European Commission

Evaluation of Innovation Activities

Guidance on methods and practices

Regional Policy

European Commission, Directorate-General for Regional Policy
Evaluation Unit C4
Marielle Riché
Avenue de Tervuren 41, B - 1040 Brussels
http://ec.europa.eu/regional_policy/information/evaluations/guidance_en.cfm#1

Please cite this study as TECHNOPOLIS GROUP & MIOIR (2012): Evaluation of Innovation Activities. Guidance on methods and practices. Study funded by the European Commission, Directorate for Regional Policy.

Cover page pictures: É Marika Sandell

Evaluation of Innovation Activities

Guidance on methods and practices

June 2012

Table of Contents

Table of Figures

Table of boxes

Preface

From 2007-13, the European Development Fund (ERDF) is investing €83.7 billion on innovation, nearly a quarter of the entire budget. In the next programming period, the wealthiest regions will be required to allocate 80% of ERDF funds to research and innovation, business competitiveness and the low-carbon economy.

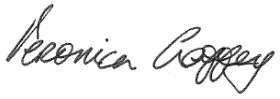
The next period will require an important effort by the European Union (EU) and the Member States, that co-finance the Structural Funds, if they are to contribute to the Europe 2020 strategy's objectives of "smart, sustainable and inclusive growth" to re-launch the European economy. The development of Europe's regions is driven by innovation, that is to say the creativity of researchers, the dynamism of entrepreneurs, and the effectiveness of governments and enterprises. It also depends on a smart specialisation strategy that takes account of a region's current performance and ability to compete in a challenging environment.

Yet, it is clear that the effects of ERDF investment on innovation are not sufficiently evaluated, whether it is the impact on direct beneficiaries, or more generally on the economy and society as a whole. The programme managers thus lack an evidence base that would enable them to improve their effectiveness and results.

In this context, the Directorate General for Regional Policy of the European Commission asked a consortium of Technopolis Group and the Manchester Institute of Innovation Research to assess the state of the art of methodological practices for the evaluation of innovation support, to analyse 15 evaluations of different types of ERDF co-funded measures, and to produce a methodological guide.

This guide is intended primarily for managers of ERDF programmes and those in charge of their evaluation. It is intended for an informed reader, familiar with innovation issues. It highlights the main questions to ask before developing technical specifications, and examines the pitfalls to be avoided, the advantages and limitations of certain methods, and the necessary conditions for a quality evaluation.

At a time when the new generation of programmes for 2014-2020 is being developed, this guide usefully emphasises the fundamental relation between a high quality programmes with clearly articulated objectives and identified means of achieving them, and the future evaluations that will analyse the results.



Veronica Gaffey

Head of the Evaluation Unit, DG Regional Policy, European Commission

1. Evaluation methods and innovation policy

Boosting regional innovation performance is a key European Union (EU) priority that will directly contribute to the Europe 2020 strategy. This goal is pursued, notably, through the Structural Funds. The Member State managing authorities are tasked with delivering and evaluating Structural Fund co-financed innovation measures. However, in some EU regions, the design and delivery of innovation measures is still a relatively novel form of policy intervention. Moreover, even in regions with a track record in innovation policy, the evaluation of innovation is far from straightforward.

Innovation is a complex phenomenon, difficult to quantify and with often long time lags before an impact can be measured. For these reasons, attribution (how much of the change is due to the policy instrument) can be a very difficult question to answer. While there is not a one-size-fits all evaluation method, this guide draws lessons from past evaluation experience. The aim is to facilitate an effective application of evaluation methods and thereby, improve innovation policy design and delivery.

1.1 Why a guide on evaluating innovation policy measures?

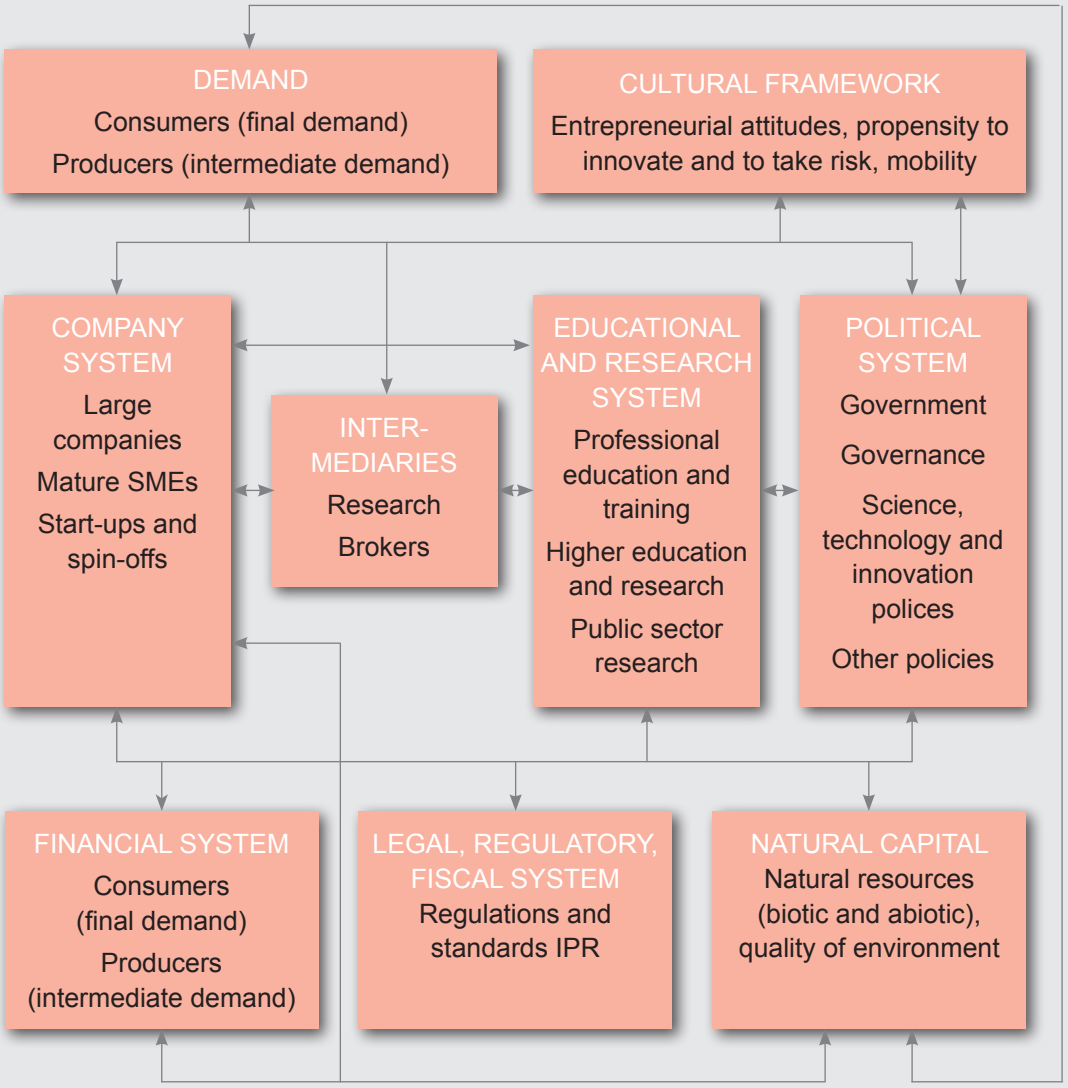
1.1.1 Innovation policy aims to improve performance of a broad system over time

Innovation is an ever-changing phenomenon. It takes place in a dynamic and constantly evolving system that is adapting to a range of internal and external factors. A mix of legislation, user needs, consumer demand, marketing strategies and new technologies and organisational practices drive innovation in manufacturing and service sectors as well as in social enterprises and the public sector.

Public policy interventions traditionally supported innovation in a single business through subsidies for research and development (R&D) or technology acquisition. However, successful innovation depends on interactions between a variety of public and private organisations, drawing on diverse skills and capabilities, including smaller and larger companies, universities, public agencies, business and innovation and financial intermediaries. Hence, innovation measures seek to motivate all these players, the ways in which they interact with each other and the collaboration modes they use to tackle socio-economic or, increasingly, environmental challenges.

Figure 1

The innovation system



Source: adapted from Arnold E. et al (2001) Evaluation of the Research Council of Norway

Thus, the measurable outcome of public intervention is more than new sales of an innovative product, reduced process costs or enhanced labour productivity. Firms may gain new partners, implement organisational change or acquire new methods and competence. Equally, there may be spillovers to other firms and society from, say, the diffusion of new technologies or organisational practices. In addition, policy makers seek to encourage investors to invest in riskier innovative ventures. Equally, they may attempt to raise the awareness of young people about careers in science, technology and innovation in order to ensure a future supply of skilled personnel.

1.1.2 A diverse range of innovation measures implies tailored evaluation approaches

To exert an influence on such a complex system, a broad mix of innovation measures is commonly implemented at regional and national levels[1]. This guide does not cover the full range of measures that could be deployed through Structural Fund co-financed national or regional operational programmes. Rather, the guide takes a more in depth look at five common forms of intervention in order to illustrate issues that may arise when launching and managing an evaluation. Figure 2 summarises the main modes and targets of funding under each of the five forms of intervention and lists some of the corresponding evaluations that were the subject of case studies (see Appendix B).

Figure 2

Overview of types of innovation measures

Type of innovation measure	Mode and target	Evaluation case studies
Science-industry co-operation networks and platforms	Funding allocated to consortia or joint projects involving enterprises and research or higher education institutes	• Danish Innovation Consortium Scheme
Strategic research programmes and research centres or infrastructures	Funding channelled to research institutions to conduct basic or applied research.	• Austrian Genome Research Programme (GEN-AU) • Irish SFI Centres for Science, Engineering and Technology (CSETs) and Principal Investigator (PI) programmes
Services (advisory, innovation management, technology transfer and training) to innovative firms	Funding of incubators, business innovation centres, business support networks, etc.	• Swedish National Incubator Programme • West of Scotland Science Park
Funding of innovative companies	Funding to businesses via grants, subsidised loans or guarantee mechanisms Provision of debt finance (loans, loan guarantee) or equity finance via venture capital funds and business angels for young innovative firms/start-ups	• Estonian Enterprise Policy 2007-13 • Flanders IWT R&D grants • Dutch Innovation Voucher
Cluster policies	Funding to cluster managers and/or groups of companies	• Danish Innovation Consortium Scheme • Finnish Programmes for Centres of Excellence in Research

[1] For a complete overview of regional innovation measures in the EU, see the Regional Innovation Monitor website: http://www.rim-europa.eu

The intervention logic of a measure explains how the support will assist the beneficiaries. For example, the anticipated results of an innovation voucher (often limited to testing the feasibility of an innovation) will be significantly different from those of an industrial R&D grant to develop a prototype. Similarly, an innovation voucher may stimulate interactions between a firm and a university researcher through a one-off contract. This is very different from the way a competence centre measure supports the development of a longer-term relationship between a number of businesses and an academic research team. The intervention logic should take account also of synergies with other measures (e.g. an incubator will depend on seed capital funds to support the growth of tenant firms). Hence, the indicators and the methods used to evaluate different measures will necessarily differ.

1.1.3 Specific challenges when evaluating innovation policy

Setting the scope of an evaluation.

Consider thoroughly the scope of an evaluation in order to design evaluation questions realistically reflecting the possible effects to be expected taking into account possible other factors.

A managing authority may wish to evaluate a programme, a portfolio of interventions (e.g. a priority axis), a group of similar projects (e.g. a measure), or a single a major project. This could be done at regional, sub-regional or national level. According to the selected scope of the evaluation, the approaches, methods and cost will vary. In the specific case of innovation policy, given the diffuse and distributed nature of the innovation process, the results of a measure are rarely constrained by geographical barriers. Hence, spillovers into adjoining regions may escape measurement, whilst spill-ins from other regions, or from national interventions, may affect the outcomes of the intervention being evaluated and how one chooses to evaluate them. Equally, the limited scale of regional interventions may mean that outcomes cannot be detected in regional economic statistics.

Attributing the effects.

For robust conclusions on the results attributable to a specific measure, use a mix of quantitative and qualitative methods in order to cross-check findings.

Secondly, the question of the attribution of effects (how to decide how much of a change in performance is due to a specific innovation measure) is complex. For instance, a growth in business innovation expenditure may be due in part to an innovation measure but will be influenced by a range of other policy decisions and measures. Equally, the propensity of businesses to innovate is subject to a range of external factors including the economic climate and the effectiveness of the innovation system in which they operate (access to information, to advice, to funds, to potential collaborators etc.). Separating the effect of an intervention from other factors is often difficult and requires, to increase the robustness of the analysis, a triangulation of evidence through a mix of evaluation methods. Moreover, the impact of an innovation support may result from one or only a few highly successful projects and this 'skew' effect needs to be taken into account when analysing outcomes.

Coping with time lags and unintended effects.

Outcomes of innovation measure are subject to varying time lags and may take varied routes: research may take years to become commercially exploited, clusters take time to develop and their composition may fluctuate, the effect on business performance may not be evident until after several business cycles. If a number of start-ups go bankrupt, an evaluator may conclude that the measure has failed, yet the people who worked in the start-ups or the technology they developed may yet benefit the regional economy in the longer-term. Conversely, a highly successful spin-off measure may be questioned if several spin-offs are later acquired by foreign firms leading to the intellectual property being appropriated and exploited elsewhere.

Appraise shorter-term results through formative evaluations and allow sufficient time before conducting an impact evaluation. Be open to unintended effects. Collect from the beginning of a programming cycle the baseline and monitoring information that captures changes in innovation activity and co-operation.

