Structural Fund evaluation as a programme management tool: comparative assessment and reflections on Germany

1 Introduction

During the 1990s, in Germany as elsewhere in the EU, the Structural Funds were an increasingly powerful catalyst for economic development evaluation. Regulatory requirements grew in response to several factors, notably: the increasing significance of the EU budget allocated to the Structural Funds; the political importance of EU economic and social cohesion policy; a growing need for accountability mechanisms due to decentralisation of responsibility for delivering the Structural Funds; and increasing awareness of the benefits of improved monitoring and evaluation.

Structural Fund evaluation activity has increased and evolved over time, with a changing menu of studies. At EC level, growing emphasis has been put on determining effectiveness and guaranteeing accountability; at the programme level, evaluation is being recognised as a tool generating insights that help to understand and improve policy design and delivery. While it has not been universally embraced, Structural Fund evaluation as a management tool has become an ever more familiar and valued option.

The 2000–2006 regulatory requirements for the monitoring and evaluation of Structural Fund programmes are more extensive and explicit than ever before. In parallel, the management and delivery of these programmes is more demanding. New provisions, notably the decommitment rule and performance reserve, require the timely and effective implementation of programmes. In addition, resources have in many cases been reduced, increasing the need to achieve value for money. Monitoring and evaluation are regulatory obligations, and an opportunity is arguably lost if this activity is not at the same time oriented to meeting the challenges of improved programme delivery. Drawing on evaluation experience across the EU, this paper aims to explore the factors which appear to condition this conception of evaluation, relating these particularly to the German context.

Following this introduction, the paper has four sections. The first reflects on the nature of evaluation for quality management, and provides a short historical review of the development of Structural Fund evaluation obligations. The second provides an international overview, relating Structural Fund evaluation practice to two framework factors: the overall status of evaluation in the Member States, and differing Structural Fund delivery arrangements. The third section specifically addresses evaluation for quality management, exploring the factors which make this type of evaluation more likely to take place, to be successful and to improve Structural Fund programming. The fifth and final section reflects on the issues posed by the further development of evaluation for quality management in Germany.

2 The development of Structural Fund evaluation for quality management

Structural Fund programmes are complex: Their design and implementation take place over a defined timeframe, relying on the participation of diverse actors from multiple levels of governance intervening in a wide range of inter-related economic development fields. Structural Fund evaluation is also complex because studies have multiple aims, whose relative significance varies between the different stakeholders. The three core aims are: ensuring accountability in the use of public funds; quantifying the economic impacts achieved by Structural Fund interventions; and generating insights to improve current and future policy design and delivery. The focus of this paper is on the third of these roles, a function of most importance to active participants in programme delivery.

The importance of improving current and future policy design and delivery is clear. Programme implementation consumes resources, so benefits from efforts to ensure it is efficient. "[B]y discarding the imple-
mentation side, important information for adjusting and rectifying the current programme management may be foregone.\textsuperscript{2} More fundamentally, the way in which programmes are executed can impact on their effectiveness as economic development tools. Addressing any problems through evaluation can therefore increase economic impacts. This may be direct, where evaluation leads to changes in programme delivery, or indirect, where it strengthens partnership by focusing participants on their shared objectives, raises awareness of wider programme activities, and engenders a questioning approach encouraging learning and innovation.

Evaluations typically have multiple objectives, each potentially demanding a distinct methodological response. To illustrate this, it is useful to distinguish between two types of evaluation activity in the theoretical literature: summative and formative. Summative evaluation determines the essential effectiveness of interventions by enumerating what has been achieved relative to targets.\textsuperscript{3} It tends to be used to generate benefit for external actors (not directly involved in programme management), for accountability reasons or to assist in resource allocation. Formative techniques, in contrast, are more concerned with improving programme design and delivery. They would typically benefit programme managers and partners, improving operational practices and in-creasing the effectiveness of the programmes delivered.

For each programming period, there has been a menu of Structural Fund evaluations, closely tied to the policy cycle. There are three main types of evaluation: (i) ex ante, sometimes termed “assessment” rather than “evaluation” and undertaken prior to programme approval to verify its quality and completeness and ensure that it is evaluable; (ii) interim, undertaken at the mid-way point of implementation to identify progress relative to targets and propose any necessary reorientations of policies or delivery mechanisms; and (iii) ex post, undertaken after a programme has been completed to ascertain how effective it has been relative to its original objectives.

Relating these evaluations to the above typology, ex ante evaluation prepares the ground for summative and formative evaluations by verifying in advance the quality of a programme and its delivery mechanisms. Interim evaluations are the most relevant stage for formative evaluation. They take place during implementation, aiming to improve policy design and delivery over the remainder of the programme. A further factor giving them a formative focus is that they are commissioned by those managing the programmes, and who are most directly interested in improving delivery arrangements. Interim evaluations also have a summative dimension, providing a mid-stage account of activity to date and its short-term impacts, relative to targets. Finally, ex post evaluations are the most clearly focused on summative evaluation. They are carried out on completed programmes, enabling impacts to be assessed and delivering accountability by making resource allocation and its outcomes more transparent. Just as interim evaluations have a summative dimension, ex post evaluations would have a possible formative one, enabling insights into operational arrangements which could improve the performance, not of the completed programmes, but of future ones.

The basis for the monitoring and evaluation requirements of Structural Fund programmes are the EU Council regulations, renewed prior to each new programming period. They have developed incrementally over the last three rounds of programming (1989–1993, 1994–1999 and 2000–2006); interpreted and amplified by the European Commission, they have led to progressively more specific and extensive requirements for programming authorities in the Member States. In parallel, there has been a reorientation in the focus of evaluation. The continuing emphasis on summative objectives has been supplemented by a greater potential role for formative evaluation and the more widespread use of evaluation as a management tool: (i) directly through a changing menu of evaluations, and in particular the growing role of interim evaluation; and (ii) indirectly through the distribution of responsibility for evaluation, including the delegation of responsibilities to those delivering programmes.

Under the 1989–1993 Structural Fund regulations\textsuperscript{4} two studies were explicitly required: “ex ante and ex post assessment”.\textsuperscript{5} These were to be undertaken “to highlight the impact of programmes with respect to
The main objectives of the Structural Funds and to analyse their effects on specific structural problems".\(^6\) As such, the initial, formal rationales for Structural Fund evaluation were firmly summative. The authority responsible for the evaluations was not specified but, in practice, relevant ex ante and ex post were funded, commissioned and steered by desk officers in the European Commission. There was a formative dimension to some studies carried out on programmes from this period, but it was limited. Taking the example of the ex post evaluation of the 1989–1993 Objective 2 programmes, its predominant emphasis, of course, was on impact evaluation, especially of employment effects. In parallel, a secondary dimension of the study was to identify whether the 1988 Structural Fund reforms improved how programmes were implemented and results achieved.\(^7\) This led to operational recommendations for current and future programmes. In addition, although not formally required, selected programmes undertook interim evaluations during this period on an ad hoc basis (eg. in the UK, the Netherlands and Germany).

