisory Council on Science and Technology's (Canada) Taking Stock of the University and College Commercialisation Efforts. In addition, parts of the funding ranking conducted by the German Research Foundation are also applicable to the national level. First, ideas provided by these evaluation projects as regards the need to generate evidence of research impacts are discussed. Second, ideas regarding the use of impact assessment as a means for identifying potential development needs are presented.

The 2003 review of scientific research in Finland discusses the societal impact of research in a qualitative manner, but does not include many quantitative indicators of the use of research results. The report submitted to the Advisory Council on Science and Technology emphasises commercialisation of university research as an important impact of science. Even though commercialisation indicators capture only a small part of the research impact, they could nevertheless be included in the Academy's measurement activity, particularly if utilisation of research results becomes a higher priority in the future.

The selection of indicators could possibly include at least the following indicators:

– Patents;
– Number of licenses yielding income;

– License income;
– Start-up companies formed.

The indicator values could be provided both as absolute figures as well as efficiency measures relative to the research expenditures in the field. Notably, the measures stress impacts from more technically oriented sciences, as patents are rarely generated by research efforts in the social sciences and humanities.

Although the funding ranking conducted by the German Research Foundation addresses the level of the funding organisation, it can also provide insights for the description of performance of the research base at the national level. Particularly, it exemplifies the use of data related to the participation in international programmes, in this case the EU Framework Programme. The review of the state and quality of scientific research in Finland already includes some indicators of the international mobility of international researchers. However, since internationalisation of the research base is becoming ever more important, participation in the Framework Programme could also be reported, and compared with other Nordic countries. Visualisation of international networks formed by the projects in which Finnish researchers participate would provide a qualitative understanding of the international linkages. A similar approach could also be included for other major instruments used to support international collaboration, such as COST – European Cooperation in the field of Scientific and Technical Research.

As for the assessment of research impacts on the well-being in society, it is difficult to conduct evaluations on a scope that covers societal impacts from the whole national research system. Although there is a range of indicators measuring well-being and the quality of life in different economic, social, cultural and environmental dimensions, the relationships between research investments and the social impacts measured with these indicators are poorly understood, making the evaluation exercise complex to carry out. A recent survey on well-being indicators conducted by the Centre for the Study of Living Standards for the Advisory Council on Science and Technology in Canada concludes that measurement of the impact of research investments on well-being should be feasible, but it would require extensive empirical work regarding the quantification of the impacts from specific research investments on specific domains of well-being (Sharpe & Smith 2005).

Another approach to generating information on the impacts of research on well-being is to conduct several qualitative, more narrowly focused evaluation exercises so that generalisation of findings to higher levels becomes possible. In the Academy's evaluation activity, this could most appropriately be done in association with instrument-level evaluations, particularly programme evaluations. This is further discussed in Section 4.4.3.

In addition to generating evidence on research impacts, impact assessments can also point to development needs in the research system. The major question regarding this is whether indicators can provide information that is disaggregated enough to reveal areas which may be underperforming, and whether the indicator scores can meaningfully be interpreted as weaknesses.

The indicators employed by the Office of Science and Technology's PSA Target Metrics 2004 indicate the general performance of the UK research base in an international context. However, the information provided by the indicators is disaggregated only to the level of main scientific disciplines, which is likely to

be too coarse a level as regards the provision of strategic information for prioritisation and for the steering of operations in individual funding organisations. From the perspective of the Academy, more detailed analyses would perhaps provide a better basis for strategic decision-making and prioritisation.

In the Sight 2006 project, the Academy of Finland has initiated a bibliometric study on the performance of Finnish research. The study could be conducted at a refined level to provide a more accurate understanding of the strengths and weaknesses of the Finnish research system. Because the analysis could prove to be excessively burdensome if carried out at the level of all sciences, specific research fields could be chosen for more detailed analysis. The analysis could focus on the fields which are considered of high importance in Finland, including those which are of special interest to industry, those identified as having potential for international breakthroughs, those in which there are centres of excellence as well as those that are expected to be problem areas. Bibliometric indicators of such fields would then provide data early on to serve as a frame of reference for the measurement of impacts in case funding is allocated to that field.

Similar approaches could also be employed to support the evaluations on research fields. The analysis could be conducted to provide background information for planned evaluations and to support the work of the peer-review panels. In addition, analyses of sub-disciplines could point to research areas that require a more thorough evaluation to provide information as the basis of improvement.

Because increasing attention in the policy discussion is given to efficient and productive use of public resources, a more extensive utilisation of efficiency measures may also be called for when addressing the national level. For the Academy of Finland, the most relevant efficiency indicators used by the Office of Science and Technology's PSA target metrics 2004 include the following:

- PhDs awarded relative to HERD in main research areas;
- Citations relative to HERD in main research areas;
- Citations per researcher.

Disaggregation to the level of sub-disciplines could again provide insights on strengths and weaknesses in the research system.

In the analysis, special care should be taken as regards the standardisation of the indicators. Research impact assessments often ignore differences in the nations' scientific productivity brought about by the variance in research expenditures, the nature and composition of scientific activity, etc. In conducting assessments based on bibliometric analyses, the Academy should apply techniques that normalise the indicator values in relation to suitable measures such as the number of researchers or expenditure in the higher education sector (HERD). When interpreting the results from the analyses, one should take into account additional factors which influence the indicator scores for different nations: time lag between research investments and outputs, differing R&D cost levels, composition of scientific activity by fields, publication language, etc. (Final report… 2002).

As a whole, the emphasis of the assessment exercises at the national level could be shifted towards identifying strengths and weaknesses in the research system and in the way it is connected to other parts of society. As for the assessment of societal impact, the focus should not be on the national level, but rather on the individual funding instruments of

the Academy. The channels through which science influences society need to be understood more fully before the societal impact of science can be assessed and communicated at the national level.