An ex-ante evaluation should assess the monitoring process and indicator system including the baselines, preparing the ground for future impact evaluations. During the programme, an evaluation may assess intermediate results and how they may lead to the expected result in the longer term. The approach should be open to unintended effects that may justify an adjustment of the programme. The evaluation commissioner should be explicit about the timing of excepted results and design evaluation questions accordingly.

1.2 Managing an evaluation: main steps and methods

This document complements the European Commission's overall guidance on evaluation for the 2014-20 period[2]. It builds on the standard conceptual framework for Structural Funds evaluations and assumes a basic knowledge of the main evaluation steps and methods. The guide may be used in conjunction with more detailed advice on specific methods available, notably, via the EVALSED portal[3].

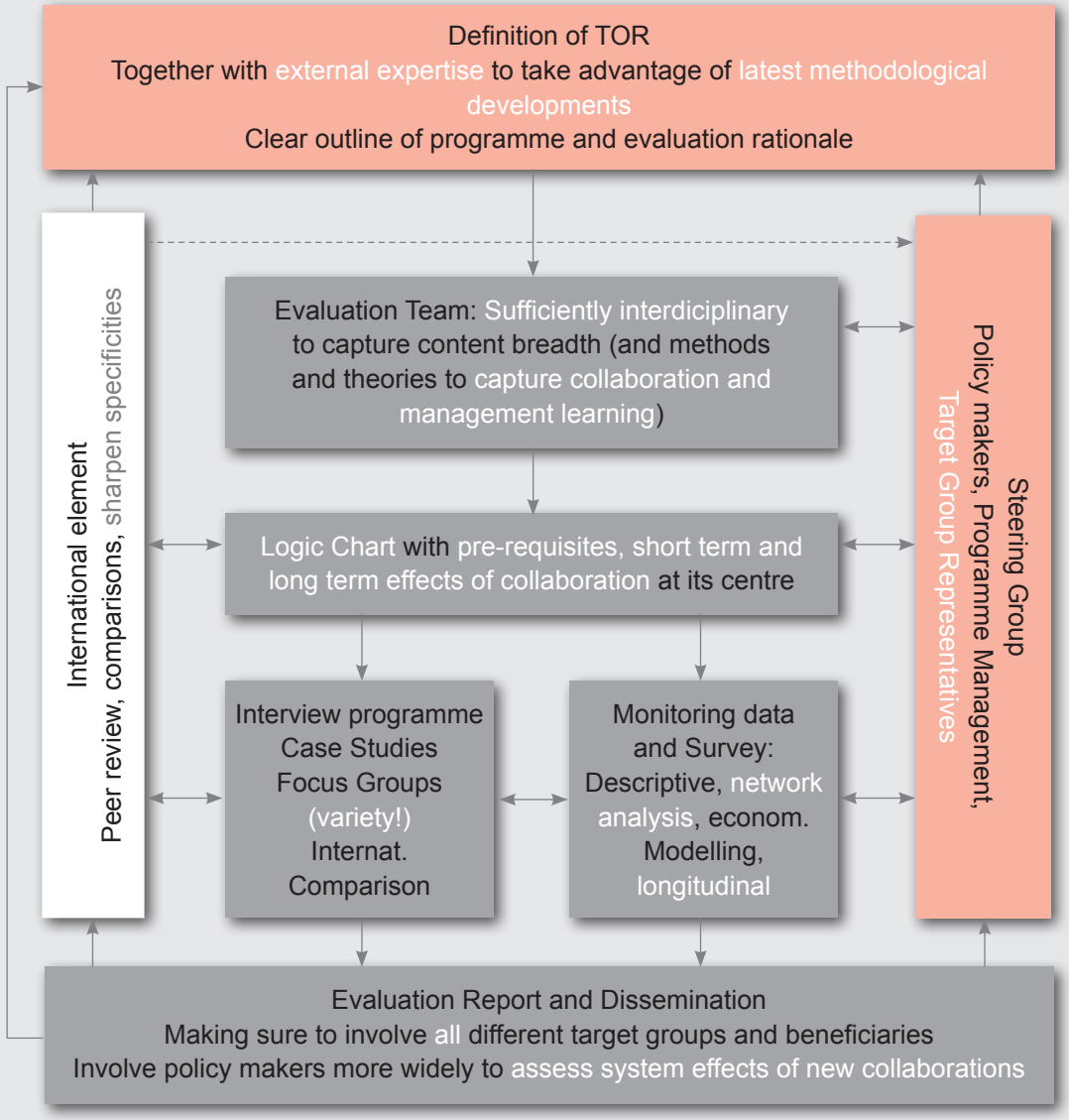
The main steps of an evaluation 'cycle' are:

- To ensure a cost-effective evaluation, the managing authority should establish, at the programme design stage, a limited set of key performance indicators as the basis for the monitoring system. However, many indicators are designed for monitoring and cannot be used directly for an evaluation of the effects or success of an intervention. Equally, if the intervention logic is unclear, the evaluators may need to reconstruct it and design new evaluation indicators, going back to available baseline data or conducting a survey. In certain circumstances, this retrospective task may prove difficult or very costly.

2 http://ec.europa.eu/regional_policy/information/evaluations/guidance_en.cfm#2
3 http://ec.europa.eu/regional_policy/sources/docgener/evaluation/evalsed/index_en.htm

Figure 3

Stylised process chart for innovation measure evaluations



- Decide on the scope (and level) of the evaluation and define the evaluation criteria and questions: should a single measure be covered or a 'portfolio', should the measure be assessed in its entirety or on the basis of its component projects.
- Manage an evaluation to deliver policy relevant outcomes: the selection of the evaluation team, budgeting[4] and timing, the involvement of stakeholders in the evaluation process, the interpretation of results into relevant policy lessons.
- Use a mix of methods to fit the needs and the context of each evaluation. There is no "magic bullet": no single method can address all the main evaluation questions or be applied across the range of policy measures.

[4] As absolute fee rates differ markedly among member states, the guidance expresses budgets in terms of the number of person days that might be needed to complete specific types of evaluation.

Rather than impose a specific tool kit of methods, the guide discusses those used in practice. In this way, management authorities and evaluation practitioners should be better informed about the advantages and disadvantages of such methods so that future evaluations may be performed effectively in various regional contexts.

1.3 How to use this guide

The heart of this guide is the five thematic sections addressing the specific approaches and methods applicable to each of the five broad types of measure:

- Science-industry co-operation (funding allocated to consortia, networks or platforms of business and higher education and research institutes)
- Strategic research programmes (funding channelled to research institutions)
- Innovation 'brokers' such as science parks, incubators and technology transfer offices providing advisory services to innovative firms (funding of intermediaries)
- Funding of innovative companies (direct financing of businesses via grant, loan/guarantee and equity modes of funding)
- Cluster policies (funding to cluster managers and/or clusters of companies)

Each of the five sections is structured as follows:

- A description of the innovation activities that are supported by the policy measure, which are as varied as the contexts and actors throughout Europe.
- A discussion of the intervention logic and expected result of the measures. Too often measures are not built on a clear intervention logic explaining the change sought and demonstrating how this will be achieved. This makes an evaluation challenging to design and implement. A significant improvement in programme design with clear objectives and indicators deriving from the intervention logic is needed for the 2014-2020 programmes.
- An overview of the main evaluation questions and indicators applicable. These questions and indicators are linked to the intervention logic and the objective of the intervention. As indicators only tell a partial picture: answering evaluation questions requests mobilising other and complementary sources of information.
- How to design and manage an evaluation for the type of measure, according to the focus of the evaluation (relevance, value for money, results and impacts).
- The main methods used to evaluate each type of measure and their pros and cons and possible other approaches.
- A summary of key 'pointers' to keep in mind during the evaluation.

The reader will find in annex a glossary of key terms used and a list of the case studies that were carried out to inform this guidance document.

The final report of the background study can be found at: http://ec.europa.eu/regional_policy/information/evaluations/index_en.cfm#2

2. Evaluating science-industry co-operation

2.1 What sorts of innovation activities are supported?

Most EU countries have a number of measures to support science and industry collaboration that fall into two broad types. The first type supports one-off, smaller scale projects, whereby researchers from a firm and research institute work together, with a clear division of labour, to achieve a scientific, technological or innovation objective. Such interventions (e.g. innovation vouchers) often focus on resolving a specific technological problem, or developing a prototype for a new product

At the other end of the spectrum, a second type of measure supports research institutes and firms to engage in longer-term strategic collaboration. These interactions can range from looser networking on a key enabling technology for the regional economy to the establishing and joint governance of a formal legal entity with firms and research institutes as shareholders. Examples of such measures include:

- Innovation networks or platforms where a condition for funding is that both science and industry stakeholders are involved.
- Competence centres supporting the development of joint research structures of firms and industry for a designated period of time.

Such measures differ in terms of their focus and in their distance to market, with some targeting pre-competitive or strategic research (e.g. some competence centres), and others supporting prototyping or marketable technologies.

2.2 What is the expected result of such measures?

2.2.1 Intervention logic

Innovation is increasingly dependent on the complementary knowledge and skills of different actors. Hence, co-operation and knowledge transfer are crucial for regional competitiveness but are also risky (e.g. requiring increased caution in handling intellectual property) and can create 'transaction costs'. In firms, these costs arise as external collaboration may disrupt 'normal' business activities and requires specific management skills (e.g. negotiating agreements, etc.). In higher education and research institutes, career incentive structures, partner searches and lack of "open to business" cultures often

impede external collaboration. Thus, although it is in the interest of research organisations and firms to co-operate, they do so less than is optimal. Public intervention helps to overcome such obstacles and reduce costs by subsidising the collaboration. Although the subsidy often covers only the additional costs of collaboration (and not actual R&D costs), this limited financial support can make the difference for small firms or even R&D units of larger companies.

The expected result is threefold and complex:

1. The actual output of an R&D co-operation project is a 'first order effect'.
2. A second important effect is a shift in the focus of firms and scientists towards more strategic (firms) and more problem oriented (science) R&D activities.
3. Most importantly, such measures develop collaborative skills and foster learning on how to engage in and sustain collective structures (post-intervention).

Such measures tend to follow a step-by-step logic: support for a partner search and preparation of projects, new networks or joint ventures, the project implementation leading to expected outputs and the subsequent academic, business and societal results. At the same time, they build absorptive capacity (to be able to learn from others), social capital (ability and willingness to engage), management skills (to steer complex projects) and thus induce a change in the routine of innovation activities.

Figure 4

Illustrative intervention logic for a science-industry co-operation measure

Inputs	Outputs	Results	Long term results
▪ Grants & financial instruments ▪ Advice	▪ Increased R&D investment leveraged by funded projects ▪ Newly established or extended networks and centres ▪ Scientists and engineers working on joint projects	▪ Increased patents / co-publications in specific technology field ▪ Prototypes developed; ▪ Enhanced capacity to manage collaboration projects in both science and industry ▪ Revenue from contract research or technological services	▪ Growth in sales &/or exports of innovative products/ services arising from collaboration projects ▪ Sustained increase in R&D investment in enterprises involved in past collaboration ▪ Increased share of private R&D funding in collaborative research centres ▪ Increased share of hi-tech manufacturing and knowledge intensive service jobs in regional economy

2.2.2 Evaluation questions and indicators

The evaluation questions and indicators should cover three main types of effects (1) concrete R&D outputs, (2) changes to collaborative practices and (3) improvements in R&D management. The relative importance of the criteria and indicators will differ in line with the focus and intervention logic of the measure.

Figure 5

Indicative evaluation questions & illustrative indicators – science-industry co-operation measures

Indicative evaluation questions	Examples of possible indicators
▪ Is there an overall level of increased research and innovation investment due to the collaboration?	▪ Additional euro spent on R&D due to the funded project by firms involved in collaboration
▪ Did the collaboration projects funded lead to high quality research results?	▪ Number of (co-)publications in peer reviewed journals and citation impact
▪ Did the collaboration result in increased innovation outputs?	▪ Number of patents, prototypes, new products/services compared to a pre-project baseline for the partners
▪ Did the projects lead to a sustained change in the type and frequency of science-industry collaboration?	▪ Number and type (bilateral, consortium, etc.) of collaborations before, during and after intervention ▪ Increased personnel mobility between research institutes and firms, etc.
▪ Did the interaction with industrial partners change the R&D management practices or orientation of research institutions?	▪ Share of applied research in total activity of research institution ▪ Increased revenue knowledge transfer (licensing, etc.)
▪ Did the projects enhance the innovation management capacities of businesses?	▪ Number of newly adopted innovation management practices, changed business models, change in recruitment patterns.
▪ Is there an observable economic impact in terms of new products or services?	▪ Share of turnover based on innovations arising from collaborative projects.

2.3 Managing an evaluation of a science-industry co-operation measure

Collaborative measures can have a major influence on an innovation system by improving the linkages between actors. Hence, when analysing the rationale, appropriateness and relevance of a measure, an evaluator needs to establish if the focus and forms of collaboration supported are justified with respect to the pre-existing patterns of co-operation. This may be done by considering the demand and need for collaboration in selected fields and evidence on the intensity of co-operation between firms, between firms and research organisations, etc. using a mix of official statistics (e.g. CIS data on co-operation, co-patenting and co-publication data), more informal data (participation of organisations to technology platforms, clusters, etc.), through interview with stakeholders or a bespoke survey. The evaluation results against target (effectiveness) should cover both the innovation output at firm and organisational level, learning within project, the broader, structuring effects on collaboration behaviour and the overall economic effects.