The 1994–1999 regulations increased evaluation obligations.\(^8\) Article 6(2) of the main regulation stated that: “In order to gauge their effectiveness, Community structural operations shall be the subject of prior appraisal, monitoring and ex post evaluation designed to appraise their impact with respect to the objectives ... and to analyse their effects on specific structural problems”. Key additions were the explicit reference to “evaluation” rather than “appraisal”, and the formal incorporation of monitoring, a process of data collection underpinning and facilitating evaluation.

The first evaluation task of the 1994–1999 period was the ex ante evaluations of the new programmes. These studies were commissioned by the EC and primarily used to guide the Commission’s desk officers in programme negotiations. They mainly addressed the strategic coherence and completeness of programmes rather than the quality of their delivery arrangements. While they informed a negotiation process which improved programmes and the rigour of their

---


\(^3\) Council Regulation (EEC) No. 2052/88 of 24 June 1988, Article 6


---

### Table 1

<table>
<thead>
<tr>
<th>Type</th>
<th>Focus</th>
<th>Summative Rationales</th>
<th>Formative Rationales</th>
<th>Contribution to policy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ex ante appraisal - prior to programme approval</td>
<td>Prior appraisal rather than evaluation per se</td>
<td>Summary of anticipated outputs, and gross and net impacts of programmes in advance. Verification of the quality of design of the programme – does it respond to the issues identified?</td>
<td>Verification of the quality of design of implementation and monitoring arrangements, proposing improvements where necessary, before the programme has been approved and launched.</td>
<td>Single loop learning – improving the design and delivery of the current programme.</td>
</tr>
<tr>
<td>Interim evaluation - mid-way through the programming period</td>
<td>Mainly formative, but with a summative dimension</td>
<td>Identification of initial results of assistance (outputs and, where possible, short-term impacts), and compare this to targets.</td>
<td>Verification of the continuing relevance of the objectives of the policy and propose adjustments if appropriate. Assessment of the operation of monitoring and implementation arrangements and propose adjustments if appropriate. Provision of information helping to prepare for any subsequent programme.</td>
<td>Single loop learning – improving the design and delivery of the current programme, in itinere.</td>
</tr>
<tr>
<td>Ex post evaluation - after the programme has been implemented</td>
<td>Mainly summative, but with a formative dimension</td>
<td>Economic impact evaluation to establish what the programme achieved (outputs and their consequent gross and net impacts).</td>
<td>Indication of factors contributing to the success or failure of implementation and the achievement of results. Use this to inform the operational arrangements of future programmes. Derivation of policy conclusions for economic and social cohesion.</td>
<td>Double (and single) loop learning – supporting the redesign of the Structural Fund policy itself, and the design and delivery of the next programmes.</td>
</tr>
</tbody>
</table>
evaluation provisions (eg. targets), the impression held by some programme-level actors was that the studies were undertaken with a lack of transparency and openness, used against programme actors rather than informing them as part of a shared strategic process.

There was no regulatory requirement for interim evaluation during the 1994–1999 programming period. These evaluations were instead introduced as an obligation during programme negotiations with the EC, through inclusion among the standard clauses appended to Single Programming Documents (SPDs) and Operational Programmes (OPs). Interim evaluations are “in some respects, the most important” of all Structural Fund evaluations. They were the first large-scale round of obligatory evaluations entrusted to the Member State level or below and were conceived as a means of deepening monitoring, to provide mid-term information improving programme efficiency and effectiveness. Undertaken primarily to enable programmes to report to the EC on progress and necessary adjustments, the opportunity was also seized in many cases to address practical questions of direct interest to programme partnerships.

Supporting this innovative activity, multiple publications and initiatives promoted the perceived benefits of formative approaches, especially as regards interim evaluations. In particular, MEANS Handbook No 110 promoted process-oriented evaluation, seeing “intermediate evaluation as a real asset to programme implementation and not just as a response to a regulatory requirement”. In parallel, under SEM 200011, the requirement for systematic evaluation to be introduced into all EU programmes was reinforced by a guide which stated that, among the principles of evaluation (alongside being analytical, systematic, reliable and issue-oriented) was that it should be user driven – providing useful information to decision-makers.12 Nonetheless, the relatively limited formal emphasis on evaluation as an instrument for quality management led one commentator13 to state that:

In the evaluation strategy for EU structural funding, the analysis of the implementation structure carrying regional policy has almost been a “blind spot” and “missing link”. The implementation stage and dimension of EU structural funding has been given little space and attention in the Commission’s guidelines for evaluation and in the relevant tenders and “terms of reference”.

Marking the increasing emphasis the EC is placing on evaluation for quality management, it is significant that the EC’s synthesis report on the interim evaluations of the 1994–1999 Objective 1 and 6 programmes was entitled: Better management through evaluation.14 The study emphasised the potential for interim evaluations to be operationally and strategically oriented, stating that:

In many Member States, programme managers and policy makers … asked, on the basis of the evaluation findings, what kind of adjustments needed to be made, particularly with regard to the current interventions.

In a further sign of its commitment, the EC selected “Evaluation for quality” as the theme of its fourth international Structural Fund evaluation conference held in Edinburgh in 2000, which once again sought to disseminate lessons from international evaluation experience.

Moving on to the regulations shaping the current programming period, the preamble to the 2000–2006 regulation draws on previous experience to emphasise that the effectiveness and impact of the Structural Funds will depend on improved and more thorough evaluation. All three stages of evaluation are now formally required: ex ante, interim and ex post. In addition, further thematic evaluations can be undertaken at the discretion of the EC and Member States, where the need arises. It is already evident that this programming period will be the most demanding yet. Partly this is because of the strengthened monitoring obligations associated with decentralised management, and partly because the outcome of evaluation is now linked to funding allocation through the performance reserve.

3 Contrasting Frameworks for Structural Fund Evaluation

The Structural Fund regulations provide a common evaluation framework, but this has been variously interpreted and applied in different contexts. Although encouraged, evaluation for quality management is not an obligation and has not been an element of all evaluations, even interim ones, in

---


(11) Sound and Efficient Management, reform programme of the European Commission of its financial and resource management


spite of its potential benefits. Two critical factors shape EU Member State approaches to Structural Fund evaluation: The first, familiar factor is the overall status of economic development evaluation; and the second, perhaps less widely discussed dimension, is the distinctive ways that Structural Fund programmes are conceived and managed. These two factors are explored below.

