

45th Congress of the European Regional Science Association

23rd – 27th August 2005, Amsterdam

**A concept for evaluating innovation related actions under the
EU Structural Funds**

Tanja Fleischhauer

University of Dortmund, Faculty of Spatial Planning
Department of Economics, 44227 Dortmund, Germany
anja.fleischhauer@uni-dortmund.de

Abstract

Against the background of a new programming period of the European Structural Funds and the anticipated increase of the financial importance of innovation-related actions, the goal of this paper is to analyze the quality of the existing (general) evaluation concept of the European Union. This analysis aims to provide insights into the strengths and weaknesses of the evaluation concept in connection to innovation-related activities.

The alternative evaluation concept that will be presented in this paper is focused on innovation-related activities and is generally based on the present EU-evaluation concept. Core changes are proposed for the impact analysis as well as the intensification of the policy advice contribution of evaluation. Similar to the existing EU-evaluation concept, the proposed concept for innovation-related activities will strongly be based on indicators. But as the discussion of appropriate and reliable indicators is too complex to be made the subject of this paper, it focuses on the theoretical structure of the evaluation concept.

Keywords: Evaluation; Innovation; European Structural Funds

JEL-Codes: O30; R58

1 Introduction

On 1 January 2007 a new programming period of the European Structural Funds is going to start (2007-2013). The negotiations about the contents and the regional structure of the funds are still in progress but the European Commission has already published her thoughts in the Third Cohesion Report and in some draft regulations (e.g. COM 2004/492 for general regulations; COM 2004/493 for the ESF and COM 2004/494 for the Cohesion Funds).

Main elements of the Commission's proposal are the concentration of funding for the least developed regions and Member States, the thematic concentration on the strategies of Lisbon and Gothenburg as well as institutional capacity building (KOM 2004: XXXVI). The Commission is willing to introduce a new priority ('Regional Competitiveness and Employment', former 'Objective 1') as a successor of Objective 2 to strengthen the regional competitiveness and the employment. The new priority follows the Lisbon Strategy and therefore mainly promotes innovation and the knowledge society (ERDMENGER/ZIEGLER 2004: 327).

The Commission's proposal makes clear, that the evaluation of structural funding is – as it has been in the past programming periods – a necessary condition to achieve the quality standards of funding. This means that ex-ante-, mid-term- as well as ex-post-Evaluations remain obligatory in the future (KOM 2004: XXXVII). These so-called overall-evaluations comprise all projects and measures of the intervention, no difference if they are environment-, employment- or innovation-related. A deeper (thematic) evaluation of all innovation-related actions is not obligatory to be accomplished and remains voluntarily.

Evaluations suffer from a lack of statistical information on the regional level as well as from difficulties in comparing the regions' funding achievements (e.g. TOEPEL 2000: 400; BEYWL/TAUT 2000: 359). These problems are especially true for the evaluation of innovation-related actions because the general problems are accumulated with problems of the measurability of innovation (PERRIN 2000: 5FF; DIEZ 2001: 912FF; AUTIO 1998: 132).

If the plans of the commission happen to turn into law on 1 January 2007, the financial importance of innovation-related actions will increase as well as the political requirement to measure and value their effects. Thus it is likely that the current general evaluation-concept and evaluation-practise will not be able to cope with the coming necessity, problems and challenges to evaluate innovation-related actions.

Against this background the goal of this paper is to contribute to a concept to evaluate innovation-related actions under the EU Structural Funds.¹ The alternative concept is based on the strengths and weaknesses of the present general EU-evaluation concept. Similar to the existing EU-evaluation concept, the proposed alternative concept will strongly be based on indicators. But as the discussion of appropriate and reliable indicators is too complex to be made the subject of this paper, it focuses on the theoretical structure of the concept.

Chapter 2 provides an overview about some basic principles of evaluation as well as innovation. To identify strengths and weaknesses of the present EU-evaluation concept, two case studies are analysed (Chapter 4), which both deal with the evaluation of innovation-related actions under the EU-structural funds (1994-1999). The analysis is based on three different criteria, which are introduced in Chapter 3. The evaluation concept itself is introduced in Chapter 5. Chapter 6 finally summarizes the main findings of the paper and draws conclusions.

2 Basic principles of the Evaluation of innovation-related actions in the EU

Referring to the European Commission, Evaluation is „judgement of interventions according to their results, impacts and the needs they aim to satisfy“ (COM 2000/1051: 3). The evaluation fulfils two general goals: First, all evaluations are to examine the efficiency of an intervention for the purpose of accountability or financial allocation. Beyond that, an evaluation is to ensure an improvement of the implementation and the management of interventions (COM 2005: 4-5).