Collaborative measures should, ideally, be accompanied by a formative evaluation that helps the beneficiaries learn from what they do and how they do it. The distinctive characteristic of collaboration measures is the aim to trigger changes in behaviour (in project partners, in R&D management, in governance of collaborative structures). Such changes, as well as more tangible outputs, should be monitored and findings fed back to the participants and the measure managers to enable timely adjustments to improve the *effectiveness of implementation. A formative approach should also* help to improve the uptake of the measure and its longer-term effects (e.g. sustained ability and willingness to co-operate).

Timing is a key issue. While projects may generate concrete results within a couple of years, the broader economic effects and especially the impact on sustained collaboration and on joint R&D activities will occur in the longer-term. Hence, it is advisable to conduct several rounds of analysis together with beneficiaries and stakeholders in order to capture the overall results. There is also a need to check for negative side effects since a positive impact on institutes could, for example, be outweighed by detrimental effects on firms. For example, an institute may be better placed to negotiate terms on intellectual property making it difficult for firms to exploit commercially the results of a joint project.

Ideally, the terms of reference of a collaborative measure evaluation should be drafted drawing on advice from external experts and representatives of beneficiaries. This facilitates the assessment of collaboration effects, especially in terms of behavioural change and the added value of new research combinations. An evaluation of R&D collaboration measures should include peer review panels and an international comparison, requiring a clear rationale for the selection of comparator programmes and regions/countries. This broadens the pool of comparator programmes and takes advantage of specific expertise as regards research content and governance structures of collaborations. Given the focus on promoting learning amongst participants, the technical specification may require evaluation teams to include experts in disciplines such as cognitive psychology and organisational studies, notably for case studies. An evaluation may lead to a tool-box for programme managers to track behavioural changes in R&D management and governance more comprehensively.

Given the range of potential outputs and results of collaboration measures, the evaluation team must apply a comprehensive and rigorous methodology. However, there is a trade off between the sophistication of methods and usability. Even if methods such as network analysis or econometric modelling are designed and executed correctly, the interpretation of results of the analysis may need so much technical expertise, that their use in evaluation practice and subsequent policy-making is limited. Hence, it is advisable to be relatively prescriptive about the methods to be applied (partly to allow costing during the tender process). Given the constant advancement of qualitative and quantitative methods, the terms of reference should not exclude the possibility for additional or alternative methods to be proposed. As the evaluation design and programme effects are complex and multi-dimensional, a structured interaction between evaluators and programme management is important (see the Berlin case). Finally, an international evaluation tender is preferable if there is need for an external perspective and to compare with other systems and learn from them, especially in large, complex and ambitious programmes.

In terms of budgets, an evaluation with a limited scope that only reviews the specific results of collaborative projects may be undertaken over six months with an input of between 60-120 person days. However, an evaluation of outputs and results at several levels (organisation, networks, regional or inter-regional linkages) and the overall result on the regional economy will be more costly and may require up to a couple of hundred person days input (particularly if it a formative evaluation over several years)

2.4 Which specific methods are most relevant?

In order to evaluate shorter and longer-term results as well as broader economic and system effects, an evaluation of collaborative measures should combine quantitative and qualitative data. An appropriate mix of methods is involved: monitoring data on R&D expenditure and outputs, (recurrent) surveys (to build up a time series, for network analysis) and an analysis of learning and networking dynamics at a project and broader system level (interviews, case studies, etc.). R&D expenditure (e.g. trends in business funding of higher education research) and output data (co-publication data, etc.) can be analysed for beneficiary firms and institutes and also benchmarked against those for non-participants or between regions. However, this requires a broad, accurate and complete data set. Moreover, the analysis of monitoring data and of surveys on co-operation allows changes in input, output and collaboration patterns to be traced but not, necessarily, explained (see the Danish Innovation Consortium case).

Beneficiary surveys are useful in measuring changes in behaviour; however, they will not be sufficient to capture the full extent of changes or the spill-over effects of collaboration on the innovation management in the organisation. Participation in collaboration programmes might, for example, trigger changes in terms of the attitude towards collaboration or the training received may enable broader and sustained collaboration. Hence, surveys help to prepare the ground for follow up interview and case studies. Equally, surveys may be designed drawing on a set of pilot interviews.

A social network analysis (SNA) can be repeated at several points in time to capture changes in collaboration patterns due to the intervention. The analysis may identify non-participants (to assess alternative networking patterns) and it may compare networks found in other programmes or regions. However, evidence from evaluations that used SNA (Austrian GENAU evaluation) highlights that interpreting the relative effectiveness of a network is challenging. As the demand and opportunities for knowledge and innovation differ between networks, a simplistic comparison can be misleading. Hence, it is essential to embed SNA in a qualitative reflection with experts and beneficiaries, to interpret the meaning of certain network developments.

Policy learning can be enhanced through case studies that build on survey or SNA findings, to deepen the understanding of how co-operation capabilities have been improved by the measure. However, the representativeness of a case study approach may prove a challenge since no single new combination of actors will have the same features as another. Hence, the overall case selection should ensure a sufficient coverage of the main forms of collaboration and the underlying technological and innovation issues being tackled by the measure. To avoid simplistic generalisation on the basis of selected cases, the evaluation should also run focus groups at various points in time to discuss interim findings with beneficiaries and stakeholders, catalyse learning and adjust the evaluation approach as required. A steering group should ensure that complex methodological approaches are translated into everyday language and the focus on pre-conditions and behavioural effects of collaboration is retained.

A comparison group approach can assist in clarifying the net benefit of collaborative projects only if it is possible to define a control group of non-participating firms and organisations that are sufficiently similar in their structure and innovation activity. Hence, for specific and complex measures, such as competence centres, such an approach is unlikely to work. In this case, it is more instructive to benchmark with comparable cases in other regions or countries. Ideally, an evaluation should include an international peer review to facilitate comparison of specific features of collaboration, most importantly the requirements on the selection of partners and inter-disciplinarity, intellectual property and legal and governance issues of collaborative entities. The panel should comprise a mix of domain (technology, sectoral), R&D co-operation and governance expertise. As selecting the peers and running a review is challenging, the evaluators must have the necessary expertise (e.g. a network of potential peers to draw on, prior experience of managing panels).

2.5 Summary – practical tips

- Use a mix of methods and ensure a disciplinary breadth to measure not only direct results on business innovation or research results, but also on enhanced capacity to manage co-operation and on collaboration patterns in the broader system.
- Set up a clear baseline of collaboration prior to the intervention and try to maintain a monitoring system capturing collaboration over time.
- Apply appropriate methods, notably social network analysis, for mapping changes in collaboration patterns with care. Interpret the advantages and disadvantages of diverse networking and co-operation structures through a focus group with a strong involvement of beneficiaries and stakeholders.
- Appraise spill-over effects and unintended consequences (e.g. disrupting existing collaborations, limiting more traditional but effective collaboration, etc.)
- A counter-factual analysis is difficult as collaboration structures are often unique to a region or sector. Favour instead an international peer review and benchmarking to provide a comparative basis for policy learning.
- A key consideration when evaluating collaboration programmes is timing: many evaluations include sustainability as a question but there are few examples of a recurrent analysis over time or follow on analysis to assess long-term results.

3. Evaluating strategic research and technology measures

3.1 What sorts of innovation activities are supported?

User-oriented or applied research funding considered of strategic importance and undertaken within universities and public research institutes is often targeted by competitive funded measures. Such measures may address fields of national or international importance where there is an opportunity to expand regional capacities to a level sufficient to create an international comparative advantage, or focus on major sectors of the economy in which it is desirable to build capacity, to meet regional or national goals. Typically, intervention is in the form of grants that fund:

- the establishment of regional research centres that may be coordinated nationally (e.g. Finnish Centres of Excellence in Research)
- technology platforms and research consortia (e.g. Norwegian Functional Genomics Programme);
- one or more large-scale research programmes (Investigating Gene Function initiative or the Quantifying and Understanding the Earth System Programme both from the, UK; Flemish Strategic Basic Research measure)
- a range of cooperative projects, network projects, pilot projects, etc. (including support for accompanying research in other disciplines) (e.g. Austrian Genome Research Programme; the UK competitive 'Genomics' research measure).

Additionally, support may be used to establish a single awarding body tasked with the distribution of targeted strategic research funding (e.g. Science Foundation Ireland, set up to support strategic research in ICT and Biotechnology and related fields).

Transferring the knowledge generated into industrial applications is not always an immediate aim, however, it tends to form part of the rationale of such measures. Hence, the intervention may be broadened to include industrial users of the strategic research to increase knowledge diffusion and the exchange of ideas between the science base and the private sector (see section 2 above).

3.2 What is the expected result of such measures?

3.2.1 Intervention logic

The rationale for funding strategic research programmes and centres is that they are expected to deliver an expanded and improved supply of trained people, intellectual property and know-how that can form a platform or niche market for regional business development in the selected field(s). In addition, such measure may seek to coordinate and bring together fragmented pre-existing and new organisations to build coherence and attain critical mass and improve the quality and quantity scientific output. This contributes to raising capacity (in terms of research capabilities and know-how enabling firms to absorb or integrate new technologies) to a level where regional organisations may successfully participate in national or European initiatives.

Strategic research programmes and centres generally seek to ensure that the development of capabilities is not done in isolation but that the individuals, groups or institutions receiving support operate in a more networked manner (e.g. technology platforms or research pools). A longer-term expectation can be that increased regional capacities may lead to the co-location of significant (national or international) businesses thereby improving the regional economic structure and performance.

A long time horizon is one issue in evaluating this type of measure. The establishment or development of a critical mass of researchers and research activities may take several years and cycles of funding (including synergies with other types of innovation measure) in order to come to fruition.

Figure 6

Illustrative intervention logic for a strategic research measure

Inputs	Outputs	Results	Long term results
▪ Grants	▪ Completed research infrastructure: labs, etc. ▪ Installed equipments and instruments ▪ Doctoral and post-graduate training delivered ▪ Funded research projects ▪ Foresight and technology road map studies (user involvement)	▪ Industrial PhDs awarded ▪ Optimal usage of installed facilities (time equipment used as % of overall capacity) ▪ Increased research collaboration ▪ Attraction and retention of scientific personnel ▪ Scientific publications (highly cited) ▪ Patent applications / patents granted.	▪ Increased revenue for universities and research institutes from knowledge transfer (licensing, etc.) ▪ Increased participation by regional researchers in European or international research programmes ▪ Increased share of science and technology graduates in population. ▪ Greater share of user-driven research

To summarise, the major aims of this type of intervention are:

- to support research relevant to regional and/or national strategic priorities
- to support research training activities increasing the supply of scientific and technically trained people

- to develop and build research capacity through
  - the creation of research facilities and centres of excellence and
  - by developing critical mass, either within a single geographic location or, via networking and coordination, across a wider (national) area.

3.2.2 Evaluation questions and indicators

For each of the above aims, specific evaluation questions can be derived along with indicators by which they may be measured, either directly or indirectly (see Figure 7).

Figure 7

Indicative evaluation questions & illustrative indicators for strategic research measures

Indicative evaluation questions	Examples of possible indicators
▪ Is the research funded relevant to the socio-economic needs of the region (country)? ▪ Have the investments been made in fields where the region is, or has potential to be, specialised?	▪ Share of funding provided by industrial partners ▪ Active stakeholder involvement in user groups (number, functions of participants, etc.) ▪ Share of funding allocated per scientific/ technological field compared to current scientific specialisation profile
▪ Has the programme resulted in a sustainable improvement in research infrastructure?	▪ % of potential running time for which installed equipment is used ▪ Share of time equipment is used by researchers from other institutions (open access) and revenue generated from this usage
▪ Have the projects resulted in high quality scientific or technological results relevant to regional stakeholders?	▪ Number of (co-)publications in peer reviewed journal and citation impact; ▪ Share of research results exploited in follow-on projects with industry and other user groups
▪ Has the measure increased scientific and technological skills and specialisation in the region?	▪ Number of new Master/PhD graduates in the priority fields; ▪ % of new graduates employed in regional businesses or research institutes ▪ Share of scientific and technological personnel trained in priority technology fields

Indicative evaluation questions	Examples of possible indicators
▪ Have the research results led to economic, social or environmental benefits in the region?	▪ Increase in licensing revenue of research institutions or new technology based firms created using research results; ▪ New (foreign) investment in the region, e.g. business R&D facilities attracted by increased R&D capacity. ▪ Innovations with demonstrated environmental impact (e.g. reduced material or energy input).

The development of indicators and their monitoring by a funding agency will often require access to specialised scientific databases (Web of science, Scopus, Patstat, etc.). The use of such data avoids collecting information directly from participants (reducing the burden of reporting and costs of surveys) but implies the need for specialist skills in analysing the data. Some regions (e.g. Flanders in Belgium, see: http://www.ecoom.be/en) have set up 'observatories' to monitor such indicators

3.3 Managing an evaluation of a strategic research programme

Drawing on the evaluation cases, two stylised objectives can be suggested:

- To assess the overall effectiveness of the support in building a top-class research system, and to examine to what extent the programmes constituted value for money and efficient use of public funds, based on evidence on outputs and results arising from the implementation of the research activities.
- To assess the relevance of the programmes, their coherence with regard to national and EU research policies, and their efficiency, based on beneficiaries' satisfaction with programme management and monitoring.

Many of the decisions about the evaluation design will be determined by the overall scope of the evaluation: should it examine the impact of the programme on individual researchers (if these are the principal beneficiaries) or on research groups or larger aggregations such as institutions?

Evaluations that focus on relevance should be undertaken early in the programming cycle in order to feed the results back to the programme management. Relevance may be assessed by examining the degree of uptake of the support, canvassing the opinions of participants and relevant non-participants, and also the broader potential beneficiaries of the measure. Relevance may also be assessed through the examination of research outputs and other research related activities.