4.3 Impacts from the Operations of the Academy of Finland

4.3.1 Perspectives to Impact Assessment at the Organisation Level

The underlying purposes for evaluating the impacts of research at the organisation level can be categorised either as external or internal (see e.g. Kuitunen & Hyytinen 2004). The term 'external' refers to situations where impact assessments are employed to generate evidence of impacts for stakeholder groups external to the organisation. The motivations for generating such information are related to considerations of accountability, the credibility and legitimacy of an organisation's mission and transparency of the organisation's operational processes. Impact assessments carried out to serve external motivations thus attempt to demonstrate that impacts actually materialise and that the process that leads to these generations is efficient.

Impact evaluations initiated on the basis of internal motivations, in turn, attempt to inform the developmental activities in the organisation. In comparison to externally motivated impact assessments, the scope of the evaluation becomes broader, as it no longer is sufficient to demonstrate merely that impacts do take place. In order to develop the organisation's operations to increase societal impacts, the organisation needs to understand how the impacts are generated and what factors influence the generation of impacts. For a funding organisation with a limited amount of resources to be allocated, this involves both understanding how the funding allocation should be prioritised so that it leads to outcomes in areas where societal impacts are thought to be most needed, as well as understanding the mechanisms that generate societal impacts for each of its instruments.

4.3.2 Academy's Evaluation Instruments

In the past, evaluation of the impacts from the Academy's activities at the level of the whole organisation has been limited to information generated for the Academy's Annual Report and Financial Statements (Suomen Akatemian Toimintakertomus…), which responds to the agreement of target outcomes with the Ministry of Education. Unlike some international funding organisations (such as the ones addressed in this study, the Australian Research Council and the German Research Foundation), the Academy of Finland has not published reports with the intent of assessing the contribution that the organisation in its totality makes to society.

Currently, there are increasing demands for ensuring accountability and efficiency in the administration of public resources. For the Academy of Finland, this presents itself as a growing need to develop a system indicating the Academy's performance as regards the generation of benefits to society.

The management system employed to supervise all public agencies and organisations in Finland – management by results – emphasises the assessment of an organisation's performance in relation to the societal benefits achieved by it and to the efficiency with which the public resources have been used (Tulosohjauksen… 2003). The approach highlights the significance of setting of objectives – target outcomes – and, in relation to these,

measuring achieved changes in society with a set of indicators. An important element in the approach is the measurement of the utilisation of outputs and services that bring about changes and that are provided by the organisation in question.

In the context of basic research the application of this framework is problematic, because the actors receiving benefits from research are numerous and the nature of the benefits varies. Due to these reasons, the benefits cannot be defined unambiguously. The methods for measuring societal benefits are inadequate to provide an organisation-level estimate based on the utilisation of research results. This limitation entails that the benefits from scientific research can, at best, be measured at the level of outcomes from the research process rather than at the level of the actual societal benefits to the users.

In the past, the indicators employed to demonstrate performance in relation to the target outcomes have been based mostly on data on the Academy's inputs to the research system. Some indicators have been based on the outputs of research, but the data employed has reflected national performance instead of focusing merely on the Academy's activities. The Academy of Finland is currently implementing a new information system for gathering evidence of research project outcomes. The system is based on final reports from funded projects. At the time of writing of the report, the information gathered from the project managers was planned to include (Linko & Sulonen 2004):

- Information for identification of the project;
- Funding received by the project;
- Research personnel;
- Outputs;
- A qualitative assessment of the societal impacts;
- Collaboration (collaborative partners and interdisciplinarity);
- Continuation of the project;
- Description of potential success stories.

Since project funding comprises the majority of funding volume of the Academy of Finland, the indicators that can be formed on the basis of the database cover most of the direct outcomes of the Academy's activities.

4.3.3 Applicability of Methods Employed in the Evaluation Projects

In the analysed evaluation projects the impact of a funding organisation is analysed and demonstrated in mainly two ways: by analysing how funding is allocated (e.g. characterising the recipients in some manner) and by assessing the impacts of the funded research.

In the past, the outcome indicators used by the Academy of Finland have mainly been of the former type. However, the recent increase in the requirements for productivity and accountability has created a need for developing an indicator system based on the outcomes and impacts from the research funded by the Academy. Indicators of this type have also been developed by the Foundation for Research, Science and Technology (FRST), New Zealand. The experiences of the FRST provide ideas as how the indicator development should be conducted in the Academy of Finland.

Similarly to the FRST, the Academy could strive to develop an indicator system which is stable in the long term, i.e. the indicators measure outcomes, which remain as valid objectives for Academy-funded research also in the long term. Individual values of the indicators do not increase our understanding on the extent

of the impacts that the Academy's operations have. It is only with a frame of reference that the performance of the Academy in generating impacts can be assessed. Thus, indicators selected should be stable so that changes in the impacts can be assessed. In addition, the indicators should be based on an underlying conceptual framework to provide a basis from which their significance could be assessed. Indicators should be selected in a way that data collection is simple and does not cause much additional burden for funding applicants. Existing databases should be used whenever possible.

The conceptual framework presented in this study is an example which can be used to guide indicator selection. Since it was constructed on the basis of an analysis of a heterogeneous group of evaluation projects, it is somewhat data-based and situation-specific. It emphasises the flows of knowledge between the researchers and the end-users of research external to the scientific community and, thus, partly neglects the utilisation of research results for scientific purposes by other researchers.

Instead of such a bottom-up approach, the indicators should be selected on the basis of a more theory-based conceptual framework that also reflects the long-term goals of the Academy. The development of the conceptual framework could be based on existing knowledge on the economic impacts of basic research, perhaps supported by Academy-funded theoretically oriented research focusing on the cultural and societal impacts stemming from the research process.