Since the 1980s the European Union has been promoting the use of evaluations in all policy ranges and – as a result – as well in the individual Member States. While evaluation as a research and policy field has gained increasing attention since the early 1990s, the effects of innovation-related activities on the regions' economic development and the increase of competitiveness are still evaluated insufficiently. As a consequence the evaluation of interventions of the Structural Funds as a whole will be obligatory in the programming period 2000-2006. In contrast, thematic evaluations like the effects of innovation-related activities are still not intended and are so far mainly integrated into the overall-evaluations.

¹ As referred to later this does not mean that an alternative concept is a guarantee to measure all effects of innovation-related activities. The aim of the proposed evaluation concept is to improve the practical usage and the quality of evaluations and find a better way to cope with the problems caused by the complex and vague character of innovation.

In addition to the integrated appraisal, evaluations of innovation-related activities are charged by the European Commission. In these cases it is not an evaluation of effects of *individual* regions, but for *all* regions promoted by Structural Funds in the European Union so as to identify the added value for the Union.

From the Commission's point of view these so-called "meta-evaluations" are not only important for evaluating the efficiency of assigned capital. The Commission is also obliged to report the performance of the Structural Funds to the European Parliament, the Council, the Economic and Social Committee and the Committee of the regions annually before 1 November (e.g. COM 2003/646). Further the Commission reports the progress of the economic and social cohesion and the contribution of the Structural Funds every three years. This happened last in 2004 with the third Cohesion Report. (COM 2004).

In fact, the thematic evaluation or the commission of meta-evaluations of structural funding is voluntary for the funded regions and the European Commission as the funding institution. But the more the financial assistance for innovation by the Structural Funds increases the stronger the necessity is to measure its effects separately and not only as a part of the overall evaluation (ERDMENGER/ZIEGLER 2004: 326-327; KOM 2004: xxv-xlii). Only then it is possible to gain an overview about effective and non-effective strategies in order to adjust future policies.

Besides the chances to gain additional insights and recommendations for future programming periods the preparation of evaluations means that evaluators have to be aware of some problems and risks of evaluation in general as well as the evaluation of innovation-related activities (Table 1; for more information see e.g. KUHLMANN/HOLLAND 1995; KUHLMANN/BÜHRER 2000; SCHUBERT/SCHUH 2004; WIECHMANN/BEIER 2004). Main problems to be referred to are the measurement of impacts and their separation from deadweight. These problems especially apply for evaluations of innovation-related actions because innovation is complex, uncertain, long-term and produces manifold effects that can not be measured linearly (DOSI 1988: 222-223; DIEZ 2001: 912).

In consequence, these problems and circumstances have lead to the realization that evaluations can not be "perfect" (ROSSI/FREEMAN/HOFMANN 1988: 94-95). The goal must rather be to achieve an optimal evaluation under the given conditions.

Table 1: Problems of the evaluation of innovation-related actions

	Problems
General problems of evaluation	• Results of evaluations cannot be compared to each other due to a great variety of methods.
	• Absence of exact objectives to measure the effects of a program (the verbalisation of objectives needs a consensus between all actors).
	• Objectives of a program are often political objectives. These are often kept consciously low in order to be achieved anyway.
	• In spite of progress in the availability of program-concerned data, main indicators are often not available.
	• Importance of qualitative or quantitative methods.
	• Disconnection of actual effects and dead-weight.
	• Subjective assessment criteria which are often not comparable to each other.
Specific problems of the evaluation of innovation-related activities	• Innovation is complex and uncertain; there is no „guarantee“ that public resources can provide innovation.
	• Innovation can only be reached on the long run, but policy and society ask for short-term efficiency.
	• Innovation causes complex and multiple effects that do not evolve linearly. Their evaluation is in a way delicate and makes a linear impact analysis impossible.

An important element to generate an optimal evaluation is the basic structure of an evaluation concept. Therefore the analysis of the quality of existing evaluation reports is crucial in Chapters 3 and 4. In Chapter 3, the criteria for the analysis of innovation-related evaluations will be characterized; Chapter 4 describes the main findings of the analysis of two meta-evaluations, in which the effects of innovation-related activities in the European Structural Funds have been assessed.

3 Criteria for the analysis of innovation-related Evaluations

The analysis of the existing meta-evaluation reports is based on three criteria groups that will be presented in this chapter:

1. General demands on evaluation,
2. EU-specific demands on evaluation,
3. Consideration of potential objectives of Innovation policy.