If the focus is on the effectiveness and/or efficiency, then it is generally better to undertake a mid-term exercise (particularly if the programme has a comparatively long life

cycle) or even *ex post*, in order to allow the generation of sufficient measureable outputs and results. For example, since research publications are one of the main outputs in this type of measure, then results that can be evaluated could be anticipated at a relatively early stage of the programming cycle. Alternatively, the outputs of research training (PhDs, trained researchers, etc.) will not emerge in suitable numbers until at least three years or more into the cycle, whilst the results of research facilities may take considerably longer to accrue. Thus, it is necessary to decide whether the evaluation focuses on changes in the publishing behaviour of the beneficiaries, on their longer-term behaviour, on the impact on the broader research communities, or on the economic impact on the regional or even national economy.

In the regional, and even the national, context, it is unlikely that several, successive strategic research measures will be launched for a specific field – they will more often tend to be 'one-time' interventions. However, it can be very useful to undertake a formative evaluation at an early stage of a programme (after the first call) in order to learn lessons (such as on programme management and administration), which may be applied to later calls or to other programmes in a different field. Equally, it is recommended to undertake a summative evaluation some time after the programme has finished (or in the case where a new research centre has been established) in order to capture results (such as publications, the formation of research networks, etc.).

If the purpose of the evaluation is primarily to provide justification (i.e. on value-for-money questions) to sponsors and politicians, for example, then a summative evaluation is more appropriate. These considerations of course do not preclude the use of a series of (more focused) evaluations designed to look at the different anticipated outcomes of the intervention over a longer time frame. A good example is that of the evaluations of the Danish Innovation Consortia, which proved to be cost effective.

In terms of budgeting, the use of sophisticated evaluation techniques and extensive data gathering exercises will dramatically increase the cost requirements but these may be offset by the implementation of robust monitoring procedures at the outset of the programme. The nature of the type of support is such that the demand for participants to supply regular updates on published outputs and other research related activities is widely accepted and can be applied at relatively little cost to both the participants and the programme management. Moreover, the data for the quantitative indicators is relatively easy to collect, mitigating the cost requirements: e.g. data on citation counts is relatively inexpensive to access thanks to the availability of sophisticated and user-friendly on-line bibliometric databases. Equally, data on career paths may be obtained through surveys although longer-term career tracking can become resource intensive.

Qualitative data (for example from peer review of outputs or peer assessments of research facilities and infrastructures) is more costly to obtain. However, the size of peer panels does not need to be large, unless their scope is broadened to include representatives of industry and other potential stakeholder groups. Likewise, quantitative and qualitative data on network activities, collaboration and engagement with stakeholders and other beneficiaries may be obtained through relatively simple (and inexpensive) on-line surveys provided sufficient information is available for the identification of the appropriate targets. Wider scale interviews, whilst more effective at generating qualitative data

Box 1

Evaluation of the Austrian Genome Research programme GEN-AU

The GEN-AU programme was set up in 2001 in order to boost research excellence, collaboration, the development of young researchers and the visibility of Austrian genome research, provided a highly useful and critical milestone for the programme management.

The design of the evaluation terms of reference can be considered good practice as it involved a moderated process that sought external advice. In terms of the evaluation approach, on the negative side, too little attention was paid to the systemic impacts of the programme whilst the positive outcomes of networking were over emphasised. Resource limitations also constrained the number of case studies.

These criticisms are outweighed by the use of a mixed methods evaluation approach (desk research, analysis of monitoring data, broad interview programme with participants and non-participants and instrument-focused case studies). Instrumental to the success of the evaluation was the use of a logic chart that allowed clear communication about the programme and its impact logic as well as a clear definition of the evaluation steps and dimensions. Other notable methodologies were the use of a social network analysis and a benchmarking with international comparators in terms of their design and management characteristics.

and for covering a broader and more finely tuned set of issues, are also possible as an alternative to surveys but entail larger costs. However, if the set of external stakeholders and beneficiaries is relatively small, the quality of information gained may outweigh the savings obtained through the use of a survey.

While a funding agency may be able to monitor the programme and run elements of a formative evaluation, assuming sufficient in-house capacity and expertise, if a degree of independence is to be demonstrated then a tender for external evaluators will be required. Some guidelines for the selection of tenders would include: a high degree of familiarity with the rationales for public support for strategic research and the research process; experience with the analysis of publication data and bibliometric approaches; use of international peers to assess research centres and research infrastructures; avoidance of approaches that overly rely on econometric techniques.

3.4 Which specific methods are most relevant?

The assessment of the relevance of the support should consider the broad context of the funding landscape by checking with stakeholders that there is a clear and important 'gap' in the provision of funding for the topic in question. It is assumed that the strategic relevance and need for the programme has already been assured. A combination of desk research and surveys (of existing research practitioners in the field or related areas)

Box 2

Impact evaluation of the Finnish centres of excellence in research

The evaluation of the societal impact of the Centres of Excellence (CoE) in Research, conducted in 2008, was commissioned by the Academy of Finland, and covered the first two funding rounds beginning in 2000. The strength of the evaluation lies in an integrated analysis of the operational achievements and the broader strategic outlook of the Finnish CoE in Research concept. It looked at the societal impact (as opposed to its impact on the quality of research outputs or on economic impacts), while placing it in the broader context of the national innovation system and assessing its significance for science policy making. This broad based approach posed significant challenges given the need to gather substantial data through a range of tools in order to reach reliable conclusions.

The methodology combined qualitative data gathering tools (a desk research review of self-evaluation material; interviews and workshops with researchers; case studies) and quantitative data gathering tools, including surveys and a review of statistical data on CoE units. The workshops held in the final stage of the study ensured that the conclusions were sufficiently robust and that the quantitative data fully utilised. They also served to strengthen the outcome of the evaluation by allowing the evaluator to present the findings and interpretation of the evidence and to gather feedback prior to the final conclusions.

are suitable methodologies. A desk research phase may be used to ascertain what the programme is actually funding and could also be used to identify the existence of similar, overlapping funds. Alternatively, a survey or interview programme with stakeholders may be used to gauge whether the measure filled an important gap and if it was still needed.

The same types of methods may be employed to assess effectiveness (and longer-term results) although interpretation of the results may differ according to the specific questions addressed. A typical methodology is the questionnaire survey. In its simplest form, the beneficiaries are asked to provide information on the papers and outputs arising from their funded research. They may also be asked if, and in what ways, their work has been improved as a result of the new funding source. Other issues might include whether the support has contributed to levels of research collaboration or if new research facilities/centres have contributed to their work. A sophisticated methodology to assess collaboration outcomes and impact is social network analysis (see sections 2 and 6). Broader issues can be addressed, such as whether recipients and other stakeholders consider the strategic research area has become a more cohesive and well-managed discipline, and whether it had become stronger internationally. Lastly, in order to give greater depth to the survey findings, an interview programme (with recipients and/or stakeholders) can be employed.

A range of bibliometric techniques (publication count, journal impact, citation count and co-publication analysis) can be used to judge the effectiveness of the support in generating scientific output. Such techniques allow a comparison of the research funded to publication outputs and citation rates in equivalent disciplines within the same

country and/or internationally. Alternatively, or additionally, the results may be used to compare publication data from the same cohort of researchers but from an earlier time frame (provided such data is readily available, for example, from a CV analysis or specialised databases). This may be useful in assessing the effectiveness of the measure in funding high impact research and in judging the impact on the international standing of the research teams. However, it may be necessary to double-check that the research outputs are, principally, due to the additional funding to avoid an attribution problem.

A common approach in the evaluation of centres of research excellence is that of a peer review. Typically, the visiting panel will contain a proportion of international experts in the relevant field of research to provide a benchmark with comparator organisations or measures. A case study approach can also be useful in providing a more detailed view of the activities undertaken by a centre of research and comparing it with a suitable example from another region or country. However, it is not always easy to identify suitable comparators and the results may have to be compared directly to the specific programme goals and the indicators derived from them in order to judge the performance of the centre in question.

The question of cost-effectiveness is rarely examined, mainly as the (economic) outputs of strategic research are hard to measure. One way is to calculate the ratio between the scientific outputs (highly-cited papers, doctoral students graduating, etc.) versus the eligible costs. Evidence may be collected using monitoring data and through a survey. This type of analysis is useful in estimating the comparative efficiency of two or more similar measures. The Value for Money Review of the Science Foundation Ireland is a comparatively rare example of such an evaluation although the cost of this 14-month study (€189,000) reflects the resource intensity of the scope and approach. The evaluators developed a set of output and short- and long term result indicators for the SFI funding activities. These were later adopted by the SFI to collect data on an on-going basis on input, outputs, and short- and long term results for the annual programme review. The data sources include: bibliometrics; monitoring data; counterfactual analysis based on survey of beneficiaries and non-beneficiaries; and case studies.

Information on the evolving direction of strategic research is crucial for informing future funding allocations and determining their management. In this respect, a peer review can play a significant role, on condition that the panels contain senior international experts in the field, including potentially the end-users of the research. The panels should receive well-organised and highly synthesised evidence and be given direct access, or the possibility to pose specific questions, to the researchers involved. An able secretariat or dynamic (but not dominant) chairperson (ideally both) can often determine the level of success of an evaluation in achieving its objectives. A less direct review may be accomplished through a benchmarking to compare the funding programme in question to similar programmes internationally. One way to do this is to compile the information required to inform the benchmarking exercise through desk research, usually by an external consultant. EU wide research and innovation benchmarking platforms provide a low cost means to scan for relevant programme. Another option is for the peers to utilise their knowledge of funding measures in their respective countries to examine the relative attributes of the programme and arrive at a series of conclusions and recommendations.

3.5 Summary – practical tips

Strategic research programmes and centres have easily definable and measurable outputs at the early stages of their development (i.e. research publications, trained personnel). However, such measures have a broader goal of capacity building and engagement with a community of stakeholders drawn from a wide range of sectors beyond the boundaries of the immediate research community. This makes it much harder to assess the longer-term results. The following key tips are suggested:

- Measuring the direct output of research activities (publications, etc.) provides only a partial picture. Hence, use a combination of indicators (publications, patents, licensing revenue, scientific and industrial collaborations, spin-offs, etc.) and methods (for example, bibliometric studies, case studies and network analyses) to assess the contribution of the overall strategic goals of the programme both from a point of view of scientific excellence and of socio-economic impact.
- Make use of peer review panels involving both national and international leading experts from the same or allied disciplines, particularly those with experience of similar measures or of establishing new strategic research centres.
- The full results of strategic research funding take time to develop and should be evaluated only at a stage when scientific and economic results can be expected (implying a three-to-five year time horizon at a minimum).

4. Evaluating support services to innovating firms: the example of science parks

4.1 What sorts of innovation activities are supported?

Governments may seek to improve the performance of their innovation systems through the creation of an intermediate layer of innovation support organisations. These organisations work to enhance the connectedness of the public/higher education research base, the business community and other stakeholders (e.g. financial institutions). The ambition is to boost the dynamism of the innovation system, producing a step change in the intensity of science-enabled innovations as well as improving the flow of commercial insight and capability available to researchers.

National and regional development agencies invest public funds in numerous types of connecting structures[5], of which the following are the most common:

- Science parks are business parks located at or close by a research university or institute. They are designed for high-technology firms and use co-location to universities to boost rates of innovation among tenants and thereby improve economic performance[6]. Public support takes the form of accelerated planning applications and coordination among various stakeholders (from utilities to local communities), on the one hand and some level of co-financing, on the other. Parks offer prospective tenants a much greater proportion of space within buildings configured for R&D work, as compared with the general business stock in a region. In addition, science parks are organised to encourage interaction between businesses and the academic community, providing a density of contacts that facilitates the cross-fertilisation of people or ideas one or two orders of magnitude higher than might be found in a conventional setting.
- Technology business incubators[7] comprise a single facility or small group of buildings located on a university campus or within an adjoining science park. They aim to accelerate the development of tech-based start-ups through an array of support resources and services, orchestrated by the incubator management team

[5] National and regional agencies may operate business innovation centres that provide more generalist support to businesses, e.g. through advice and consulting on market entry strategies or business-to-business networking to facilitate new innovation partnerships. These more general support structures typically provide a spectrum of assistance open to all businesses and are not the focus of this chapter.

[6] See www.iasp.ws/publico/intro.jsp

[7] See http://www.spica-directory.net

and offered both in the incubator and through its network of contacts. Much of this support is provided free or at a nominal cost and most cases rents are below the market price for such high-quality and highly-serviced premises.

- Technology transfer offices (TTOs) aim to improve the connections between the research base and business. TTOs are typically part of a university or research institute (they may be separate legal entities to overcome limitation in the university statutes on the giving of financial advice or the holding of certain types of investment) and are responsible for commercialising university-owned intellectual property through: attracting and assessing invention disclosures; patenting and other forms of intellectual property protection; licensing; spin-out company formation; material sales; managing seed funds. Some TTOs may also incorporate a function that helps researchers sell their time as expert consultants.

While increasingly co-located, each type of intermediary structure can be evaluated separately. For reasons of brevity, the rest of this chapter focuses on science parks.

4.2 What is the expected result of such measures?

4.2.1 Intervention logic

Science parks promote the economic development and competitiveness of regions and cities by creating a series of synergies, with public grants, buildings and science park management producing a sequence of benefits as shown in Figure 8.