3.1 The status of regional policy evaluation

Across the EU, the status of evaluation differs. To illustrate the strength of commitment to economic development evaluation, Table 2 takes the example of the evaluation of national regional aid schemes. Although evaluation activity has increased across all Member States because of Structural Fund obligations, they are nonetheless still very differentiated. Three groups can be identified: (i) countries with a long tradition of public policy evaluation,

<table>
<thead>
<tr>
<th>Member State</th>
<th>Evaluation tradition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Austria</td>
<td>Little tradition of evaluation, but several regional policy evaluations in recent years. Emphasis is broad, covering institutional and procedural issues, as well as quantitative evaluations of effectiveness.</td>
</tr>
<tr>
<td>Belgium</td>
<td>No regular evaluations of regional policy in Flanders or Wallonia. Structural Fund experience is leading to some impact evaluation studies of individual measures.</td>
</tr>
<tr>
<td>Denmark</td>
<td>No national regional policy measures since 1992, so no evaluation required.</td>
</tr>
<tr>
<td>Finland</td>
<td>Periodic evaluations of regional and industrial policy measures during the 1980s and 1990s, especially interview surveys of companies receiving government subsidies.</td>
</tr>
<tr>
<td>France</td>
<td>Public policy tradition of evaluating “public actions” but no known evaluation of regional policy measures during the 1980s and 1990s.</td>
</tr>
<tr>
<td>Germany</td>
<td>Evaluation integral to the regional economic policy framework. System focuses on tracking the outputs of regional aid and analysing the performance of eligible areas. Only intermittent analysis of the effectiveness of regional policy measures.</td>
</tr>
<tr>
<td>Greece</td>
<td>No regular evaluations of regional policy. Some official and academic evaluations of regional policy measures in the 1980s. Superseded by emphasis on Structural Fund evaluation.</td>
</tr>
<tr>
<td>Ireland</td>
<td>Little tradition of evaluation, although some studies since the 1970s involving a mix of quantitative modelling and survey methods. Increasing emphasis on Structural Fund evaluation in the 1990s through modelling by the Economic &amp; Social Research Institute.</td>
</tr>
<tr>
<td>Italy</td>
<td>No tradition of regional policy evaluation apart from Structural Funds. The creation of evaluation units in all national and regional administrations is changing the picture.</td>
</tr>
<tr>
<td>Luxembourg</td>
<td>Infrequent regional policy evaluation linked to policy redesign phases.</td>
</tr>
<tr>
<td>Netherlands</td>
<td>Regular regional policy evaluation, undertaken every 4-5 years in line with reviews of regional policy legislation. Methods involve evaluation of efficiency and effectiveness through surveys and modelling.</td>
</tr>
<tr>
<td>Portugal</td>
<td>No tradition of regional policy evaluation apart from Structural Funds.</td>
</tr>
<tr>
<td>Spain</td>
<td>No recent tradition of regional policy evaluation apart from Structural Funds. Some regional policy evaluation in the 1970s, but practice ceased.</td>
</tr>
<tr>
<td>Sweden</td>
<td>Tradition of regular regional policy evaluation, mostly through surveys of firms receiving government subsidies.</td>
</tr>
</tbody>
</table>
where evaluation is seen as integral to policy-making and where the outcomes of regional policies are regularly examined; (ii) those where evaluation is an occasional exercise, for example associated with policy redesign; and (iii) those viewing evaluation as of limited value and largely irrelevant to mainstream policy practice. The lack of a tradition of regional policy evaluation among this final group in some cases reflects the limitations of evaluation resources and skills in systematising policy design and revision, but can represent a broader scepticism towards evaluation as a policy assessment tool. Evaluations here tend to be one-off exercises, rather than integral stages in policy making and delivery.

Where evaluation is integral to economic development activity, Structural Fund programme managers benefit from the fact that evaluation is already a credible and widely understood practice (with many aware of the rationale for, and methods of, evaluation, including from direct experience) and from the availability of experienced evaluators. However, the domestic context alone is not sufficient to encourage Structural Fund evaluation supporting quality management. In particular, while summative evaluation may be effective, it does not follow that policy delivery will be conceived as an ongoing learning process, and that evaluation for quality management will be an element of the evaluation culture. Germany provides an illustration of a Member State with structured provisions for regional policy evaluation: The regional policy Framework Plan (the Rahmenplan) has included evaluation provisions since its inception in 1969. The primary focus of activity, however, has been on tracking the outputs and impacts of regional aid, analysing the performance of eligible areas and ensuring accountability. The structured remit for evaluation has not included scope to appraise policy management and delivery mechanisms. The formalised context may have led to a narrow conception of evaluation impeding the emergence of formative practices. Even impact evaluation in Germany, whilst built into the policy calendar, has been undermined by acute awareness of the inherent methodological shortcomings.

3.2 Structural Fund delivery arrangements and evaluation

The second factor influencing the status of Structural Fund evaluation for quality management among EU Member States is the way that these programmes are conceived and run. For the practice to be seen as useful, a programme needs to be understood by its participants as a clear entity, with distinctive policy objectives and delivery structures – as “something to evaluate”. There also needs to be scope for actors to change policies and/or modes of delivery in the light of evaluation findings. These conditions are not met in all Structural Fund contexts.

To explore this, it can be helpful to describe a Structural Fund management continuum. At one extreme are those programmes delivered through dedicated systems established solely and specifically for deciding on Structural Fund co-finance (which can be termed “differentiated”). Examples include programmes in Belgium, Denmark, the Netherlands, Sweden and the United Kingdom. Here, project appraisal is carried out by Structural Fund secretariats or other competent organisations and/or committees, while final project decision-making is typically the responsibility of a committee representing the Structural Fund programme partnership.

At the other extreme are programmes using existing, domestic policy decision-making channels for Structural Fund resource allocation (characterised as “subsumed”), eg. Germany, Austria and Spain. Here, projects are appraised and selected using established domestic resource allocation routes, with Structural Fund budgets integrated alongside domestic ones, and decision-making often delegated to single competent agencies rather than reserved for dedicated Structural Fund committees. Falling between these two extremes, several Member States (eg. France, Finland and Italy) have “composite” systems relying on domestic appraisal and decision-making systems but with some additional, Structural Fund-specific elements.

Differentiated Structural Fund systems may lend themselves most readily to formative evaluation. Since there are Structural Fund specific policies, structures and channels, the programme to be evaluated has a clearer profile, and distinct, dedicated structures and systems whose performance can legitimately be assessed and modified.
on the basis of evaluation findings. Evaluation for quality management fits less well with subsumed systems, although it may still take place (as Austrian, German and Spanish examples illustrate). The interventions that constitute the programme belong to a complex, wider picture of domestic policy-making. There is no strong “programme identity”. It is difficult to disentangle Structural Fund elements from the broader economic development picture, which it is not the remit of the Structural Fund study to evaluate or to call into question.