If a set of outcome indicators is selected on the basis of the conceptual framework despite its limitations, the selection could involve the following measures that indicate societal impacts from the Academy's operations:

Similarly to the FRST outcome indicators, the analysis of these indicators should be disaggregated to the level of scientific disciplines and, whenever possible, to the level of different stakeholder

Indicator	Corresponding channel for knowledge flow
Number of publications targeted for users produced in Academy-funded projects	Knowledge transfer through publications
Number of patents produced in Academy-funded projects	Knowledge transfer through intellectual property rights
Number of services provided for users in Academy-funded projects	Knowledge transfer through consulting and interaction with users
Number of collaborative contracts between Finnish project participants	Formation of linkages that increase the transfer of knowledge within the innovation system in the long term
Number of international collaborative contracts alongside with Academy-funded projects	Formation of international linkages in the research system
Number of follow-up projects (including projects running alongside with Academy-funded projects and utilising the results)	Knowledge transfer through other research projects and accumulation of experience
Number of researchers funded by the Academy that have been employed by industry or other stakeholder groups	Transfer of skills and knowledge through researcher mobility

Table 3. Indicator selection based on the conceptual framework.

groups utilising the outcome. The indicator scores could also be analysed as proportionate to the amount of research funding allocated in the field.

The indicator selection emphasises the distinction made in the conceptual framework between the scientific community and the stakeholders outside the scientific community. The indicators attempt to measure the various channels through which knowledge produced in the research projects flows to other societal actors. Because the measurement of changes brought about by the use of knowledge cannot be directly measured, the approach settles for measuring the outcomes of research as proxy measures for societal impact.

The analysed evaluation projects have exemplified both approaches in which the data is collected from individual projects (e.g. the impact assessments focusing on the Austrian Science Fund and the EU Framework Programmes) and approaches where the data is gathered from organisations (the Foundation for Research, Science and Technology). Implementing systems for the collection of data with both approaches would be excessively burdensome, so in practice this implies that data on research outcomes is collected with either one or the other approach. This decision influences the data that can be acquired. In short, the difference between the approaches can be characterised as follows: When the data is gathered from organisation level, the outcomes of funded research cannot be attributed to a particular research effort, and all outcomes during that year are recorded. When the data is gathered from the project level, it is known which research projects produce the outcomes, but the impacts that materialise after the data is collected are not recorded.

Depending on the needs of the organisation, either one of the approaches can be preferred. If the purpose of the indicators is only to demonstrate that impacts do take place, then an organisation-level approach to data collection may be appropriate. If the funding organisation is also interested in investigating, which types of projects produce the impacts, then project-based data collection should be preferred. The shortcoming of this approach is that it is not able to measure the outcomes and impacts that are realised after the data collection takes place. Therefore, indicators based on project-level data can underestimate the actual outcomes of the funded research. This is especially the case when project outcomes are reported shortly after the project has ceased to receive funding.

There are two strong arguments in favour of the project-based approach in the Academy of Finland: First, the Academy is already implementing a new system for collecting project final reports from research projects. Second, there are pressures to develop the Academy's operations so that the utilisation of research outputs is greater and the societal impacts from the research more pronounced. The project-based system for data collection provides more opportunities for analysis of project impacts to support this goal.

Due to the shortcomings of project-based systems described above, it is of paramount importance to investigate and develop the validity of the chosen measures in relation to their ability to indicate societal impacts. Instead of developing complex sets of impact indicators, attention should be focused on improving our understanding of the phenomena involved so that the relationship between research outcomes and societal impacts

would become more fully understood. To this end, the Academy should supplement its attempts to select indicators with qualitative studies that investigate the significance and validity of the selected indicators.

At least three focus areas for such studies can be identified. First, the Academy of Finland should investigate how researchers report the information that the indicator value is based on. In relation to the more "solid" indicators, such as those based on immaterial property rights, this is not as important. However, with indicators measuring more ambiguous concepts such as collaboration, it becomes a question of what constitutes a collaboration and what does not. Investigating the types of understandings researchers have on these concepts is thus one way in which the validity of the indicators can be improved.

Second, it is important to investigate the extent to which the information reported covers all of the outcomes that can be attributed to the project. In the Academy of Finland, final project reports are submitted on average six months after the project funding has ended. It is likely that many of the project outcomes materialise only after the reporting of data. If information gathered through final project reports is used to compose the indicators for the Academy's performance, this should be taken into account. For some indicators this may be more important than others. For instance, knowledge flows brought about by researcher mobility may take place much later than those generated through the project's publications.

Most importantly, the Academy should investigate how the measured scores of the research outcome indicators reflect benefits that accrue from the utilisation of research. In practice, this would involve more theoretically oriented studies on the relationship between research outcomes and societal impacts as well as empirical studies on different outcomes as regards their utilisation, influences on decision processes, and changes brought about by them in user organisations. The contribution of the investigation would be two-fold: the study would both increase the understanding of the extent to which the outcome measures reflect societal impacts (thus increasing the validity of measures as impact indicators), and it would increase understanding on the mechanisms with which the knowledge is utilised.

An important aspect that needs to be taken into account is the non-linear manner in which impacts from research often materialise. Knowledge produced by research efforts is shared among different actors, projects and organisations, and accumulated over time. The flows of knowledge from the researcher to the user are often indirect, instead of being direct results from a single, identifiable project.

In most of the analysed evaluation projects this issue has been disregarded, partly because of the emphasis on indicators. Most of the indicators employed are based on a linear notion of the process through which impacts from research efforts are realised. In contrast, a non-linear notion of research process prompts the use of methods, which attempt to identify and track knowledge flows between researchers and (often simultaneous) research projects, and which focus on researcher mobility between organisations. This is extremely difficult and often even impossible.

4.4 Impacts from the Funding Instruments of the Academy of Finland

4.4.1 Academy's Evaluation Instruments

The Academy has three major types of instruments in addition to core project funding: research programmes, centre of excellence programme, and different personal-level schemes. The instruments address different types of needs in the research system, and thus their purposes, or societal missions, also differ:

Research programmes aim to enhance research activities in a specific research area by initiating a collection of research projects that address the research topics in the field. Research programmes also respond to needs identified in society by generating scientific knowledge as the basis for problem solving. Of the three types of instruments employed by the Academy, research programmes are most oriented towards the end-users of research.