Figure 8

Illustrative intervention logic for a science park evaluation

Inputs	Outputs	Results	Long term results
▪ Loans for capital investment ▪ Grants for strategic planning and management ▪ Grants for service provision	▪ Newly built or renovated buildings and equipped labs ▪ Fully equipped incubators ▪ Services delivered to tenant and incubated companies.	▪ Value of (inward) investment in R&D intensive facilities ▪ Growth in turnover and employment of incubated companies ▪ Enhanced capacity of tenant firms to manage innovation projects ▪ Increased collaboration between tenants and between tenants and regional universities	▪ Increased share of hi-tech manufacturing and knowledge intensive service jobs in regional economy ▪ Share of science park tenants and trend increase in regional business R&D investment

An additional argument for public support is a market failure in the provision of quality premises with specifications required by high-technology firms. Commercial developers tend to build and manage premises with a specification that suits the needs of the majority of businesses. Intelligent buildings with flexible and highly serviced spaces, wet labs and pilot plants increase development costs by an order of magnitude. The market will only very rarely provide this kind of accommodation without some public incentive. By contrast, a science park can provide many of these high-cost spaces and services on a shared basis, thereby bringing rentals closer to general market rates. Moreover, businesses will pay a premium for the increased business opportunities and brand image that follow from locating on a successful park. However, the initial capital investment needs substantial public subsidies and it can take 10 years for even the most successful parks to become financially self-sustaining.

4.2.2 Evaluation questions and indicators

Figure 9

Indicative evaluation questions & illustrative indicators for a science park evaluation

Indicative evaluation questions	Examples of possible indicators
▪ Is there evidence of a market failure?	▪ Companies reporting difficulties in renting equipped laboratories, etc. (industrial property surveys) ▪ Cost of park services compared to private market rates
▪ Has the park attracted and retained tenants by providing quality infrastructure and services?	▪ Profile of science park residents: R&D intensity, share of scientists and engineers in workforce, geographic origin of companies, etc. ▪ Rate of occupancy of facilities over time ▪ Rate of satisfaction of tenants (annual surveys)
▪ Has the park leveraged sufficient funds to be financially sustainable?	▪ Additional public or private (e.g. equity) investment secured by park residents
▪ Has the park had a wider impact on the regional economy and innovation system?	▪ Intensity of co-operation between park residents and university or public research teams (compared to regional average) ▪ Rate of growth of employment and value added of tenants (compared to regional average for similar firms) ▪ Number of spin-offs 'graduating' from incubators hosted in the science park

Based on the intervention logic, a number of specific questions may be used to judge the effectiveness and value-for-money of a science park:

- The volume of finance deployed overall and the share of private funding secured by early public investment as well as ongoing support and reliefs (financial capacity, financial gearing, financial sustainability).
- The quality of infrastructure and operations, which might be revealed by above-target lettings and rental income and high levels of tenant satisfaction with facilities, services and business opportunities.
- Higher levels of informal interaction and joint projects than would be typical in a similar group of businesses distributed across the region.
- The rate of innovation expenditure as a share of turnover of park tenants, and the resulting trends in employment and sales.
- The volume and rate of growth of the tenants, on average and in sum, as compared with technology companies more generally in the region.
- Unintended spillovers both positive ones such as major investments (e.g. attraction of a new research institute) and negative ones (e.g. environmental impact of increased congestion; fall in occupancy of commercial business parks).
- A positive rate of return on the public investment, i.e. compared with the rental income and taxes from the net additional growth in the economy.

4.3 Managing a science park evaluation

The scope of a science park evaluation will reflect whether the client has a policy or financial interest in a science park or many, and whether that park is relatively new or a long-established facility pursuing an extension and refurbishment programme. Moreover, there is a potential for these investments to attract substantial additional inward investment and ultimately transform the structure and performance of a region. From this perspective, it is desirable for evaluators to look at the effects on regional stakeholders as well as the changing international perceptions of the region.

Perhaps most importantly, parks differ markedly in scale from one to another. In numerical terms, the majority of parks are small, perhaps accommodating 20-30 businesses and providing an extension (grow-on space) for start-ups moving on from the university or research institute's incubator. These smaller operations might be able to transform the commercialisation of public research for a given institution, but only larger parks offer the prospect of transforming a regional economy.

The starting point for the design of any evaluation is the evaluation objective. Is this an early look at how things are working at a newly opened park, commissioned by the regional agency that provided the initial capital investment? Or is it a more substantive and comprehensive assessment of the social and economic impacts of a network of science parks commissioned by the minister for economic affairs?

A typical science park evaluation is likely to identify, count and judge achievements to date (the summative component of the evaluation) and offer recommendations as to how the science park should be developed going forward in order to improve its efficiency

or effectiveness (the formative component). In some cases, an evaluation may have an explicit requirement to draw more general lessons for policy makers or for new parks. Moreover, a good evaluation should satisfy the needs of several groups of key stakeholders, from regional politicians, to investors, to university councils, to the park operators and above all to the tenants and their academic collaborators.

In terms of the core questions, the evaluation of a science park, or science parks, is no different to any other evaluation of a publicly funded programme:

- Relevance. Was this science park the best solution to a demonstrable problem or opportunity, and is there still a case for public support in some form?
- Effectiveness. To what extent has the science park delivered on its objectives? Has it attracted a critical mass of high-tech tenants, working together and with the university on innovation projects? Has the rate of innovation increased? Has the park boosted income and jobs in high value, high tech sectors?
- Efficiency. To what extent was the science park established in an orderly and timely manner? Have anchor clients been recruited, is the management team in place, have bridging functions been set up, are support services fully operational and working well, etc? Are the capital and operational costs in line with expectations and comparable parks elsewhere?
- Impact. An evaluation must test the extent to which a science park is doing something additional for the area, that there has been a net increase in innovativeness as compared with the pre-existing situation or indeed the current situation in analogous areas of the region or country that have universities and tech businesses conducting themselves without such intermediary structures
- Displacement. An evaluation should identify unintended consequences, wherein, for example, a new park possibly blights other business districts in the region.
- Value for money. Is this is an appropriate initiative that adds value to the regional innovation system and does not displace other business activities. Do the benefits realised represent good value for money? Might it have been more productive to invest those funds in an alternative measure?

Budgeting for a science park evaluation will be contingent on a number of factors, such as the size of the park(s) and the scope of the evaluation questions. However, evaluation clients should recognise and take into account that there will be internal and external costs, even where work is being commissioned externally. The internal workload for running a substantive evaluation is unlikely to be less than 20 person days. The external evaluation budget, given well-defined nature of these structures and the availability of good performance statistics, will range from 30-50 person days and a three-month timetable should be enough to evaluate a smaller park.

An evaluation of five parks might cost twice as much as an evaluation of one park, assuming the client is content for any consultations or fieldwork to be done on the basis of sampling. The evaluation of a regional or national network of science parks would require significantly more time, 6-12 months rather than 3-4 months, to allow all key parties to contribute. Similarly, if the scope of the evaluation is to determine the regional added value of the park, this will be more demanding and might require double the investment in workload.

Box 3

Evaluation of the West of Scotland Science Park

The West of Scotland Science Park (WSSP) is a small park (around half a kilometre square) with around 30 tenants and 1,000 on-site employees, situated to the north west of Glasgow and providing links to the Universities of Glasgow and Strathclyde. The evaluation was an ongoing review, and sought to gauge views on strategic relevance and future development as well as counting basic economic cost and benefits and comparing those with other science parks in Scotland.

It was the first formal evaluation and was carried out in 2008, some 10 years after completion of the original development. It comprised desk research and a limited consultation of stakeholders and tenants. The study had a budget of around £40,000 and involved a team of four analysts working part time for four months. The contractor had previously carried out similar evaluations of other science parks in Scotland and was able to pass on the benefits of that experience, and pre-existing data collection tools for example, to the client, in the form of a more economical and insightful study.

Science park evaluations tend to be carried out by an external contractor on behalf of the public agency that 'owns' the scheme and provides the financial and management support to the park or parks. Outsourcing means that the managing authorities can commission work from organisations with substantial prior experience of assessing science parks as well as bringing greater independence than an internal team. There is also a case for European level procurement to attract bids from the relatively small number of professional evaluators able to give clients access to proven data collection tools and analytical frameworks that have been developed elsewhere.

Given how much is known about these structures' outcomes and processes, it makes sense for the terms of reference to specify the evaluation questions reasonably fully and to prescribe the overall methodology to a large degree. With the more substantive evaluations, there is latitude for competition among prospective contractors on the detailed design of for example the counterfactual analysis and the investigation of the science park's contributions to the wider innovation system.

4.4 Which specific methods are most relevant?

Monitoring data provides the baseline case for tenants and may also provide good time series data on some or all of the key metrics. Beneficiary surveys are typically used to extend the monitoring data to capture facts and figures on behavioural changes, innovations and innovation-related commercial achievements. Beneficiary surveys may include open questions to capture information on unexpected achievements. For instance, the evaluation should also consider the risk that tenants simply decant from one site to another to take advantage of subsidies and improved facilities and transport. In the evaluation of Polish Technology Parks the evaluators sought to test the impact of the

national network of technology parks by comparing the commercial performance of tenants with that of a control group. The statistical analyses proved inconclusive due to difficulties with matching the treatment and reference groups (see Box 4). Since establishing a control group is difficult due to the specific characteristics of park residents, the intelligent and qualitative use of a comparison group (e.g. comparing performance of park residents with firms with a similar profile in the region or with firms located in a science park in another region) is recommended.

Indeed, science park evaluations typically include a benchmarking exercise (see for instance the West of Scotland case), which draws on standard metrics tracked by most parks and often aggregated by national or international science park associations. These tend to focus on occupancy rates, turnover and employment levels, and provide a neat, economical means by which to compare basic operational performance. The diversity of European science and technology parks and their monitoring systems means these comparative analyses tend to miss the key performance dimensions of increased technological innovation and innovation-related income and employment.

This low-key approach arguably downplays any consideration of the science park's role in the wider innovation ecosystem and also tends to tackle the critical evaluation questions about impact and displacement simply by asking beneficiaries to self-assess the net effect of the science park / innovation support on their business. Assessing the impact requires a more comprehensive approach that takes account of displacement effects (e.g. firms moving from an existing location into the park, quantifying collaboration outcomes with regional firms, tracking the effect of the park tenant firms on recruitment and retention of skilled personnel in the region).

Box 4

The experience of evaluating technology parks in Poland

A 2008 study reviewed the impact of the development of technology parks and innovation on Polish business. The evaluation sought to quantify the impact on innovation and economic output by comparing the performance of a sample of enterprises located across the national network of technology parks and the performance of Polish technology businesses more generally. It began with a programme of desk research to typify the portfolio of national technology parks in terms of their size and services. The study team then ran a survey of the 352 business located at the selected technology parks and used the results from the 92 businesses that provided a complete return as the basis for preparing a matched sample of 71 businesses (size, age, sector) outside the technology parks. The comparative analyses revealed few statistically significant differences between the two populations in terms of income and employment, however the evaluators came to the view that the matching process had been somewhat problematic and that the two populations were perhaps not a good match. Moreover, the technology park population included established firms and start-ups and had therefore gathered what amounted to current facts and figures about innovation activities and income and expenditure, and these data (single point in time) were thought to be a source of variability and were less robust than time-series data would have been.

One challenge is tracing the evolution in a park's performance over the sort of period (5-10 years) needed for the facility to reach maturity. The natural rates of turnover in tenants and the churn of key staff within those businesses can cause difficulties in piecing together the story. Perhaps more importantly, some of the most successful businesses will move away into larger bespoke premises and can easily be 'lost' to the evaluation – a problem compounded by mergers and acquisitions, a common route through which young businesses are developed and ultimately grow. This may require evaluators to track firms through several stages of growth using public or commercially managed databases holding company information.

4.5 Summary – practical tips

A successful evaluation of the effectiveness and value for money of a science park will be more likely if one observes the following rules of thumb:

- A participative approach that involves public funders, park managers and tenant representatives in overseeing the conduct of the study. This include a commitment to include at least some questions in the evaluation that have been framed by managers and beneficiaries and to feedback the preliminary results in a public seminar with all stakeholders and otherwise openly publish the main findings.
- The central evaluation question is the impact of co-location on individual tenants' behaviour, informal relationships and new ventures. A science park must facilitate dramatically higher levels of business-to-business and business-to-university interaction, as compared with more widely distributed tech firms. If it is not delivering high levels of connectedness, it is just another business park.
- When it comes to the overall results of the science park, an evaluation should address the wider effects on the regional innovation system. This should include any unintended consequences, particularly the negative effects, by conducting a survey of key stakeholders, comparing data on occupancy, employment, etc. of other business and science parks in the region, etc.
- For evaluations of larger science parks or national programmes, there is arguably a case for investing in a more robust experimental methodology, with a before and after measurement of innovation related income and employment of tenants, and a control group of some sort, in order to test the counterfactual

5. Evaluation of innovation funding for companies

5.1 What sorts of innovation activities are supported?

Direct financial support to enterprises to undertake product development, enhancing product design, prototyping, process innovation, technology acquisition, organisational change, improvements to product marketing, etc. is possibly the most prevalent innovation measure in industrialised countries. In the EU, State Aid rules[8] limit the scope of subsidies to business R&D projects or equity financing, in early-stage financing of innovative firms, provided from Government funds. Whilst the forms of support vary, all aim to reduce the risk for businesses to innovate:

- grants covering a percentage of business R&D costs, which can be awarded for defined activities on either a first-come-first-serve or competitive bidding basis;
- soft loans provided either directly by a government agency or through commercial banks or other financial intermediaries. In some cases, such loans may be reimbursable only under specific conditions (e.g. in the event that a product development project is successful and the company generates new sales);
- government loan guarantees intended to facilitate the granting of business loans by commercial banks or other financial intermediaries reducing the need for businesses to provide collateral when applying for a loan.
- government support to seed capital, business angel networks and early stage venture capital funds, which may take one of several forms: creation of a fund-of-funds, co-investment, etc.