3.3 Differing conceptions of Structural Fund evaluation

Given the differing contexts described above, it is not surprising that the emphasis placed on evaluation as a management tool varies between and within Member States. Table 3 seeks to summarise responses by distinguishing between evaluation focused on meeting regulatory requirements, (i.e. carried out with a summative focus, justifying the resources used and impacts achieved, and potentially with limited follow-up locally) and evaluation which, while undertaken at the instruction of the EC, has also been appropriated to generate information enabling programme actors to improve programme design and/or delivery.

As the table illustrates, a more minimal interpretation of evaluation appears to have been the norm in Portugal, Spain and Greece for example. Among the Member States where evaluation has been exploited most enthusiastically are the UK, Denmark, and Ireland. Elsewhere, the picture is more mixed, with strong contrasts in regional approaches in Germany, France and Italy, for example.

Table 3
Overall approach to Structural Fund evaluation in EU Member States

<table>
<thead>
<tr>
<th>Member State</th>
<th>Approach to evaluation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Austria</td>
<td>Most evaluation restricted to regulatory requirements. However, gradually being valued more as a management tool and integrated into the programming process.</td>
</tr>
<tr>
<td>Belgium</td>
<td>Evaluation goes beyond regulatory requirements. Investment in networked monitoring system and modelling for impact assessment.</td>
</tr>
<tr>
<td>Denmark</td>
<td>Evaluation goes beyond the regulatory requirements. Studies have been conducted independently to improve the effectiveness of the programmes.</td>
</tr>
<tr>
<td>Finland</td>
<td>Evaluation generally restricted to regulatory requirements, but active efforts to improve quality, relevance and robustness of methods.</td>
</tr>
<tr>
<td>France</td>
<td>Evaluation organised to meet regulatory requirements, but examples of regions using studies also to address programming issues.</td>
</tr>
<tr>
<td>Germany</td>
<td>Evaluation organised to meet regulatory requirements, but some examples of regions using studies also to address programming issues.</td>
</tr>
<tr>
<td>Greece</td>
<td>Evaluation generally restricted to regulatory requirements.</td>
</tr>
<tr>
<td>Ireland</td>
<td>Evaluation goes beyond regulatory requirements. Institutionised management of evaluation, and extensive investment in modelling for impact assessment.</td>
</tr>
<tr>
<td>Italy</td>
<td>Evaluation generally restricted to regulatory requirements.</td>
</tr>
<tr>
<td>Luxembourg</td>
<td>Meet regulatory requirements.</td>
</tr>
<tr>
<td>Netherlands</td>
<td>Evaluation generally restricted to regulatory requirements. Effort to establish national frameworks.</td>
</tr>
<tr>
<td>Portugal</td>
<td>Evaluation generally restricted to regulatory requirements.</td>
</tr>
<tr>
<td>Spain</td>
<td>Evaluation generally restricted to regulatory requirements.</td>
</tr>
<tr>
<td>Sweden</td>
<td>Evaluation generally restricted to regulatory requirements, but pro-active responses.</td>
</tr>
<tr>
<td>United Kingdom</td>
<td>Evaluation organised to meet regulatory requirements, but regions frequently use studies to address programming issues.</td>
</tr>
</tbody>
</table>
4 Facilitating evaluation for quality management

Within this overall context, the next question is to consider what practices and approaches facilitate the use of evaluation to improve programme management and delivery. There are several immutable features which arguably impede the quality and utility of evaluation. Among them are the timing of interim and ex post evaluations, which is imposed by the regulations, and the breadth of Structural Fund programmes and thus of their evaluations. Studies have to address the wide range of interventions which programmes include, implemented by diverse actors, and to take horizontal dimensions into account. This context, combined with time and resource limitations, carries the risk that the quality of some or all aspects of the evaluation will be compromised by its complexity. This section focuses on three aspects of evaluation practice where there appears to be room for manoeuvre: first, how the conditions for effective evaluation of any sort can be met, namely the availability of appropriate skills and data sources; second, how the style and organisation of evaluations can improve the ability to contribute to quality management objectives; and, finally, how to ensure that evaluation results genuinely feed into improving the quality of programme management and delivery.

4.1 Pre-conditions for effective evaluation

4.1.1 Evaluation capacity

For effective evaluation, the first fundamental condition is the availability of appropriate knowledge and skills among evaluators and those commissioning and participating in studies. As a result of the Structural Fund obligations, more people have direct experience of evaluation. The evaluation capacity of programme participants (and of many evaluators) is more likely to derive from direct experience than formal training, since this is underdeveloped across Europe. Two key questions in this context are: (i) how can evaluation capacity among non-specialists be enhanced; and, (ii) how is it possible to ensure the availability of qualified evaluators?

At EC, Member State and/or programme level, targeted awareness raising, including through guidance and exchange of experience can accelerate the development of evaluation capacities. The EC has been active in developing evaluation competencies among Structural Fund programme participants. Among its contributions, the Directorate General for Regional Policy (DG Regio) has published a variety of documents, analysis and guidance on evaluation, the recent dissemination of which has been facilitated through the Inforegio website. Through the “MEANS” initiative, launched in 1994, DG Regio also sponsored a sustained programme of research providing methodological support facilitating compliance with increasing Structural Fund evaluation requirements and encouraging a common language for evaluation (although the initiative was initially weakened by insufficient dissemination).

National level initiatives to develop evaluation capacities have included guidance, conferences and the formalisation of channels to diffuse information. Austria lacked a strong evaluation culture before EU accession. To accelerate learning about evaluation and the Structural Funds, the Austrian Federal Chancellery instituted a common evaluation resource, Checkpoint EVA, to facilitate access to relevant documentation. This consists of a library with more than 600 evaluation-related documents and a web site including a database of evaluation projects and evaluators, a news page and a discussion forum. The service brings increased visibility to evaluation in Austria, and ensures that there is an easily accessible source of relevant information for experts and novices alike.

France has used various methods to develop evaluation capacity. Among the most innovative are the interactive “days of exchange” following each round of major Structural Fund evaluations. The days unite relevant actors to discuss the results of evaluations and draw out lessons - for programme design, programme management and the conduct of evaluation. These events are co-ordinated by a national agency, DATAR, and involve various stakeholders, including those commissioning the evaluations, national experts and evaluators. They present several advantages: (i) the active involvement of all the main players in evaluation means that any resolutions have the
commitment of all key actors; (ii) those procuring evaluations meet a range of evaluators, and can discuss their work with organisations which have already used their services; (iii) all managing authorities and evaluators benefit from experiences matured in other contexts; and (iv) evaluations comes to be perceived as part of an on-going process and not as one-off exercises.