The second of the Academy's instruments, the centre of excellence programme, has been created to provide more stable conditions for internationally promising research units to execute their research strategy. The goal of the programme is to generate creative and dynamic research environments which attain a leading position in their field of research and provide excellent conditions for training of skilled researchers. Societal impacts are also addressed by the programme's objectives, although they are given less weight in the overall mission of the programme (National Strategy for Centres of Excellence…. 1997).

The Academy has several funding schemes that are targeted at individual researchers. With these instruments, the Academy supports doctoral training, postdoctoral research careers and international mobility of researchers.

The Academy's evaluation instruments can be viewed as corresponding to its funding instruments. Ex-post impact evaluations have so far been conducted only for the assessment of research programmes. Impact assessments of the centre of excellence programme or the personal funding schemes have not yet been carried out. The first centre of excellence programme spanning the years 2000–2005 is ending at the time of writing this report, and the planning of the programme evaluation is well underway. The evaluation will be conducted in the near future.

As for the research programme instrument, all Academy programmes are evaluated after they have finished. The evaluations are carried out with a peer-review methodology. The evaluations panels review the programme on the basis of project self-evaluations, publications generated by the programme as well as presentations and interviews. In recent evaluations assessing the Sunare, Telectronics II, Syreeni and Life research programmes, the typical issues addressed have involved assessments of (Research Programme on Sustainable… 2005, Research Programme on Biological… 2004, Research Programme on Marginalisation… 2004 & Research Programme for Telecommunication… 2004):

- programme delivery (planning, resources, coordination);
- scientific quality of researchers;
- attainment of specific programme goals (interdisciplinarity, increased cooperation, internationalisation, societal impacts);
- added value of programme and significance to the Finnish research system.

Recently, increasing emphasis has been given to the evaluation of societal

impacts. This is reflected in the research programme strategy of the Academy of Finland (Academy of Finland Research… 2003, p. 19) as well as in the guidelines for research programme evaluations (Suomen Akatemian tutkimusohjelmien arviointiohjeistus 2005).

4.4.2 Applicability of Methods Employed in the Evaluation Projects

Research Programme Evaluations

Assessment of the impact of funded research on the end-users of research results is an essential goal in the Academy's evaluation activity. Achieving this with evaluations that focus on the national and the organisation level is difficult, because it is complicated to identify the users of the research in a comprehensive way. In research programme evaluations the challenge of identifying the end-users of research still exists, but the problems associated are greatly reduced because the thematic area the research addresses is defined more narrowly. The number of potential users is much smaller and stronger assumptions can be made regarding the ways in which research results influence the users and, thus, the influences can be identified more easily. As the assessment of impacts on research users is one of the main goals in the Academy's evaluation activity, the Academy should emphasise impact evaluation particularly in research programme evaluations, because there the context is most favourable for generating evidence of research impacts and the mechanisms through which the impacts are generated.

Furthermore, assessment of user impacts in programme evaluations is also suggested by the argument that impact evaluations should focus on issues which reflect the societal mission of the funding instruments. In the research programme strategy published by the Academy in 2003, more effective utilisation of research is identified as one of the key objectives for future research programmes. Regarding the role of utilisation in the implementation of the programme, the strategy states that (Academy of Finland Research… 2003, p. 19):

"Planning for the utilisation of research results shall be started from early on in the programme, and the needs for integration, analysis and popularisation of the research results shall be addressed as an integral part of coordination. It is important that the results are published not only in high-quality scientific publications but also on other forums."

In its future evaluation activity, the Academy of Finland should emphasise evaluation research that attempts to improve the utilisation of research results in its programmes. Better understanding of the factors that influence the extent to which research results are utilised is central to this task. The factors involved may include aspects from the whole programme process: identification of needs that are addressed by the programme, programme planning, implementation of the programme, and follow-up activities. An important factor is also the capacity or willingness of the society's different actors to assimilate and apply knowledge produced in the research process.

As part of the typical research programme evaluations, the Academy should initiate studies that focus on the interaction between research and its users. The evaluation project commissioned by the Economic and Social Research Council (ESRC), UK, provides ideas as to how this could be achieved. First, for the identification of users of the research results, the 'snowball approach' could be employed. For each of the programme's projects, researchers should be requested to indicate potential

users both outside the scientific community as well as in research organisations in order to take into account non-linear knowledge flows as well. The users identified in this fashion should be interviewed regarding the relevance and benefits of the research programme, and asked to indicate other potential users in their own organisation or elsewhere. These, in turn, would be similarly interviewed. In theory, the process could be continued further but, the experience in the project commissioned by the ESRC suggests that two subsequent rounds are sufficient to identify the most important users of the research.

In addition to identifying users, this approach would also provide a picture of the uniformity with which the research projects' results are being utilised. In addition, it would provide an initial understanding of the flows of knowledge between researchers and users. However, because of the way the identification of the users was conducted, the approach may yield a picture of interaction which overemphasises one-directional knowledge flow from researchers to users. In order to better understand the factors which contribute to utilisation, a case study of the research utilisation should be devised. The cases should include both situations in which the use of research results leads to successful impacts in the user organisations and situations in which potential for high impact on the user organisation existed (at some point in the research process), but utilisation of research results did not eventually take place. The cases could be selected so that they impose variance on factors that potentially influence the utilisation, such as participation of the users in the planning of the research project, the extent of interaction during the project, or the dissemination activities targeting potential user groups. The study should also investigate what the importance of programme level activities is to the utilisation of project level results. Both management of research programmes and research projects is likely to influence the extent to which research results are utilised.

Other approaches employed in the analysed evaluation projects to assess programme level impacts, such as the economic case estimations conducted by the NIST, are not recommended unless the purpose of the study is to further investigate the appropriateness of the instruments for specific needs. If an attempt to employ the NIST's economic case estimation approach is made, it should focus on a programme with direct relevance for a specific industry. An example of such a programme is the Telectronics II programme with only few research projects and relatively well-definable user industry (Research Programme for Telecommunication… 2004).