5.2 What is the expected result of such measures?

5.2.1 Intervention logic

The classic argument for Government support to business innovation activity is the existence of a 'market failure': a company that invests in innovation is unable to capture the full returns as it cannot stop other firms from copying or further developing the technology. This leads to a socially non-optimal level of investment in R&D as well as non-technological forms of innovation (design, etc.). Market failures and their impact on private incentives to undertake R&D are thought to be less critical the closer research is to the market. Public support to R&D tends to be less generous to firms the closer R&D is to the commercial use.

[8] http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=CELEX:52006XC1230%2801%29:EN:NOT

Hence, public funding of innovation projects aims to assist firms to do materially more development work than would be the case otherwise, producing more innovation (in turn resulting in increased sales/profits for assisted businesses, increased productivity gains, etc). Offering grants, credits, or loan guarantees directly to selected businesses (open to all comers, but selectively investing in the best proposals and teams) may be preferable to an R&D tax credit[9] in that it is targeted (on promising opportunities) and causes participants to work harder in pursuit of their project goals. Such selective assistance creates rivalry between the assisted and unassisted and should cause an upward adjustment of average R&D investment for a sector or region.

Figure 10

Illustrative intervention logic for a business innovation financing measure

Inputs	Outputs	Results	Long term results
▪ Grants ▪ Subsidised loans ▪ Equity financing (subordinated loans, seed capital, funds of funds, etc.)	▪ Increased business R&D investment leveraged by public funds ▪ Acquisition of new technology ▪ Equity (co-) investment in new or existing innovative firms	▪ New products or services launched ▪ New or upgraded production lines ▪ New hi-tech firms established ▪ Increased collaboration with universities, etc;	▪ Growth in sales and exports of innovative or hi-tech products and services ▪ Increased labour productivity rates ▪ Increased share of hi-tech manufacturing employment and knowledge intensive service jobs in total employment

Governments (whether national or regional) rarely explain a decision to intervene in such terms and often the intervention logic is not made explicit. In general, strategy and programming documents advance one or more of the following arguments:

- Insufficient levels of business R&D expenditure resulting in a low value added composition of the regional economy. The intervention logic is to tackle a perceived market failure, even if more complex factors may be at play.
- Outdated or inappropriate use of technology by regional firms leading to relatively lower (manufacturing) productivity rates. The intervention aims to assist firms to upgrade their production technologies and improve regional competitiveness
- Helping regional firms shift to higher value added activities or developing new business sectors. The intervention will often target export-intensive firms with a mix of R&D and marketing support. A second approach is to foster the creation and growth of new technology based firms through a mix of direct grants, equity financing and incubation and high-growth business services.

[9] R&D tax credits are not eligible for Structural Fund support. For a review of evaluation methods see: http://ec.europa.eu/invest-in-research/pdf/download_en/rd_tax_incentives_expert_group_report2008_rtd_final1.pdf

- Optimising the socio-economic impact of public funding for research by encouraging commercialisation of results through spin-off companies or licensing. The justifications for intervention include: academic incentive systems that do not reward commercialisation of results, low propensity to entrepreneurship, or an 'equity gap' that limits access to seed capital for high potential young firms. The forms of intervention including proof of concept, start-up grants, seed funding, support for technology transfer and incubators (see section 4), etc.

5.2.2 Evaluation questions and indicators

Indicators for assessing the outputs and outcomes of business innovation measures should permit an analysis of the quantitative impacts on business innovation activities and results and the verifiable changes in co-operation patterns within other actors in the innovation system. They should also facilitate a qualitative assessment of the extent to which the intervention has improved in-house capacities of firms to continue to invest effectively in innovation and to extend or strengthen knowledge acquisition and exchange (with other firms, etc.). Illustrative evaluation questions and indicators that may be used to focus an evaluation are set out below.

Figure 11

Indicative evaluation questions & illustrative indicators for funding for business innovation

Indicative evaluation questions	Examples of possible indicators
▪ To what extent is the measure focused on firms or sectors of the regional economy facing specific difficulties to innovate or with a specific potential? ▪ Is the measure reaching firms with a latent potential to innovate?	▪ % of firms assisted which previously reported negligible R&D or innovation expenditure ▪ Renewal rate (% of previously non-assisted enterprises supported)
▪ Is the public funding being disbursed using the least possible (human and financial) resources by the implementing agency? ▪ Are the application, selection and funding procedures managed so as to minimise the cost to beneficiaries?	▪ Managerial efficiency (e.g. management cost per euro disbursed compared to benchmark programme). ▪ Stakeholders assessment of programme management (qualitative) ▪ Satisfaction of beneficiaries with programme procedures (survey/interview returns)
▪ Has the funding provided generated additional innovation activity in the beneficiary firms? ▪ Have the projects outcomes improved competitiveness of the beneficiary firms?	▪ Trend in R&D intensity (R&D expenditure as a share of turnover) compared to baseline (pre-intervention) ▪ Trend in performance indicators such as sales from new products/services; growth in productivity, etc.

Indicative evaluation questions	Examples of possible indicators
▪ Has the funding induced learning and/or built capacity in beneficiary firms enabling them to maintain their innovation intensity? ▪ Have new co-operation linkages been developed between beneficiary firms (and/or with other innovation system actors: financial intermediaries, etc.)?	▪ Post-project change in innovation expenditure, retention or additional hiring of qualified personnel to manage innovation; ▪ Identification of new co-operation patterns (survey or monitoring data).

5.3 Managing an evaluation of business innovation support

Evaluations of business innovation measures seek to elucidate the impact of one or more funding measures on the innovation activity of the target enterprises. Often funding measures are launched as a 'suite' of support with, at least on paper, an inter-linkage between say a small 'innovation voucher' (used to develop an initial idea in collaboration with a public research institute) and a more substantive grant for a full R&D project. The resulting product may then benefit from further support for patenting or for export promotion. Similarly, a project to assist a firm to acquire and install new technologies may be twinned with a training grant or a consultancy support. Hence, it generally makes sense to evaluate a 'portfolio' of measures. Even, when the evaluation focuses on a single measure, it should consider the coherence of the intervention with other financial measures (e.g. a regional R&D grant may 'compete with' or complement national grants or R&D tax credits) and non-financial measures (e.g. the success of a business innovation measure may depend on the support provided by a network of intermediaries).

Focusing an evaluation on specific business sectors or geographic areas helps to limit the scope and allow a more in-depth assessment. However, at a regional level (or in smaller countries), the focus on a specific sector may lead to difficulties in conducting a robust counter-factual evaluation or to carry out statistical analysis (see the Estonian case). The choice of sectors needs to be carefully thought through since there can be a tendency to focus an evaluation on 'hi-tech' sectors irrespective of the overall intervention logic; which may on the contrary focus on increasing innovation 'across the board' including in 'low-tech' sectors. The latter often still dominate economies, in terms of jobs, etc. A geographic focus is relevant if a measure covers diverse areas ranging from urban zones through university towns to rural hinterlands. Hence, understanding whether a measure has successfully supported innovation in similar firms in different zones can be important for designing future interventions.

Where the evaluation covers a 'portfolio' of inter-linked measures this adds to the complexity of the analysis. An issue that can arise (see the Estonian and Polish cases) when the evaluation scope covers both grant and loan funding are the differing modalities and

intervention logics. Loans (and guarantees) are often delivered by commercial banks or other financial institutions so that beneficiary firms may only be vaguely aware that they have received public support, leading to a lower inclination to reply to surveys, etc. Equally, the firms supported may be less 'competitive' than those receiving grant funding: loan type measures typically target firms with a weaker financial position (e.g. absence of collateral, poor financial results) that find in securing financial support difficult. This should be kept in mind when forming a comparison-group.

An evaluation budget will vary depending on the scope (number of measures evaluated, number of beneficiary firms and total funds disbursed, time period covered, etc.) but also the choice of methods (e.g. an Internet based survey will cost less to administer than telephone or face-to-face interviews). Broadly speaking, the core evaluation work examining the policy relevance, coherence, etc. may need up to 10-20 person-days (depending on number of interviews, etc.) and the design and delivery of a survey a similar number of days. Depending on the mode adopted (telephone versus face-to-face), interviews with stakeholders and firms may take between one to three hours (plus time for writing up notes). When drawing up terms of reference, MA should calibrate the number of interviews to be carried out to the available budget.

A rough rule of thumb for budgeting would be a range from 60 to 200 hundred person-days depending on the scope of the evaluation. The time-scale for this type of evaluation ranges from a minimum of six months (for a smaller evaluation of a single measure or a limited number of total awards over a number of measures) up to a year. If a survey of beneficiary firms is to be conducted professionally then the evaluation time frame will need to be at least six months (see the Polish case).

Box 5

Key data to maintain on beneficiary firms

In order to facilitate surveys and or statistical analysis of beneficiary firms, the managing authority should ensure that there is a minimum set of data maintained on beneficiary firms:

- Identification:
  - Full legal name (as used in business registers, etc.)
  - Enterprise registration number and/or VAT number
  - Name, phone and email address of company representative (wherever possible)
- Baseline data (should be ideally gathered at stage of application)
  - Turnover / value added (in euro/national currency)
  - Employment (full time equivalents) (year preceding application)
  - R&D expenditure (and if possible percentage of R&D contracted externally)
- Outcome data (indicative)
  - Values for all baseline indicators updated to latest available year
  - Sales (turnover) from new products

Most government departments have neither the expertise nor the resources to carry out an in-house an evaluation of a business innovation measure. However, the case studies suggest that the quality of the monitoring data on supported enterprises is a key parameter influencing the capacity to undertake a cost-effective evaluation. In an 'ideal world', the funding agency will maintain a database on beneficiary firms including baseline information (collected at the application stage) allowing an evaluation team to study change in performance before and after the intervention.

In addition to prior experience in the evaluation of similar measures, the evaluation team should have expertise on the analysis of business and regional economic statistics, survey techniques, in-house business innovation processes, financial engineering, etc. Although an international comparison of the functioning and outcomes of similar measures can be useful when framing the evaluation methods and drawing conclusions, the need to survey and interview regional firms implies that the evaluation team tends to be nationally sourced for this type of evaluation.

5.4 Which specific methods are most relevant?

5.4.1 Evaluation methods

Relevant evaluation methods include one or more of the following:

- Use of structural business statistics or data from tax authorities (company accounts, etc.) to compare beneficiary performance over time with those of a comparison group of non-assisted enterprises. Such data can also be used to assess multiplier effects of the public funding (gross-value added, etc.).
- Bespoke surveys of beneficiaries (either a sample or the entire population) extended to a comparison group (or double comparison group, i.e. non-applicant and unsuccessful applicants) to allow for a counter-factual analysis.
- Counterfactual econometric analysis of micro-data (e.g. the Community Innovation Survey data from Eurostat) or national/regional panel data (e.g. the Mannheim Innovation Panel in Germany). An econometric analysis is generally only appropriate when an evaluation covers a large enough number of beneficiary firms for which a reliable and complete economic time series can be obtained
- An in-depth analysis of a sub-set of beneficiary firms applying interview/case study methods to understand the synergies between a innovation measures. This can be done by tracking over time companies that have received a 'package' of support (e.g. innovation voucher, grant for R&D, prototyping and follow on investment, training and export grants). By covering the full 'project life cycle' from the firm's viewpoint, the evaluation avoids the risk of project fallacy (assuming that a grant, which may only cover part of the product development phase, leads to a direct and verifiable outcome for the firm).

The application of econometric methods to look at how the performance of recipients of funding compares to a counterfactual situation is a technique that is still used correctly in only a handful of evaluations. The advantage of a counterfactual approach is

Box 6

Evaluation of equity financing measures

Government support for equity funding of (young) innovative firms is very different from grant/loan type funding. Hence, there is a need for a specific methodology to examine their performance and outcomes. Over the past decade, public policy has switched from taking equity stakes in innovative firms to support for 'hybrid funds': public sector agencies co-invest as limited partners (in addition to private investors) in privately administered venture funds.

Key evaluation questions that need to be addressed include

- Is the supported fund of sufficient scale to be viable? Funds that are too small are ultimately non-viable on account of the high proportion of their funding that is absorbed by running costs and their limited ability to provide follow-on funding.
- Has the public intervention succeeded in leveraging (multiplier effect) additional private investment beyond the level that would have been invested in under 'market' conditions? Constructing a counter-factual argument is difficult since it is likely that most early-stage funds in the region/country will have benefited from public support. Hence, the methods adopted will be generally qualitative through interviews with fund managers and invested firms to understand investment trends.
- Have the investee firms supported by hybrid funds performed more strongly than a control group of firms or with respect to regional (or sectoral) business growth trends, in general?
- Are the investments supporting a 'structural adjustment' of the regional economy towards higher value added (more productive) sectors or more knowledge-based firms (including in service sector)?

Various methods are available to consider the overall impact on a regional economy of public support for equity finance. For instance, the 2009 evaluation 'From funding gaps to thin markets: UK Government support for early-stage venture capital', used econometric techniques to quantify the impact of venture capital support, comparing the commercial performance (impact of financial support on subsequent employment growth) for almost 800 client businesses (using micro-economic data provided by six venture capital funds) with the equivalent performance for almost 8,000 unsupported businesses (a matched control sample). See: www.nesta.org.uk/library/documents/Thin-Markets-v9.pdf

the (potential) rigour and accuracy of the results. However, such methods are strongly dependent on the availability of sufficiently robust and complete (over-time) data on the use of different sources of funding by a large enough panel of enterprises[10]. Hence, in many cases, evaluators adopt an approach based on survey data of beneficiaries and comparison of beneficiary performance with the average performance of firms in their sector, etc. using standard economic statistics.