A further national impulse comes from evaluation societies which, where they exist, benefit evaluators and those commissioning evaluations. There are evaluation societies in six EU Member States: Germany, France, Finland, the UK, Italy and Belgium. Not all have a specific focus on Structural Fund or even economic evaluation, but they provide a range of opportunities for development, including through conferences and working groups, alert services, and targeted networking.

While there are many examples of positive capacity building initiatives, they have mainly been sporadic and could be intensified and better structured to further develop the competencies of evaluators, managers, partners and other stakeholders.

The second aspect of evaluation capacity is whether suitably qualified and experienced evaluators are available. The evaluation community is still differentiated across EU Member States (Table 4). Three main determinants help to explain the patterns. The first is the length of time a Member State has been in the EU. For example, the evaluation community in Sweden, which joined the EU in 1995, is still limited to a small number of evaluators, although newcomers are now penetrating the market. The second is the overall market size for evaluation services. In Denmark, the small size and number of Structural Fund programmes limits the evaluation market and community. In contrast, Germany, France and the UK are among those larger countries which have sizeable and active evaluation communities and a wide range of Structural Fund evaluators. Lastly, the overall status of evaluation

<table>
<thead>
<tr>
<th>Member State</th>
<th>Availability of evaluators</th>
</tr>
</thead>
<tbody>
<tr>
<td>Austria</td>
<td>Limited - some use of foreign consultants required. New Member State without its own strong tradition of evaluation. Gearing up rapidly.</td>
</tr>
<tr>
<td>Belgium</td>
<td>Adequate - some use of French evaluators in Wallonia.</td>
</tr>
<tr>
<td>Denmark</td>
<td>Limited due to limited national regional policy evaluation and small number of Structural Fund programmes.</td>
</tr>
<tr>
<td>Finland</td>
<td>Limited and apparently declining. New Member State. Some entrants to the field were disillusioned. However, positive initiatives to promote high quality evaluation (methodological research, national level working groups, etc.).</td>
</tr>
<tr>
<td>France</td>
<td>Good – active community. Large number of regions and programmes. Established position of evaluation.</td>
</tr>
<tr>
<td>Germany</td>
<td>Good – active community. Large number of regions and programmes. Relatively established position of evaluation.</td>
</tr>
<tr>
<td>Greece</td>
<td>Limited. Low status overall for evaluation.</td>
</tr>
<tr>
<td>Ireland</td>
<td>Limited but high quality. Small Member State, but high status for evaluation.</td>
</tr>
<tr>
<td>Italy</td>
<td>Good – active community. Large number of regions and programmes. Established position of evaluation.</td>
</tr>
<tr>
<td>Luxembourg</td>
<td>Limited, but additional availability in neighbouring Member States.</td>
</tr>
<tr>
<td>Netherlands</td>
<td>Good – active community. Small Member State, but active evaluation culture.</td>
</tr>
<tr>
<td>Portugal</td>
<td>Limited community.</td>
</tr>
<tr>
<td>Spain</td>
<td>Adequate with variable quality</td>
</tr>
<tr>
<td>Sweden</td>
<td>Limited – use of foreign consultants required</td>
</tr>
<tr>
<td>United Kingdom</td>
<td>Good – active community (but limited numbers in Scotland and Northern Ireland)</td>
</tr>
</tbody>
</table>
matters. Spain has been a member of the EU since 1986 and has a large demand for Structural Fund evaluation, but the evaluation community is still limited because of the relatively low status of evaluation.

A large evaluation community is important for Structural Fund evaluations since studies are often undertaken simultaneously across many regions. As a general trend, it has been difficult to balance the need to ensure that markets are open to new evaluators with the need to hire sufficiently experienced practitioners. The methods used to select evaluators can limit or extend the community. The use of restricted calls for tender may present strong access barriers for newcomers, and lead to market dominance by a small number of experienced operators. In contrast, initial open calls, followed by invitations to a shortlist of tenderers, can open up access, but may also increase the administrative burden, risk and time inputs of those commissioning studies. A further way to extend the community is through brokerage facilities, with an appropriate body maintaining an openly available list of evaluators, updated using submissions from evaluators indicating their experience and skills.

In addition to the size of the evaluation community, their proximity to relevant programmes is also important if evaluation is to be a management tool. Evaluators need familiarity with the economy and institutions of the relevant area, and sufficient accessibility to partners to undertake partnership-based, interactive evaluation. Where local evaluators lack the necessary skills (which are often concentrated in capital city regions), one solution, used in Italy and France, is to encourage a transfer of expertise through a partnership between less experienced local operators and more experienced evaluators based elsewhere.

The third important feature of an evaluation community is their skills. To conceive Structural Fund evaluation as part of a wider programme management process requires not only proficiency in economic evaluation techniques, but also awareness of organisational and management science, and – critically – highly developed interpersonal and political skills.

Especially in the mid-1990s, when interim evaluations were carried out extensively for the first time at programme level, there was some lack of specialist expertise, with inexperienced evaluators “learning on the job”, and often seeking simply to stay “one step ahead” of those commissioning the evaluations, who were also often inexperienced. “Learning by doing” has remained the primary route to gaining evaluation skills: There is still a lack of formal training, and the level of professionalisation is often therefore relatively low, although it has developed significantly over time. Experience in evaluation has increased, but methodological weaknesses remain, among them: a lack of application of sufficiently robust economic impact evaluation techniques; and, difficulty in addressing the wide range of thematic issues encompassed in any single Structural Fund programme, including not only the range of intervention types but also the horizontal themes (notably environmental sustainability and gender mainstreaming). One solution to this latter issue is to select multiple evaluators for a given study, each with proven skills in specific areas. This practice is most applicable in large programmes given its budgetary implications, and has, for example, been used in Nordrhein-Westfalen and Yorkshire and Humberside. Smaller programmes still face a particular problem in contracting evaluators with sufficiently broad expertise to evaluate all dimensions of the programme to the same standard.

4.1.2 Data and information – a key resource

The second condition facilitating evaluation as a tool for quality management – and also enhancing management potential on a day-to-day basis – is the availability of reliable, up-to-date and meaningful programme data, reporting on both the financial and physical state of advancement of the programme and its constituent projects.

In the 1989–1993 programming period, monitoring was poorly developed: Data was patchy or non-existent, there was an absence of targets and indicators and monitoring systems (for physical indicators at least) were rudimentary and unsystematic. The situation improved significantly in 1994–1999, often in fact as a result of evaluators, especially in the course of the first 1994–1999 interims, having to divert resources away from added value
Informationen zur Raumentwicklung
Heft 6/7.2001 351

evaluation tasks and towards establishing basic data sets and, in the least developed cases, selecting the necessary indicators for which data should be collected and assisting in the design and establishment of initial database solutions.