Evaluations on Centres of Excellence

Even though only one of the analysed evaluation projects focuses on the evaluation of research organisations, and none of them directly addresses the evaluation of centres of excellence, one can nevertheless present ideas regarding the evaluation of the centres of excellence based on the conceptual framework developed in the present study.

As the Academy of Finland has not yet commissioned ex-post evaluations on the centres of excellence, little is known about their impacts. Thus, in contrast to programme evaluations, it is not practical for the first of the evaluations to focus on the analysis of factors that influence the materialisation of impacts. Rather, the evaluation should first attempt to evaluate how the centre of excellence programme has succeeded in producing those impacts it strives for.

Because the goals of the centre of excellence programme emphasise the creation of dynamic, internationally recognised research environments (National Strategy… 1997), this should also be the main focus of the impact evaluation. Relating this to the conceptual framework of the present study, the primary focus of the evaluation would then be on the impacts of the funding instrument on the research organisation, and only secondarily on the impacts of the research on actors external to the organisation. Key questions would then include: What developments has the centre of excellence programme funding enabled in the research activities? Is the research environment conducive to the generation of new innovative ideas? How does the research organisation contribute to the advancement of skills and capabilities of individual researchers? Has the research become more renowned internationally?

However, the goals of the centre of excellence programme also include elements of broader societal impacts. Main questions related to collaborations and impacts on user organisations include: How relevant is the research conducted from the perspective of different stakeholder groups? Does the centre contribute to the development of research capacities available to the stakeholder groups? How well is the research organisation networked, both domestically and internationally, with business enterprises, other research organisations, and government agencies? What are the services provided for different stakeholder groups?

Because the assessment of scientific quality and significance of research from an international perspective plays a large role in the evaluation, the appropriate method for evaluating the centre of excellence programme is a peer-review by international experts conducted through site-visits. The site-visits could be supported by self-evaluations provided by the centres, focusing on similar issues as the ones used by the Dutch Universities, the Netherlands Organisation for Scientific Research and the Royal Netherlands Academy of Arts and Sciences to evaluate research organisations in Netherlands (see Section 2.3.10).

In order to assess the programme as a whole, common elements should be included in all centre reviews. The purpose of these common elements would be to enable comparisons between centres of excellence, thus providing an understanding of the distribution of different outcomes and impacts within the whole programme. For this reason, the peer-review panel probably needs both international experts in the given research field that participate only in individual centre reviews as well as Finnish experts that participate in all of the reviews to enable comparable assessments to be made.

Quantitative indicators could be employed to support the overall assessment as well as the comparisons between individual centres, although with reservations regarding the differences between the scientific disciplines. The indicators would serve at least four purposes: 1) the assessment of the centres' productivity in relation to typical research in the given scientific discipline, 2) assessment of the total impact of the programme, 3) assessment of the distribution of different outcomes across the centres, and 4) assessment of trends in the research organisation's productivity. The last point is related to assessing the impacts of funding. Because the main goal of the programme is to generate internationally leading research environments, this should also be noticeable as a trend of increasing the number and quality of outputs. The indicators employed could

address, for instance, publications in high-quality journals, educational impacts, participation in international collaborations, etc.

Evaluations on Individual-Level Funding Instruments

In the past, assessments focusing directly on the impacts of personal funding schemes have not been carried out by the Academy of Finland. However, researcher training and research mobility – which are central goals to many individual-level instruments – are one of the primary channels through which knowledge flows both internationally and between researchers and actors external to the scientific community. An impact evaluation could 1) provide information on the impacts of these instruments, 2) provide ideas as how to develop them further as well as 3) increase our understanding on knowledge transfer through researchers.

The focus on individuals has an important benefit to the assessment exercise. When the evaluation focuses on individuals rather than projects, the reconstruction of the influences from research activity is easier to accomplish, because the person "carries" the knowledge and skills with him/her. This mitigates the challenge of assessing the appropriation of benefits from research efforts typical to evaluation exercises with long-term focus.

Because long-term development of researchers is also a central goal to the personal funding schemes, it is suggested here that impact evaluation focusing on individual-level funding schemes should take a long-term perspective. The central issues addressed in the evaluation could include the following: How have researchers utilised the knowledge and skills learned with the support of Academy funding during their career? What have the benefits from international mobility been? Has international mobility resulted in improved opportunities, or possibly challenges, later in their career? How have researchers utilised their knowledge and skills when they have been employed outside academia?

The evaluation projects conducted on the instruments of Wellcome Trust, UK, and the Economic and Social Research Council, UK, provide ideas for the methodology of the evaluation. The analyses of publication output employed by the Wellcome Trust could also be used in evaluations of the Academy's funding instruments. The purpose of the analyses would be to provide information on the subsequent publication activity of the funded researchers, on the distribution of publications among the cohort members, and on international co-publication. Similarly, analyses on mobility, both internationally and between research and business sectors, could be conducted to reveal patterns of knowledge flow between the Finnish research base and actors external to it.

Interviews with researchers should also focus on the more qualitative issues. Illustrative and interesting examples of research knowledge utilisation could be identified and investigated more thoroughly. The emphasis should be on understanding the ways in which Academy-funded research has contributed to researchers' subsequent work. This could involve, for instance, investigating how researchers apply the skills and methods adopted during their research training, and how the international experiences have contributed to the subsequent career.

4.5 Summary

The impact evaluation activity at the Academy of Finland can be conceptually

structured by identifying three different levels as the objects of the evaluation activity: evaluation of the impacts of science at the national level, the impacts from operations of the Academy, and the impacts from the instruments that Academy employs in order to respond to societal needs. As a whole, the evaluation activity should be viewed as a system that provides strategic information and supports long-term development of the organisation.

The evaluation activities can be further characterised by dividing evaluation research into two domains based on the purpose they serve. First, evaluation can aim at measurement of impacts from research efforts, often initiated based on motivations external to the organisation. Second, evaluations can aim to enhance existing instruments through identification of factors that influence the processes with which impacts are generated, typically initiated based on internal motivations. The external and internal motivations influence evaluation activity at different levels to varying degrees.