In general, most evaluations will use a mix (triangulation) of several of these methods in order to build up a complete picture of the extent the funding has fostered additional innovation activity and generated economic returns (value for money). Indeed,

[10] See http://ec.europa.eu/regional_policy/information/evaluations/impact_evaluation_en.cfm for information on a pilot study on applying the method to support to innovation and R&D.

where the evaluation budget permits, there is a strong case for complementing statistical and survey methods with case studies of specific supported enterprises. This method can help to clarify complementarities (coherence) between different measures (or measures administered at different levels of governance). A case study approach can also pinpoint behavioural changes within beneficiary firms and consolidate conclusions on sustainability (the ability of firms to maintain the innovation activity).

5.5 Summary – practical tips

- To limit survey costs and avoid 'bothering' beneficiary firms, maintain a database on beneficiaries including baseline business performance data.
- Extend the intervention logic of the measure(s) being evaluated to cover interactions with other financial and non-financial measures that may influence the outcome. The inter-play between regional funding and national measures (including tax incentives) should be taken into account.
- Be explicit from the stage of programme design, and consequently when framing evaluation questions, on the time scale in which the expected outcomes are likely to occur. Such 'time-bound' targets enable an evaluation team to focus effort on those that should have occurred in the period covered by the evaluation study.
- Apply a mix of methods avoiding only statistical or survey methods that may fail to identify 'behavioural' effects on the capacities of supported firms to innovate.
- Seek to understand through more in-depth analysis of a sub-set of firms, applying a case study approach, the outcome over time of a 'package of support'.
- If attempting an econometric counter-factual approach (you can have qualitative "theory-based" counter factual approaches), be careful to apply it to a sufficiently large group of beneficiary firms and to check for combined effects of different types of funding. If sufficient data is not available for an econometric analysis, at least attempt to compare beneficiary performance trends with those of the broader economy (sectoral level, age of firm, etc.).

6. Evaluating clusters policy

6.1 What sorts of innovation activities are supported?

Innovation clusters are groupings of independent undertakings (innovative start-ups, small, medium and large firms as well as research organisations) operating in a particular sector and region and designed to stimulate innovative activity by promoting intensive interactions, sharing of facilities and exchange of knowledge and expertise and by contributing effectively to technology transfer, networking and information dissemination within the cluster'[11].

Cluster measures are a policy intervention aimed at creating, mobilising or strengthening one or more clusters, often supported by specialised cluster managers[12]. Cluster measures can take the following forms:

- Funding a cluster organisation (with an office/cluster manager) to undertake activities to strengthen co-operation between businesses, intermediaries, etc.
- Funding projects involving organisations from a cluster initiative (or a set of clusters). Such interventions aim at boosting business innovation by influencing the intensity of co-operation.
- Strengthening the framework conditions for clusters development such as support to human resource upgrading; improving the business environment, research infrastructure; support to internationalisation etc.

6.2 What is the expected result of such measures?

6.2.1 Intervention logic

Cluster policies aim to stimulate innovation by addressing co-ordination and information barriers that prevent knowledge and technologies being diffused, transferred and used in the economy. Two main results are expected:

- the creation or an increased intensity of formal and informal co-operation and knowledge exchange between firms, research institutes, public agencies, etc.

[11] European Commission. (2006). Community Framework for State Aid for Research and Development and Innovation (2006/C 323/01).
[12] See for instance: http://www.cluster-excellence.eu

- An increase in the intensity and quality of innovation within a cluster that will enhance competitiveness through productivity increases and a stronger knowledge base resulting in a long-term boost to the regional economy.

More specifically, the intervention logic of clusters development measures will typically address one or more of the following expected results:

- Overcome barriers to cooperation that limit the frequency and intensity of interaction. In particular, measure aim to provide incentives to a greater flow of knowledge and technology between research organisations and enterprises.
- Induce behavioural change, e.g. increase the propensity of firms to collaborate with other firms or enter into partnership agreements with research institutes.
- Improve the innovative capacity of enterprises putting them in a position to commercialise know-how and intellectual property.
- Strengthen and expand user-oriented research and thereby accelerate technological breakthroughs in key areas.
- Develop a critical mass of innovation excellence in emerging areas with strategic potential for cluster firms.

Different types of cluster support measures will target specific combinations of the above results as summed up in the following table.

Figure 12

Illustrative intervention logic for cluster support measures

Inputs	Outputs	Results	Long term results
▪ Grants ▪ Loans (infrastructure investments) ▪ Advisory services and cluster management	▪ Increased investment in collaborative projects in the cluster ▪ Newly established and/or extended networks ▪ Joint innovative, marketing and export activities	▪ Enhanced capacity to jointly develop products and services by cluster partners ▪ Increased R&D co-operation between enterprises and associated research institutes ▪ Sustained increase in joint business investment by cluster partners	▪ Growth in employment, sales & exports attributable to the cluster partners (as share of regional total) ▪ Improved international business linkages

6.2.2 Evaluation questions and indicators

Cluster policy evaluations[13] tend to focus on the direct and intermediate effects of cluster interventions, given that they are more concrete and easier to measure, whereas long-term results on the wider regional economy and innovation system are diffuse and tend to be more difficult to capture. A challenge for the evaluation of cluster support is the selection of appropriate evaluation criteria and indicators to assess the effects of a cluster measure. They differ depending on the change that interventions aim to induce and include:

- Introduction of new processes and new products;
- Increased sales and productivity of cluster enterprises;
- Increased collaboration between firms or between firms and universities or public research institutes (change of behaviour).

Figure 13

Indicative evaluation questions & illustrative indicators for cluster measures

Indicative evaluation questions	Examples of possible indicators
▪ Is cluster support focused on current or future strategic sectors or technologies?	▪ Sectoral and geographic origin of companies involved in cluster.
▪ Does the cluster manager provide efficient and effective support to cluster firms?	▪ Management cost per euro disbursed (benchmarked against other cluster measures). ▪ % of cluster participants (firms, research organisations, etc.) actively involved
▪ Has the cluster measure led to sustainable new co-operation patterns between firms, research institutions and public sector organisations?	▪ Change in intensity of co-operation between cluster participants (compared to regional average for innovation co-operation) ▪ Recorded change in intensity of co-operation of cluster participants for innovation or business development/export.
▪ Has the cluster led to improved innovation potential and economic performance of participating firms?	▪ Number of new products/services developed through joint cluster projects ▪ Increased hi-tech exports of regional firms
▪ Is the cluster contributing to the attractiveness of the region as a location for R&D and innovation?	▪ Rate of growth of employment and value added of cluster participants (compared to regional average for similar firms) ▪ Additional R&D intensive inward investment

[13] See Schmiedeberg C. (2010) Evaluation of Cluster Policy: A Methodological Overview, Evaluation 16: 389

6.3 Managing a cluster measure evaluation

Given the multi-layered nature of cluster policies, there is a need to distinguish between two possible objectives when **scoping the remit** of an evaluation:

- The effectiveness of a cluster development measure (or portfolio of measures): the extent to which the measure(s) achieved the expected outputs and intended results
- The impact of cluster policy on one or more supported clusters: an analysis of trends in the performance of (firms constituting) the cluster over time.

The timing of the evaluation is a second key issue both in terms of the programme cycle and the development phase of the cluster. An interim evaluation of a cluster measure will support policy learning and provide insights into the role of the cluster manager, methods to increase co-operation amongst cluster actors, etc. However, it is generally not worthwhile to focus on tangible economic effects on the firms in the cluster at this stage. Since a measure may support clusters at different stages of development, it is possible to 'stagger' a series of specific evaluations of individual cluster initiatives (this was done, for instance, in Wallonia, Belgium). A 'meta-evaluation' can then sum up and cross-analyse the findings of the specific evaluations.

Depending on the scope and the timing, the evaluation may focus on:

- The efficiency and effectiveness of the cluster manager in implementing actions to strengthen the cluster (e.g. the number of new participants in cluster actions);
- The output in terms of new forms of co-operation within the cluster or between the cluster participants and other regional or inter-regional clusters in complementary sectors or technologies.
- The impact of the cluster measure on the innovation activity and resulting economic performance of the firms.

When drafting the technical specifications, a first step is to review the key parameters of the cluster measure and how they influence the scope:

- The range of cluster participants (SMEs, large corporations, research or higher education institutes, etc.)
- The nature of the sectors or technologies targeted by cluster initiatives
- Direct beneficiaries of funding (cluster managers; funding to groups of companies)
- Specific activities/ types of support covered by the measure (technology transfer, networking, information dissemination)
- Underlying policy objectives (increased competitiveness; cooperation; knowledge transfer and technology dissemination; etc.)
- Expected effects of specific actions (increase innovation activity; increase in contract research by firms with research institutes, IP licensing revenues, etc.)

In a second step, the types and level of results to be analysed should be fixed, e.g.:

- On company performance (at the level of individual enterprises)
- On co-operation between businesses and/or between businesses, research institutions and other supporting organisations.
- On the impact on regional competitiveness through the cluster(s) supported.

There are specific methodological challenges associated with the evaluation of the results of cluster measures, for example, when trying to:

- Capture the knowledge spillovers both within and beyond the cluster supported;
- Analyse intended as well as unintended effects of a cluster measure;
- Estimate the net-effects on the cluster and/or the regional economy.

Evaluating the economic impact on a cluster will invariably require a significantly greater time lag than the evaluation of a cluster development measure (providing funding to cluster managers), where behavioural changes can often be established shortly after the end of the intervention. Moreover, an evaluation of the overall impact of a cluster measure requires more sophisticated statistical methods (such as social network analysis) whereas an interim evaluation can be run through standard techniques including statistical analysis, interviews, surveys of participants, etc.

The cost of the evaluation will vary depending on the range of objectives and methods specified in the technical specifications but broadly speaking a cluster evaluation budget may range from 30-50 days (for a focused evaluation of a single cluster initiative) up to several hundred for an evaluation of a broader cluster programme.

Like other innovation measures, there is a basic need for programme managers to maintain a reliable database of baseline and monitoring information on activities and participants. In the case of cluster measures, the cluster manager may be tasked with collecting data on cluster firms, co-operation projects, new participants, etc. which would facilitate the evaluators task and help to limit the evaluation budget.

Ideally, cluster measure evaluations should adopt a participatory approach. Indeed, prior to the official launch of the evaluation, the evaluation design should be shaped through a stakeholder dialogue so as to foster a common understanding of issues and questions to be addressed. A participatory approach also implies stakeholder consultation during the early stages of the evaluation, which may take the form of semi-structured interviews with key representatives of cluster members. Ideally, the interviews should gather information on (a) operational and organisational issues, (b) the overall functioning of the measure, and (c) the expected results as viewed by different stakeholders. However, the need to ensure 'buy-in' and coordinate a participatory process may increase the overall cost of the evaluation study.

Evaluations of cluster measures can be demanding even for experienced evaluators as they require an understanding of the concepts and intervention logic underpinning clusters. The primary criterion in selecting an evaluator should therefore be the experience of the proposed team in conducting previous cluster policy evaluations. The evaluation

team should include people with an expertise in the sectors or technologies targeted by the measure.

6.4 Which specific methods are most relevant?

Two main groups of methods can be applied for the analysis of cluster policy. Firstly, the use of programme monitoring and reporting data, surveys, case studies and econometric analysis help to analyse how well the cluster measure met targets and can be used to analyse the extent to which individual cluster participants benefited from their involvement. Secondly 'systemic' approaches (including input-output or social-network analysis) may give a more complete insight to the question of whether the cluster has positively influenced the regional economy.

The analysis of **firm-level data** (from the applicant records or from official statistics) is a necessary starting point for evaluations of cluster development measures. Such data provides a baseline for the analysis of economic impacts. Normally, the cluster manager or the programme management will maintain such data or commission a specific survey as part of the scoping of the cluster initiative. If not, then it will be necessary to include it as part of the desk-research in the evaluation.

Available reporting and monitoring information on the cluster actions, whether it kept by the programme manager or cluster manager, is an important foundation for both implementing survey of actual and potential cluster participants, selecting case studies of actions funded and for the use of more advanced statistical techniques to analyse the comparative economic performance of the cluster firms.

Where applicant and monitoring data are not available, evaluations will depend on a survey of the beneficiary population, where a low response rate might jeopardise the overall validity of findings. A wide range of cluster stakeholders should be consulted to ensure representative findings. Online questionnaire surveys (sent directly to cluster members or by the cluster manager) are a cost-effective method and simplify the process of deployment and collection of results. The downside of closed question format surveys is that they are not as flexible as semi-structured interviews and do not provide an opportunity to explore the full range of qualitative 'connections' that may arise via cluster initiatives. Ideally, a sample of respondents should be selected for telephone or face-to-face interviews or for case study analysis.

Evaluations of cluster measures that attempt to assess the effects on the beneficiary firms may apply econometric (counter-factual) methods. Such methods can quantitatively test the effects of cluster policy (mainly on single actors within the cluster), however, data requirements and methodological capabilities are significant and positive results will be found only several years after the measure ends. Indeed, the complex interaction between actors in clusters and the lack of clear cause-effect relations is a particular challenge to evaluators of cluster interventions.

Box 7

Applying social network analysis in cluster evaluations

Social network analysis assesses the cluster as a social system. The cluster is characterised as a network of vertices and edges representing the actors within the cluster and the ties between them, respectively.

An SNA is based on an interaction matrix containing data on the relationships between the members of the cluster. The required data can be drawn from monitoring data (e.g. participants to projects, active members of web-based platforms, etc.), R&D collaboration or commercial relations between firms; surveys asking the actors about their relations with other actors; communication flows measured by email traffic, co-authorship or co-patenting for academia and science-industry co-operation.