An important issue here is that the monitoring data for exploitation by evaluators should be defined not only to facilitate the analyses required by the higher tiers involved in overseeing the Structural Funds, but should also respond to the information needs of programme managers and their partnerships. In practice, this implies much greater levels of detail.

In the current programming period, a step-change is taking place in the quality of monitoring provisions which augers well for future evaluation. Prompted by the stricter regulations and EC pressure, as well as the efforts of some Member State and regional authorities (and evaluators), monitoring capabilities are developing across four dimensions: database solutions; indicators; skills; and the processes which bring systems to life.

Investment in monitoring databases has involved further exploiting IT solutions for data gathering, storage, processing and analysis. The design of monitoring systems has tended to be country- or region-specific to fulfil a range of particular conditions, such as compatibility for electronic data exchange with the Commission. There has been an evolution in available solutions over time moving from paper-based monitoring systems, through spreadsheets to more sophisticated programmed databases and finally to web-based systems.

While advancing database technologies open potential for more efficient and effective programme monitoring, they pose significant technical challenges. Few regions or Member States have found the design and “rolling out” of new systems to be easy.

The quality of monitoring data has been improved by identifying more meaningful indicators, and interpreting them more uniformly, a pre-requisite for aggregation and comparison. Further enhancements to data have come from the progressive establishment of a common language and hierarchy of indicators for use at all levels of programming. The EC has sought to clarify indicator terminology and to present a clear relationship between programme objectives, delivered and analysed from the top down, and programme outcomes - measured by a hierarchy of financial indicators and physical indicators (outputs, results, impacts) - delivered from the bottom up and tracked through monitoring. Most regions are trying to apply the EC guidance in practice, and to ensure that there are at least a few core indicators common to all Objective 2 programmes, although the consistency of definitions still varies greatly. Remaining problems include the credibility of baseline indicator values, complicated by the fragmented geography of Objective 2 areas and the lack of up-to-date data, and the quantification of targets more generally, since programmes may still in practice apply inconsistent methods.

Underpinning enhanced data capabilities, national and regional authorities are investing heavily in human capacity, using seminars, guidance notes, brochures and manuals to raise awareness, inform and motivate partners and applicants and promote common monitoring procedures. Making monitoring systems work requires all the actors generating or processing data to share common understandings of methods and the definitions being applied. Otherwise, the reliability of data is undermined.

Just as with programme management and evaluation, monitoring has involved an evolutionary process of learning and adaptation of systems and procedures. In some cases, the monitoring of Structural Fund programmes is already sophisticated and way ahead of the monitoring of domestic policy interventions. In other regions, progress has been slower. Improvements to monitoring systems in this round will bring potentially significant benefits for the day-to-day management of programmes, and for the quality of their evaluation, feeding into a deeper rooted process of programme learning.

4.2 Evaluation for quality management – process issues

The third set of conditions for effective evaluation relates to how evaluation itself is managed. Key factors are: (i) conceptualising evaluation as part of a wider process of strategic management; (ii) defining relevant objectives and then an appropriate method; (iii) conceptualising evaluation as a “process”, rather than a product; (iv) prioritising participation to
facilitate collective and self-learning; and (v) ensuring acceptance for evaluation among the programme’s stakeholders.

4.2.1 Conceptualising evaluation as part of wider strategic management processes

Attitudes towards evaluation and the expectations from it affect how evaluation is undertaken and used. A fundamental principle is that evaluation be conceptualised as part of a wider integrated and cyclical process of Structural Fund management – the “virtuous circle of programme delivery”. This requires programme managers and partners to see the programming cycle as a continuum. It begins with the development of the programme, itself built on input from previous programme experience. Programme marketing involves the active generation of projects (interacting with applicants to bring forward projects meeting strategic objectives and priorities) and providing pre-application advice to improve the relevance and technical quality of project applications even before submission. Then, systematic approaches to project selection require formal and systematic appraisal and selection procedures. Thereafter, it is important to see project award as the start of a process of close and continuous monitoring and project aftercare, supported by effective computerised systems able to track projects from initial proposal submission to completion. The experience generated in the stages of the programming cycle, processed through evaluation, in turn feeds back into refining the strategy and programming practice.

What does the programming cycle imply for evaluation itself? Two overall approaches can be contrasted. In the first, minimal, approach Structural Fund evaluation is organised at the last minute, primarily to respond to an impending regulatory deadline. It is perceived as diverting resources from programme delivery, and therefore undertaken using methods which meet obligations, but minimise the demands on programme actors. Once complete, the study is dispatched to those required to see it, and programme actors return to programme delivery. At the other extreme, evaluation is viewed quite differently, as a useful and integral part of programming, underpinning strategic decision-making. Here, it is a high priority at the programme design stage to ensure that the necessary elements are in place for the programme’s activities to be “evaluable” at a later date, setting out targets and ensuring mechanisms are in place to monitor progress towards them. Subsequently, when evaluation takes place, the accent is on generating findings which meet regulatory requirements but also supply the strategic information which programme actors need. The latter environment embodies the idea of evaluation for quality management.

4.2.2 Defining own evaluation objectives

Actors at different levels of governance have different objectives from Structural Fund evaluation. These can be pursued in parallel, enabling the same study to meet distinct information needs. First, however, these needs have to be identified and tabled. In particular, where the programme level is responsible for undertaking an evaluation, it is important that it does not simply generate and deliver information passively to satisfy the requirements of the organisation imposing the evaluation obligation – the EC – but also uses the exercise for its own benefit, examining whether it has further or different objectives to fulfil. It is only then that objectives relevant to programme level actors will be identified and appropriate methods selected to deliver them.
In addition to passivity, the dominance of existing visions of evaluation can mitigate against evaluation identifying and fulfilling local objectives. In their analyses, MEANS referred to a “reluctance to adopt new methods” and to a “form of evaluation ritualization that may limit its range”\(^{20}\), where dominant techniques and actors inhibit innovation. Examples of dominant approaches include the UK “value for money” standard, the Italian preference for cost-benefit analysis and the influence of close ties between German Federal ministries and economic research institutes.

One of the resulting primary weaknesses of Structural Fund evaluation is the risk of them being “stuck in the middle”, not achieving excellence in terms of “traditional” impact evaluation, nor in facilitating learning among stakeholders through qualitative approaches:\(^{21}\)

**Impact analyses - in the sense of appraising net effects - are barely carried out, and if so they are driven by very pragmatic ... approaches. Little scope remains for innovative, qualitative evaluation approaches which are often more suitable for complex regional development programmes in particular with regard to the stimulation of a learning process within the programme partnership.**

A more overt debate about the multiple objectives of Structural Fund evaluations could be beneficial, identifying aspirations more clearly, and linking them more directly to methodological choices. This process implies one of explicitly making compromises, prioritising certain expectations over others and acknowledging choices made for pragmatic reasons.