When the twelve international evaluation projects are analysed from these different perspectives, thoughts and ideas regarding the development of the Academy's evaluation activity can be presented. The emerging ideas address both general directions towards which the evaluation activity at the Academy of Finland could be steered as well as provide concrete suggestions regarding evaluation projects that could be initiated.

At the national level, the indicators employed by the review of the state and quality could be complemented with indicators of research commercialisation in universities and research institutes to reflect that aspect of research utilisation. As a whole, the accent of the evaluation activity should be shifted towards identification of development needs in the research system. This would provide more information for steering the Academy's funding allocation towards areas with more potential for impacts in relation to observed societal needs.

At the level of Academy operations, the evaluation activity is influenced by the needs for developing indicators that measure the Academy's productivity and performance in relation to its target outcomes. The development process should emphasise validity, relevance and stability of indicators. This implies that the focus of the development process should be on investigating the significance of the selected indicators. Particular attention should be paid on the ability of the outcome indicators to reflect the utilisation of research and societal impacts more broadly. In order to complement the information provided by the selected primary indicators, a more comprehensive assessment of the Academy's outcomes can be conducted based on the project final reports.

At the instrument level, the main emphasis of the Academy's evaluation activity should be on qualitative assessment of research impacts and the mechanisms that enhance the generation of these impacts. In research programme evaluations, impact assessment should focus on factors and conditions that influence the utilisation of research by users. As for centres of excellence, the impact evaluations should create evidence for assessing how the centres have achieved their goal of catalysing vibrant research environments. Evaluations of personal funding schemes, in turn, should investigate how research funded by these instruments has contributed in the long term through the skills and knowledge acquired by the researchers to subsequent research efforts both in research organisations as well as in business enterprises.

References

About ESRC. http://www.esrcsociety-today.ac.uk/ESRCInfoCentre/about/ 15 May 2005.

Academy of Finland Strategy 2003. 2003. Academy of Finland, Helsinki.

Academy of Finland Research Programme Strategy 2003. Suomen Akatemian Julkaisuja 2/03.

ACST Home. http://acst-ccst.gc.ca/home_e.html 4 June 2005.

The Allen Consulting Group 2003, A Wealth of Knowledge: The return on investment from ARC-funded research, Report to the Australian Research Council, September 2003.

Arnold, E. & Balázs, K. 1998. Methods in the Evaluation of Publicly Funded Basic Research. Technopolis, Brighton.

ATP Awards Summary Data. http://www.atp.nist.gov/eao/atp_funding.htm 27 May 2005.

DFG Mission. http://www.dfg.de/en/dfg_profile/mission/index.html 15 May 2005.

Evaluation FWF: Impact Analysis. Background report 4.2. March 2004. Joanneum Research – Institute of Technology and Regional Policy, Vienna.

Final report of the expert group on "Benchmarking S&T Productivity". 2002. European Commission, June 2002.

Finnish Geosciences Evaluation Report. 2003. Publications of the Academy of Finland 14/03.

Funding Ranking 2003: Institutions – Regions – Networks. 2003. Deutsche Forschungsgemeinschaft, Bonn.

FWF Austrian Science Fund. http://www.fwf.ac.at/en/portrait/portrait.html 27 May 2005.

Godin, B. 2003. The emergence of S&T indicators: why did governments supplement statistics with indicators? Research policy 32: 679–691.

Government Resolution on the Structural Development of the Public Research System. 7 April 2005. Helsinki.

The Impact of the EU Framework Programmes in the UK. An independent report for the Office of Science and Technology by Technopolis Limited. July, 2004. Office of Science and Technology, London.

International Evaluation of the Academy of Finland. 2004. Publications of the Ministry of Education, Finland 2004:16.

Kestävä ja dynaaminen kumppanuus. Yliopistojen, tutkimuslaitosten ja yritysten välinen tutkimusyhteistyö ja tutkijankoulutus. 2005. Publications of the Academy of Finland 3/05.

Kostoff, R.N. 1998. Science and Technology Metrics. http://www.onr.navy.mil/sci_tech/special/354/technowatch/docs/050510_metrics_main_text.pdf 20 June 2005.

Kuitunen, S. & Hyytinen, K. 2004. Julkisten tutkimuslaitosten vaikutusten arviointi. Käytäntöjä, kokemuksia ja haasteita. VTT Tiedotteita 2230.

Linko, S. & Sulonen, K. 2004. Tutkimusraportoinnin ja siitä saatavien tiedeindikaattorien kehittäminen. Loppuraportti. Suomen Akatemia, Helsinki.

Main Science and Technology Indicators 2005/1. http://www.oecd.org/document/26/0,2340,en_2649_34451_1901082_1_1_1_1,00.html 25 June 2005.

Martin, B., Salter, A., Hicks, D. Pavittt, K., Senker, J., Sharp, M. & Von Tunzelmann, N. 1996. The Relationship Between Publicly Funded Basic Research and Economic Performance: A SPRU review. HM Treasury, London.

Molas-Gallart, J., Tang, P., Sinclair, T., Morrow, S. & Martin, B. 1999. Assessing Research Impact on Non-Academic Audiences. Final report submitted to the Research Evaluation Committee, Economic and Social Research Council. SPRU, Brighton.

National Strategy for Centres of Excellence in Research. 1997. Publications of the Academy of Finland 6/97.

NIST: General Information. http://www.nist.gov/public_affairs/general2.htm 1 June 2005.

Nursing and Caring Sciences Evaluation Report. 2003. Publications of the Academy of Finland 12/03.

Outcome Indicators 2003/2004. 2004. Foundation for Research, Science and Technology, Wellington.

PhDs in Finland: Employment, Placement and Demand. 2003. Publications of the Academy of Finland 5/03.

Report on 'Products, Processes and Services for Users' as an Outcome Indicator. September 2004. Foundation for Research Science and Technology. http://www.frst.govt.nz/evaluation/downloads/outcome/Outcome_Indicator_Products_Process_and_Services.pdf 25 May 2005.