Case studies can be a useful tool to help tell a story and highlight the experience of cluster participants. They can illustrate the dynamics and processes that drive knowledge exchange and diffusion of innovation in clusters, and flesh out the different ways in which cluster actions influence innovation in businesses. However, it is difficult to single out representative organisations within a cluster given the number and variety of participants and therefore to generalise from the results.

Other more advanced tools can be applied to assess the effect on the wider economy. A cluster measure may have multiplier effects on the regional economy that are difficult to capture in traditional evaluation models focusing on net-effect assessments in monetary terms. Social network analysis may provide quantitative results on cluster performance, but requires detailed, high quality data.

In contrast, benchmarking can give an insight into good practices and critical factors for cluster development. An **international panel** of cluster practitioners is a method commonly used. Such panels should ideally include a mix of policy makers, programme managers and (academic) experts who can contribute different perspectives. This method can provide an external, neutral judgement on the design of the measure and how it compares to similar measures in other countries (e.g. in terms of a comparison of cost-effectiveness, which is otherwise difficult to do). A panel can also provide valuable inter-disciplinary expertise throughout the evaluation process, which is particularly useful given the multi-dimensional nature of clusters. Ideally, panel members should have some experience of the evaluation of similar cluster interventions so as to help fine-tune the evaluation approach and review key evaluation findings. The quality and experience of the panel members is crucial for this tool to be effective. Such a panel has a relatively high cost (fees, travel costs).

6.5 Summary – practical tips

Cluster evaluations should:

- combine quantitative and qualitative research methods, e.g. a review of cluster development statistics complemented by a beneficiary survey, as well as beneficiary and stakeholder interviews that can be used to develop case studies, which probe into the quality of cluster interactions.
- be participative and ideally draw on the expertise of cluster practitioners, academics and policy makers. Evaluators should ensure that the opinions of the different stakeholder groups, notably business views, are captured and codified.
- Attempt to benchmark cluster development against that of clusters in the same country or from another country at a similar stage of development.
- Reflect in a realistic budget and timeframe the complexity of an impact evaluation of cluster interventions in terms of methodological design and research tools.

Appendix A Glossary

A.1 Evaluation concepts

Term	Description
Analytical technique / approach	Used for analysing different sets of data in order to draw out relevant findings and to address the evaluation questions.
Baseline	The value of the indicator before the policy intervention at stake is undertaken.
Common indicators	A list of indicators with agreed definitions and measurement units to be used where relevant in Operational Programmes, permitting aggregation to the national and EU level.
Evaluation	Evaluation is the systematic collection and analysis of information about the characteristics and results of programmes and projects as a basis for judgments, to improve effectiveness, and/or inform decisions about current and future programming.
Evaluation criteria	Public interventions in different policy fields must be assessed based on the specific results they are designed to achieve, hence the evaluation criteria applied must be specific to the policy area (e.g. innovation).
Formative evaluation	An evaluation that is intended to support programme actors, i.e., managers and direct protagonists, in order to help them improve their decisions and activities. It mainly applies to public interventions during their implementation (on-going, mid-term or intermediate evaluation). It focuses essentially on implementation procedures and their effectiveness and relevance.
Impact	The change that can be credibly attributed to an intervention.
Indicator	A variable that provides quantitative or qualitative information on a phenomenon. It normally includes a value and a measurement unit.
Input	Financial or budgetary resources mobilised for the implementation of an intervention.
Intervention	Any action or operation carried out by public authorities regardless of its nature (policy, programme, measure or project). The term intervention is systematically used to designate the object of evaluation.
Intervention logic	Shows a hierarchy of objectives and their assumed cause-effect relationship. The aim is to align inputs, activities and outputs with the result to which the intervention is intended to contribute.

Term	Description
Methodological framework / Evaluation methodology	A methodological framework establishes the parameters for an evaluation study based on the primary focus of analysis. The term refers to the combined set of research tools and analytical techniques used in an evaluation study. Methods are families of research tools and analytical techniques that fulfil different purposes. They usually consist of procedures and protocols that ensure systemisation and consistency in the way evaluations are undertaken. Methods may be quantitative or qualitative and may attempt to describe, explain, predict or inform action. The choice of methods follows from the nature of the intervention, the evaluation questions being asked and the mode of enquiry – causal, exploratory, normative, etc.
Output	Product of the intervention's activity.
Output indicator	An indicator describing the 'physical' product of spending resources through policy interventions. Examples are: the length, width or quality of the roads built; the number of hours of extra-teaching hours provided by the intervention; the capital investment made by using subsidies.
Programme objectives	Desired results of the intervention.
Research / data collection tool	Used for gathering qualitative or quantitative data for later analysis using appropriate *analytical techniques*. Research tools gather evidence that allows the evaluator to address the evaluation questions.
Result	The specific dimensions of the well being of people that motivates policy action, i.e. that are expected to be modified by the interventions designed and implemented by a policy. Examples are: the *improvement in mobility* pursued by building transport infrastructures; the *increased competence* pursued by providing additional or modified training; the *reduced rationing of SMEs* pursued by providing them with subsidised loans.
Result indicator	An indicator describing a specific aspect of a result, a feature that can be measured. Examples are: the time needed to travel from W to Y at an average speed, as an aspect of mobility; the results of tests in a given topic, as an aspect of competence; the share of firms denied credit at any interest rate, as an aspect of banks' rationing.
Study design	Comprises the methodological framework (research tools, analytical techniques and overall evaluation methodology) as well as the work plan (timing of tasks, milestones, deliverables, resource allocation) for an evaluation study. It also links the research tools, analytical approaches, and methodological frameworks to the evaluation questions to be addressed.
Summative evaluation	A summative evaluation examines the effects of a measure by describing what happens subsequent to delivery; assessing whether the measure can be said to have caused the outcome.

A.2 Data collection tools

Method	Description
Bibliometric or patent database studies	Searches of scientific publications (and sometimes their citations) and patents from bibliometric and patent databases.
Document and literature searches	Use of documents and literature directly or indirectly related to a programme. May include, for example, administrative manuals, application forms, assessment forms, existing evaluation reports and broader policy reports.
Focus groups, workshops, group meetings, etc.	A panel of people selected for their knowledge on a topic of interest, brought together to discuss the topic with the assistance of a facilitator. The discussion is used to identify important themes or to construct descriptive summaries of views and experiences on the focal topic.
Non-participant interviews	Interviews (either face-to-face or by telephone) conducted with those who have not participated in a measure (e.g. recipients of funding) or who have not benefited from the activities or services provided by a measure. May involve a structured interview format but allows scope for investigating issues that arise during the interview itself.
Non-participant surveys	Surveys conducted with those who have not directly participated in, or are not the main intended beneficiaries of, a measure. Usually involve the completion of a structured questionnaire (paper or on-line).
Participant interviews	Interviews (either face-to-face or by telephone) conducted with those who have participated in a measure (e.g. recipients of funding) or those who have benefited from the activities or services provided by a measure. May involve a structured interview format but allows scope for investigating issues that arise during the interview itself.
Participant surveys	Surveys conducted with the participants or beneficiaries of a measure. Usually involve the completion of a structured questionnaire (paper or on-line).
Peer reviews	Evaluation or assessment of programme activities or programme outcomes/outputs involving qualified individuals within the field.
Use of existing monitoring data collected during programme lifetime	Use of data and other information relating to the programme's administration, activities or performance systematically collected during the lifetime of the of the programme, usually by the programme management or administration.
Use of existing surveys or databases	Generally collected for purposes external to the evaluation and the measure (e.g. Community Innovation Survey data, opinion polls, business expenditure surveys, etc.).

A.3 Analytical techniques and approaches

Method	Description
Before/after group comparison approaches	Approach that compares data on participants/ beneficiaries collected before the intervention with that collected after the intervention.
Case studies	Methods of inquiry that focus on detailed data collection and analysis and which focus on a restricted number of participants/ beneficiaries.
Cost benefit approach	Procedure for determining the economic efficiency of a programme, expressed as the relationship between costs and outcomes, usually measured in monetary terms.
Counter-factual approaches	Approach that compares the state where no intervention has (or is assumed to have) taken place and the state where there has been an intervention. This approach typically uses a comparison group of non-treated units in order to be able to estimate the impact.
Descriptive statistical analysis	Use of basic descriptive statistics to analyse the data (e.g. uptake analysis, meaning the extent to which target beneficiaries have taken up the support provided by an intervention/ support measure).
Input/output analysis	Method used to characterise economic activity in a given time period, and to predict the reaction of a regional economy to stimulation, for example, from increased consumption or changes in government policy.
Micro-economic modelling	Micro-economic modelling refers to modelling behaviour/ performance of individual economic actors, most often businesses but also households, consumers, etc. In the context of evaluation, micro-economic modelling would be used to try to understand the effects (or lack thereof) of public interventions on the behaviour of a business (or other economic actors). The usefulness of the model depends on whether it can be generalised.
Network analysis	Analysis that aims to map the social interaction between the subjects of an evaluation including the beneficiaries.
Other econometric analysis	The use of other techniques drawing on advanced statistical methods such as regression analysis.

A.4 Glossary of innovation policy terms & concepts

Term	Description
Citation impact	Citation is the process of acknowledging or citing the author, year, title, and locus of publication (journal, book, or other) of a source used in a published work. Such citations are counted as a measure of the impact of the cited work. The citation impact factor is the average citation count for a journal.
Contract research	Contract research constitutes an important source of research revenue for higher education or public research institutions. It is research funded by external sources including government, foreign funders, and industry or charitable trusts.
Organisational studies	Organisational studies investigate the impact that individuals, groups and structures have on behaviour within an organisation. It is an interdisciplinary field that includes sociology, psychology, communication, and management.
Proof-of-concept	Research results or an innovative idea often need to be tested or turned into a prototype before entering the market. A proof of concept grant funds the steps necessary to establish the commercial viability of a new product, process or service.
Technology road map	A technology roadmap is a plan that matches short-term and long-term goals with specific technology solutions to help meet those goals. It is a plan that applies to a new product or process, or to an emerging technology. It helps reach a consensus about a set of needs and the technologies required to satisfy those needs; it provides a mechanism to help forecast technology developments and it provides a framework to help plan and coordinate technology developments.
Soft loan	Financing that offers flexible or lenient terms for repayment, usually at lower than market interest rates. Soft loans are provided customarily by government agencies or via financial institutions.

Appendix B Case studies

MS	ERDF co-funded	Evaluation title	Year pub-lished
AT	no	Interim programme management evaluation Austrian Genome Research Programme	2005
BE	no	A look into the black box: What difference do IWT R&D grants make for their clients?	2006
CZ	yes	Assessment of economic effects and the programme settings of the support programmes Innovations, Cooperation and Potential within the Operational Programme Enterprise and Innovations (OPEI)	2011
DE	yes	Evaluation of the Berlin Innovation and Technology Support	2010
DK	no	An analysis of firm growth effects of the Danish Innovation Consortium scheme	2010
EE	yes	The impact of the State's enterprise support on the competitiveness of the Estonian economy	2010
FI	no	Impact Evaluation of Finnish Programmes for Centres of Excellence in Research 2000-2005 and 2002-2007	2009
FR	yes	Thematic evaluation of the ERDF Operational Programme (OP) and the State-Region Programme (CPER) in Franche-Comté – Innovation, Research and Technology Transfer	2010
HU	yes	Ex-post evaluation of three measures of the Economic Competitive-ness OP	2010
IE	no	Value For Money Review of the Science Foundation Ireland	2008
NL	no	The effectiveness of the Innovation Voucher 2004 and 2005: effect on innovative inputs and innovative output of companies	2007
PL	yes	Ex-post evaluation of enterprise support instruments within the Integrated Regional Operational Programme (IROP) and the Sectoral Operational Programme 'Improvement of the Competitiveness of Enterprises' (SOP-ICE) in the 2004-2006 programming period (West Pomeranian Region, Poland)	2010
SE	no	Mid-term evaluation of the Swedish National Incubator Programme	2008
UK	no	Evaluation of the Grant for R&D and SMART	2009
UK	yes	Evaluation of the West of Scotland Science Park	2009

The full case studies can be downloaded from:
http://ec.europa.eu/regional_policy/sources/docgener/evaluation/pdf/eval2007/innovation_activities/innovation_case_studies.zip

Authors
TECHNOPOLIS GROUP
Todor Kamburow
Alasdair Reid
Paul Simmonds

MANCHESTER INSTITUTE OF INNOVATION RESEARCH
Paul Cunningham
Jakob Edler

Acknowledgments
Thanks are due to the officials of the European Commission that supervised the study and provided helpful assistance and support throughout the drafting of this guidance document: Veronica Gaffey, Daniel Mouqué and Marielle Riché.

The authors would also like to thank the following for their feedback on a draft version of the guide: Tito Bianchi, Ministry of Economic Development, Italy, Gordon McLaren, ESEP Ltd, Scotland, Kieran Moylan, Border Midland and Western Regional Assembly, Ireland, Ondrej Ptacek, Ministry of Industry and Trade, Czech Republic, Doris Schnitzer, Department of European Affairs, Vorarlberg State Government, Austria, Michal Miedzinski, Technopolis Group.

This guidance is based on an assessment of the state of the art of methodological practices for the evaluation of innovation activities supported by public funds. It is intended primarily for managers of ERDF programmes which support innovation and those responsible for their evaluation. It highlights the main questions to ask before developing technical specifications and examines the pitfalls to be avoided, the advantages and limitations of certain methods, and the necessary conditions for a quality evaluation.