### 4.2.3 Focusing on the process, not the product

Structural Fund evaluations address programmes designed and delivered on a partnership basis, relying on the contribution and expertise of a wide vertical and horizontal partnership. To encourage change to systems, behaviours and relationships, the “traditional” evaluation approach, where a detached analysis is undertaken and conclusions are presented for consideration, can be relatively ineffective. Instead, formative evaluation requires dialogue with a wide range of actors – not only to understand programming processes but also to directly engage those actors who are best able to initiate and implement change. To ensure this, the process of participating in an evaluation, and the associated interaction, reflection and learning, should be seen as being as important as the final evaluation report. The evaluation process should facilitate discussion, self-discovery and self-diagnosis among programme partners. Under this model, much of the benefit of the study will already have been derived before the final report is issued.

For evaluation to generate information valuable to programme management, its organisational parameters need to be compatible with its ambitions. The first element is to ensure that sufficient time is built in for a participative process. Key stages include the evaluator becoming familiar with programme delivery mechanisms and partnership organisation, engaging in intensive dialogue, verifying impressions and finally disseminating conclusions. If all constituencies are not consulted, or if there is no time for those who are criticised to respond, then relationships may be damaged and conclusions perceived as ill-founded, counteracting the benefits.

The second condition is to ensure that appropriate evaluators are selected. Independence is a pre-requisite for all Structural Fund evaluators, especially since multi-sectoral, multi-level partnerships need to perceive the evaluation as a collectively owned process of analysis. In addition to the perceived independence of evaluators, their style is influential to establishing trust and openness. In an audit role, the evaluator may need to work as a detached expert, processing quantitative data to generate irrefutable conclusions; for programming purposes, a more effective role for an evaluator may be as a facilitator, involved in an active process of communication in which programme actors will in many cases be helped to generate their own insights and agree their own solutions. In this context, communication skills, including the ability to listen, and a highly developed awareness of the complexities of partnership working are as important as specific technical capabilities. Evaluators need to engage actors, inspire their trust and encourage reflection, and at the same time be seen as responsive yet impartial. In the Aquitaine 1994–1996 Objective 2 interim evaluation, the measured and sensitive approach of the evaluators enabled actors to engage very quickly in a new process of dialogue and...
reflection, setting a positive precedent for subsequent evaluations, and encouraging additional thematic analyses to be undertaken.

4.2.4 Prioritising participation and facilitating learning

The multi-level, multi-sectoral objectives of Structural Fund programmes suggest that it is important for evaluation studies to be owned by and undertaken for the benefit of the partnership, not just the EC, Member State or programme manager. An appropriate signal can be sent about ownership by establishing evaluation steering groups to accompany evaluations. These tend to work under the Monitoring Committee, and ensure that partnership perspectives are taken into account in the design and execution of the evaluation. The nuance in ownership is critical: Evaluation which is expected to contribute to improving programming must not be experienced as an externally driven check on performance, with the threat of punitive action. Instead, it should be seen as a more positive process, undertaken with universal participation for universal benefit.

Evaluation steering committees are common in France, the UK, Belgium, Denmark and Finland. At their most developed, they are permanent structures, addressing monitoring and evaluation issues on an on-going basis, but meeting more frequently when an evaluation is underway. Such structures are not universal. Germany and Spain are among the Member States where this approach is not routinely used, evaluations instead usually being managed directly by programme managers and/or secretariats. This can be successful where summative evaluation is the focus of a study, but limits the ability to interpret the patterns identified through economic impact analysis and poses challenges for the feasibility and likely success of process evaluation.

Several techniques can be used to organise the contribution of programme actors to the evaluation itself, and to stimulate reflection. Surveys of programme participants are frequently used and would typically be undertaken by post, face-to-face or over the telephone, addressing process, activity and impact-related questions. An oral survey arguably inspires spontaneous reflection more flexibly than a written one (although also introducing the challenges of greater interpersonal dynamics), while a face-to-face interview also enables the evaluator to see the context in which the subject is working, gathering a range of insights not available through more remote means, and building a relationship between the evaluator and programme actors.

Interviews imply partners providing information individually to the evaluators who remain the privileged interpreters and analysts. A stimulating and cost-effective complementary approach can be to exploit participative evaluation techniques including the organisation of workshops and focus groups to tackle specific issues. This can raise the profile of priority themes and encourage the sharing of information, perspectives and ideas, leading to preconceptions being challenged and new insights emerging. The value of the opinions and ideas expressed by participants is judged by partners rather than by the more remote evaluator. In the Western Scotland Objective 2 programme, the 1994–1996 interim evaluation included a survey of programme participants which addressed programme management and delivery issues alongside questions about the effectiveness of interventions. The subsequent interim evaluation of the 1997–1999 programme, undertaken by the same evaluators, built on this experience, supplementing the survey approach with working groups to identify ways forward on priority issues.

4.2.5 Gaining acceptance for evaluation as a management tool

For evaluation to be used successfully as a management tool, programme actors have to have a clear understanding of the aims of this type of evaluation and their role in it (this process also dispelling mistaken expectations derived from more "traditional" evaluation practices). For example, programme actors should not perceive that they are “being evaluated”, but must instead feel that they are participating in a jointly owned analysis in which they have a stake. This helps to instil the trust and mutual respect which is essential for participants to criticise and be criticised. In addition to discussing the shortcomings of others, all participants should ready to examine their own implementation and management arrangements and to consider potential change.
Once partners have participated in an evaluation, their inclusion in future studies becomes easier, provided the initial process was beneficial. However, gaining acceptance and setting the tone for a first evaluation conceived as a management tool can be difficult. In illustration, during the 1994-1999 period, many partnerships participated directly in evaluations for the first time, and it took time to familiarise them, with evaluators often having to allay the fears of those from whom they were collecting information (e.g. reassuring them that the studies had different objectives, methods and status than audit and control exercises).

A range of approaches have been used to forewarn partners about the scope of evaluations, their expected participation, and how their feedback will be used. The utility of formative evaluation can be emphasised by stressing that evaluating the processes used to deliver an economic development programme is not an end in itself, and is only undertaken to address operational weaknesses and so make the programme itself more effective as an economic development instrument. In Tuscany, to underline the importance of formative evaluation in the 1994-1999 Objective 2 interim evaluations, and to give credibility and visibility to these dimensions, the evaluation team were presented as “process consultants”. As a result, their formative activities were seen as a central part of their remit rather than an incidental dimension of their work, helping to ensure that participants had clear expectations.