Review of Wellcome Trust PhD Training: Career paths of a 1988–1990 Prize Student cohort. 2000. The Wellcome Trust, London.

Ruegg, R. November 2001. Performance Rating System for ATP-Funded Completed Projects. Unpublished report.

Opetusministeriön ja Suomen Akatemian välinen tulossopimus vuosille 2004–2006. 2003. Opetusministeriö, Helsinki.

PSA target metrics for the UK research base. October 2004. Office of Science and Technology, London.

Research in Business Disciplines in Finland Evaluation Report. 2005. Publications of the Academy of Finland 2/05.

Research Programme for Telecommunication Electronics II 2001–2003. Evaluation report. 2004. Publications of the Academy of Finland 2/04.

Research Programme on Biological Functions Life 2000. Evaluation report. 2004. Publications of the Academy of Finland 4/04.

Research Programme on Marginalisation, Inequality and Ethnic Relations in Finland. Evaluation report. 2004. Publications of the Academy of Finland 3/04.

Research Programme on Sustainable Use of Natural Resources (SUNARE) 2001–2004. Evaluation report. 2005. Publications of the Academy of Finland 6/05.

Riddle, C. 2003. Taking Stock of R&D Across Three Sectors. Taking Stock of University & College Commercialization Efforts: Advising on the Way Forward. A Study for The Advisory Council on Science and Technology. Preliminary Report. Advisory Council on Science and Technology, Ottawa.

Ruegg, R. & Feller, I. July 2003. A Toolkit for Evaluating Public R&D Investment: Models, Methods, and Findings from ATP's First Decade. NIST GCR 03–857, National Institute of Standards and Technology, Gaithersburg.

Salter, A.J. & Martin, B.R. 2001. The Economic benefits of publicly funded basic research: A critical review. Research policy 20: 509–532.

Science and Engineering Indicators 2004. http://www.nsf.gov/sbe/srs/seind04/start.htm 25 June 2004.

Scientific Research in Finland. A review of its quality and impact in the early 2000s. 2003. Oksanen, T., Lehvo, A. & Nuutinen, A. (eds.) Publications of the Academy of Finland 10/03.

Sharpe, A. & Smith, J. 2005. Measuring the impact of research on well-being: A survey of indicators of well-being. CSLS Research Report Number 2005–02.

The Societal impact of applied health research: Towards a quality assessment system. 2002. Council for Medical Sciences, Royal Netherlands Academy of Arts and Sciences, Amsterdam.

Standard Evaluation Protocol 2003–2009 For Public Research Organisations. January 2003. Vereniging van Universiteiten (VSNU), Netherlands Organisation for Scientific Research (NWO), Royal Netherlands Academy of Arts and Sciences (KNAW).

Suomen Akatemian Tieteen- ja tutkimusala-arviointien ohjeistus. 2005. Suomen Akatemia, Arviointitoiminnan ryhmä. Unpublished material.

Suomen Akatemian tutkimusohjelmien arviointiohjeistus. 2005. Suomen Akatemia. Unpublished material.

Suomen Akatemian Toimintakertomus ja Tilinpäätöslaskelmat 1.1.2003–31.12.2003. 2004. Academy of Finland, Helsinki.

Tassey, G. 2003. Methods for Assessing the Economic Impacts of Government R&D. Planning Report 03-1. National Institute of Standards and Technology, Gaithersburg.

Tulosohjauksen terävöittäminen. 2003. Valtionvarainministeriön työryhmämuistioita 9/2003.

Valovirta, V. 2001. Tieteenalojen arviointi Suomessa – Kehittämisen vai vakuuttamisen väline. Suomen Akatemian julkaisuja 6/01.

Webber, D. 2003. Protecting and Managing Intellectual Property: Patenting Decisions of Publicly Funded Research Providers.http://www.frst.govt.nz/evaluation/downloads/outcome/2003_Intellectual_Property_Report.pdf 25 May 2005.

Wellcome Trust: Our Work. http://www.wellcome.ac.uk/node7110.html 13 May 2005.

Annex A. Reviewed Organisations

Australia

Academy of the Social Sciences in Australia (ASSA)
Australian Academy of the Humanities
Australian Academy of Science (AAS)
Australian Academy of Technological Sciences and Engineering (ATSE)
Australian Research Council (ARC)
Department of Education, Science and Training

Austria

Fonds zur Förderung der Wissenschaftlichen Forschung (FWF)
Österreichische Akademie der Wissenschaften (ÖAW)

Belgium

Fonds National de la Recherche Scientifique (FNRS)
Fonds voor Wetenschappelijk Onderzoek – Vlaanderen (FWO)

Canada

Advisory Council on Science and Technology (ACST)
Canadian Institutes of Health Research (CIHR)
National Research Council (NRC)
Natural Sciences and Engineering Research Council of Canada (NSERC)
Networks of Centres of Excellence (NCE)
Social Sciences and Humanities Research Council of Canada (SSHRC)

Denmark

Danish Research Agency
Danmarks Grundforskningsfond
Det Kongelige Danske Videnskabernes Selskab

France

L'Agence francaise de l'innovation – The French Agency for Innovation (ANVAR)
Centre National de la Recherche Scientifique (CNRS)
Comité National d'Evaluation de la Recherce (CNER)
Observatoire des Sciences et des Techniques (OST)

Germany

Bundesministerium für Bildung und Forschung

Centrum für Hochschulentwicklung
Deutsche Forschungsgemeinschaft (DFG)
Hermann von Helmholtz-Gemeinschaft Deutscher Forschungszentren e.V. (HGF)
Landesministerium für Wissenschaft, Forschung und Kunst Baden-Württemberg
Max-Planck-Gesellschaft (MPG)
Wissenschaftsrat
Wissenschaftsgemeinschaft Gottfried Wilhelm Leibniz (WGL)

Greece

National Hellenic Research Foundation (NHRF)

Hungary

Hungarian Scientific Research Fund (OTKA)
Magyar Tudomanyos Akademia

Iceland

Rannsoknarrao Islands

Ireland

Forfas
The Irish Research Council for Humanities and Social Sciences
Irish Research Council for Science, Engineering and Technology (IRCSET)
Health Research Board
Royal Irish Academy

Israel

Israeli Ministry of Science & Technology

Italy

Consiglio Nazionale delle Ricerche (CNR)
Istituto Nazionale di Fisica Nucleare (INFN)
Istituto Nazionale per la Fisica della Materia (INFM)

Japan

Japan Science and Technology Agency (JST)
Japan Society for the Promotion of Science (JSPS)
Ministry of Economy, Trade and Industry
Ministry of Education, Culture, Sports, Science and Technology

National Institute of Science and Technology Policy
The Science Council of Japan

Korea

Korea Science and Engineering Foundation (KOSEF)

Luxembourg

Fonds National de la Recherche (FNR)

Netherlands

Consultative Committee of Sector Councils for Research and Development (COS)
Foundation Quality Assurance Netherlands Universities (QANU)
Koninklijke Nederlandse Akademie van Wetenschappen (KNAW)
Netherlands Organisation for Scientific Research (NWO)
SenterNovem
Vereniging van Universiteten – VSNU

New Zealand

Foundation for Research, Science and Technology (FRST)
Ministry of Research, Science and Technology (MORST)
The Royal Society of New Zealand

Norway

Norges Forskningsråd
The Norwegian Academy of Science and Letters
NIFU STEP

Spain

Consejo Superior de Investigaciones Cientificas (CSIC)
Oficina de Ciencia y Tecnologia (OCYT)

Sweden

Forskningsrådet för arbetsliv och socialvetenskap (FAS)
Forskningsrådet för miljö, areella näringar och samhällsbyggande (FORMAS)
Kungliga Vetenskapsakademien
Kungliga Vitterhets Historie och Antikvitets Akademien (KVHAA)
Swedish Agency for Innovation Systems, VINNOVA
Vetenskapsrådet

Switzerland

Rat der schweizerischen wissenschaftlichen Akademien/Conseil des Académies Scientifiques Suisses (CASS)
The Swiss National Science Foundation for the Promotion of Scientific Research

UK

Arts and Humanities Research Board (AHRB)
Biotechnology and Biological Sciences Research Council (BBSRC)
The British Academy
Council for Science and Technology (CST), Department of Trade and Industry (DTI)
The Council for the Central Laboratory of the Research Councils (CCLRC)
Department of Trade and Industry (DTI)
Economic and Social Research Council (ESRC)
Engineering and Physical Sciences Research Council (EPSRC)
Higher Education & Research Opportunities in the United Kingdom (HERO)
Medical Research Council (MRC)
Natural Environment Research Council (NERC)
Office of Science and Technology (OST), Department of Trade and Industry (DTI)
Particle Physics and Astronomy Research Council (PPARC)
The Royal Society
Technology, Economics, Statistics & Evaluation (TESE), Department of Trade and Industry (DTI)
Wellcome Trust

US

Advanced Technology Program (ATP) – Economic Assement Office (EAO)
Defense Advanced Research Projects Agency (DARPA)
Government Accountability Office (GAO)
Lasker Foundation
National Institutes of Health (NIH)
National Institute of Standards and Technology (NIST)
National Science Foundation (NSF)
Technology Administration (TA)
Research Institutes & Consulting Agencies
The Centre for Science and Technology Studies (CWTS)
Centre interuniversitaire de recherché sur la science et la technologie (CIRST)
Centrum för utvärderingsforskning (UCER)
CHI Research Inc.
Fraunhofer ISI
Georgia Institute of Technology – Office of Assessment
Georgia Institute of Technology – Technology Policy and Assessment Center (TPAC)

l'Institut national de la recherche scientifique (INRS)
Maastricht Economic Research Institute on Innovation and Technology (MERIT)
L'Observatoire des sciences et des technologies (OST)
Policy Research in Engineering, Science and Technology (PREST)
Research Evaluation and Policy Project (REPP), The Australian National University
Science Policy Research Unit (SPRU)
Swedish Institute for Studies in Education and Research (SISTER)
Technopolis Limited
TIA Consulting – Technology Impact Assessment
Zentrum für Europäische Wirtschaftsforschung GmbH (ZEW)

Other Organisations

The American Association for the Advancement of Science
EU Commission
European Science Foundation (ESF)
The National Academies
Organisation for Economic Co-operation and Development (OECD)
Platform Research & Technology Policy Evaluation

Annex B. Analysis Template

Name of the evaluation project:
Commissioned by:
Conducted by:

Background	
Purpose	
Research questions	
Objectives	
Evaluation topics	
Scope	
Unit of analysis	
Main level of aggregation	
Published reports	
Other remarks	

	Yes/No	Description
Interviews		
Questionnaire surveys		
Case studies		
Document analyses		
Bibliometric analyses		
Economic analyses		
Statistical analyses (other than bibliometrics)		
Expert judgement / peer review		
Historical tracing		
Sociometric/social network analyses		
Analytical / conceptual methods for modelling programme theory		

Topic analysed	Indicator used	Remarks	Major findings
Funding characteristics			
Outputs & impacts on performing unit			
Transfer, interaction & cooperation			
Exploitation & changes in stakeholder organisations			

The purpose of this study is to generate information on the methods and indicators with which the impacts of basic research, as well as the impacts of the operations of research funding organisations, have been assessed.

The study responds to the need for developing ex-post assessment of impacts of basic research funding in the Finnish context. It also provides information on the approaches of impact evaluation adopted internationally in organisations similar to the Academy of Finland.

The conclusions and recommendations of this study offer suggestions and possible new directions for developing and complementing current impact evaluation activities at the Academy of Finland. These suggestions are based on the ideas emerging from the analysis of international evaluation projects.

ACADEMY OF FINLAND

Vilhonvuorenkatu 6 • PO Box 99, FI-00501 Helsinki
Tel. +358 9 774 881 • Fax +358 9 7748 8299
www.aka.fi/eng • viestinta@aka.fi