4.3 Squaring the circle: results informing future programming

Evaluations have no practical value if their recommendations are not followed through. Three factors contribute to this. The first, already discussed above, is actively involving the partnership in the process of evaluation rather than relying on the passive receipt of a final evaluation report prepared with little interaction. Two further factors are maximising the perceived relevance of evaluation and building in active dissemination and follow-up.

In the past, some Structural Fund evaluations at Member State or regional level were only conducted to fulfil EC requirements, and, once complete, were simply “signed off”, submitted to the EC to account for activity to date, and shelved. Follow-up at the programme level is more likely where evaluation is conceptualised as an integral part of the programming cycle, designed to feed relevant and timely information into the programming process, whether it be about the design of measures and projects or the management of the programme. If an evaluation has addressed issues of genuine concern to programme implementers and managers, then its recommendations are more likely to be eagerly awaited and implemented. Relevance can be ensured at the design stage and then later during evaluation when key issues emerge from discussion and analysis.

In terms of follow-up, formal steps can be taken to flag commitment to this and to show evaluation as part of a structured learning process. They range from adding follow-up to the agenda of Monitoring Committees to more elaborate approaches. In one Spanish 1994-1999 Objective 1 Operational Programme, for example, a panel of key partners was organised to discuss the outcomes of the interim evaluation, using five, structured three-hour sessions of discussions to give detailed consideration to evaluation findings.22

The wide dissemination of Structural Fund evaluation results to the diverse audiences involved in programming is an important dimension of follow-up. It is fortunately increasingly rare that evaluations are treated as sensitive documents for limited circulation, since such approaches undermine trust and openness and sacrifice opportunities to galvanise partnerships and harness their energy for delivering more effective programmes. Dissemination can be structured through a formal communication plan23, and might involve a range of activities. Any written materials should be tailored to different audiences depending on their range and level of interest in the programme: Only a few actors will want to see the whole study. Others may be more interested in a summary of the whole evaluation and/or detailed information about their areas of involvement. Dissemination should take this into account, including through the production of user-friendly summaries bringing out key findings in a lively way. An example is the Pays de la Loire 1994-1996 interim evaluation, whose findings were summarised under five key themes (pro-

---


gramme clarity, impacts, implementation, strategy and partnership) and organised succinctly under the headings of strengths, weaknesses and recommendations.

Among the most effective ways to bring an evaluation to life is to allow the evaluator to present the outcomes of a study in person to the partnership. This raises the profile of the evaluation and opens up dialogue. In some cases, these fora have been restricted to Monitoring Committee members, while in others, evaluation results have been presented to wider groups in the course of high-profile partnership events (e.g. in Western Scotland). A further practice facilitating the dissemination of evaluation findings – and reinforcing their position as part of a longer-term, structured process of learning – is to ensure that studies gain permanence by being published, for example in a relevant regional journal or website. This gives greater visibility and credibility to evaluation reports, and also increases transparency, making studies available on a systematic and permanent basis. Selected German regions are exemplary in publishing evaluations in the journals of regional research institutes.

For dissemination to have positive outcomes, it is critical that evaluation findings are verified for accuracy and fairness before a study is finalised and distributed, especially where the programme relies on the collaboration of its partners. In a complex context like the Structural Funds, if the findings of an evaluation study do not accord with partners’ perceptions of a programme, or if they contain inaccurate information, then the study is seriously undermined, and its potential to be accepted and followed up is diminished.

5 Evaluation for quality management: the issues for Germany

As set out above, the new round of Structural Fund programming will involve even greater evaluation activity than in the past, accompanied by increased pressure for effective management. The development of evaluation as a management tool will hold increasing potential.

In the German context, this practice has already seen uneven development, and overall has flourished less than in some other contexts. Evaluation is built into national regional policy and also undertaken where necessary to meet Structural Fund obligations, and the evaluation community is strong. In other ways, the context is more problematic. In particular, Germany has a “subsumed” system of Structural Fund delivery, the accent in many programmes being on delivering the Structural Funds through domestic institutional and policy channels, and minimising the need to create new or distinctive methods to deploy these resources. In turn, this means that Structural Fund programmes often have little profile as entities in themselves, limiting the chances that evaluating their design and management will be a high priority among the stakeholders or that this will be seen as a legitimate task for evaluation.

Nonetheless, this does not necessarily mean that evaluation conceived as a management tool is not possible in the German context. Notable exceptions demonstrate this (e.g. in Nordrhein Westfalen and Bremen, and under the Interreg programme). This focus is perhaps more likely when certain conditions apply: where the Structural Funds are conceived as an opportunity for regional policy innovation i.e. not entirely coupled with the main domestic regional policy instruments; where programmes have required the establishment of new management structures eg. Interreg programmes; and, where the Structural Fund financial allocation is larger, bringing greater political significance and visibility to the Funds and making the deployment of resources more complex and the potential cost of failure much greater.

A first step in improving process-oriented evaluation practices is to continue promoting understanding of the value and position of this type of evaluation. It is important at the same time that the view promoted is balanced – in particular emphasising that process evaluation complements and supplements rather than replaces impact evaluation. A second step is to continue to increase the credibility of all evaluation activity, both formative and summative. This includes improving the rigour of all methods used and promoting a better understanding of the strengths and limitations of different evaluation techniques so that evaluation options can
be chosen in an informed yet pragmatic way. It should be remembered here that participative evaluation techniques also present weaknesses and risks which have tended to receive limited attention.24

At a deeper level, changes in the way that evaluation is perceived could be beneficial. It needs to be seen not just as an externally driven product (required by the regulations) but as an on-going process meeting not only exogenous but also endogenous information needs. For this to be achieved, evaluation has to be appropriated, with every stakeholder asking what the evaluation can and should be doing for them. To facilitate this reorientation, it is highly recommended that programmes begin to integrate evaluation explicitly into their longer-term planning. The expected role of evaluation could be made an explicit theme of discussion from the earliest stages of programme development, and should remain on the agenda of all the structures established to steer and manage the programme in question – even where there is only one, in the form of the programme Monitoring Committee.

Finally, Structural Fund evaluation has an opportunity to influence the culture of regional development in Germany. Among some German regional policy-makers, there is considerable scepticism towards EU regional policy, especially at federal level, a scepticism which has perhaps obscured some of the positive policy effects of regional development programming under the Structural Funds. In the debates about the future of German regional policy, there is a growing recognition of the need to reform the Gemeinschaftsaufgabe. While there is no consensus on the way forward, Structural Fund evaluation can provide some objective insights into the effectiveness and efficiency of new ways of managing, resourcing and delivering regional development policies.