The first two criteria groups are standards of a high-quality evaluation, published by the European Commission and a voluntary self-commitment of evaluators, respectively. The third criteria group belongs to the field of innovation and has been developed to verify if the whole range of possible innovation policy is considered in the evaluation. In other words, the evaluation's ability to draw exact conclusions from single parts of innovation policy is checked.

Criterion group 1: General demands on evaluation

During the 1980s and 1990s various work has been published to propose demands on evaluation mainly in the United States (BEYWL/TAUT 2000: 360). The best known work are the Standards for programme evaluation, published by the „Joint Committee on Standards for Educational Evaluation“ (JOINT COMMITTEE 1981; JOINT COMMITTEE 1994). Meanwhile, they are the acknowledged consensus of the evaluation research community regarding the quality control of evaluation. The German as well as the Swiss association on evaluation e.g. adopted or modified these standards.

The standards² shall improve and ensure the quality of evaluations. In addition, they give concrete advice about planning and accomplishment of evaluations (DEGEVAL 2002: 6). The standards are based on the presumption that evaluations have to be useful, feasible, proper and accurate to meet the demands of policy, science and the public (WIDMER 2004: 93). The US-standards have 30 single standards to describe their four main objectives. A detailed list of all standards is shown in Table 2.

All seven *Utility Standards* are to ensure that an evaluation serves the information needs of intended users. Main aspects are the identification of all stakeholders involved, the evaluator's credibility and the identification of the evaluation's values. Further, the selection of information and the clarity of the report are of great interest as well (JOINT COMMITTEE 1999: 21-24).

An evaluation is useful, if the user of an evaluation effectively uses the results of an evaluation to improve the programme. This in a way meets the formative or summative function of an evaluation (BEYWL/JOAS 2000: 87).

The three *feasibility standards* are intended to ensure that an evaluation is realistic, prudent, diplomatic and frugal. This means that the evaluation procedure should be adequate, as well as the role of the cost-effectiveness-relation and the production of new information. In addition, the political viability is important as this calls for acceptance of all concerned and involved persons (JOINT COMMITTEE 1999: 87-106; DEGEVAL 2002: 26-28). An evaluation is feasible if the basic conditions of the programme implementation are regarded in a way that

² This applies not only to the American but as well to the German and Swiss standards.

the intervention is possible without major interferences of e.g. problems in the data-collection. An evaluation is as well feasible if the used methods are practicable (BEYWL/JOAS 2000: 87).

Table 2: The general evaluation standards at a glance

The general evaluation standards at a glance	
Utility	Feasibility
Stakeholder Identification	Practical Procedures
Evaluator Credibility	Political Viability
Information Scope and Selection	Cost Effectiveness
Values Identification	Accuracy
Report Clarity	Program Documentation
Report Timeliness and Dissemination	Context Analysis
Evaluation Impact	Described Purposes and Procedures
Propriety	Defensible Information Sources
Service Orientation	Valid Information
Formal Agreements	Reliable Informationen
Rights of Human Subjects	Systematic Information
Human Interactions	Analysis of Quantitative Information
Complete and Fair Assessment	Analysis of Qualitative Information
Disclosure of Findings	Justified Conclusions
Conflict of Interest	Impartial Reporting
Fiscal Responsibility	Metaevaluation

Source: Joint Committee (1999: 47-220)

There are eight *propriety standards* which shall guarantee that an evaluation is conducted legally, ethically and with regard to the welfare of those involved in the evaluation as well as those affected by its results. This also means that the results of an evaluation are accessible for all involved and concerned persons (JOINT COMMITTEE 1999: 107-154; DeGEval 2002: 28-30). An evaluation is ethical, if it meets the laws of a country, cultural characteristics and respects the values and beliefs of all persons involved. If one person acts criminally, this behaviour must not be covered by the evaluation nor encouraged (BEYWL/JOAS 2000: 87).

Twelve *accuracy standards* are intended to ensure that the evaluation will reveal and convey technically adequate information about the features that determine worth or merit of the programme being evaluated. In this regard it is necessary to use valid and reliable methods to gain data and to secure the analysis of technical questions of an evaluation. Qualitative and quantitative information have to be systematically analysed just as well by taking into account professional criteria (JOINT COMMITTEE 1999: 155-220; DEGEVAL 2002: 31-36). An evaluation is accurate if the used methods meet the quality-criteria of empirical social research, such

as validity and reliability. This is true even if the rating of the criteria has to be adjusted to the special practical needs (BEYWL/JOAS 2000: 88).

Within the scope of the analysis of existing evaluation reports it is to be verified, in which way the general evaluation standards are met. The assessment is accomplished verbally, because neither the description of demands, nor the evaluation report itself is an adequate basis for a formal or quantitative assessment.

Criterion group 2: EU-specific demands on evaluation

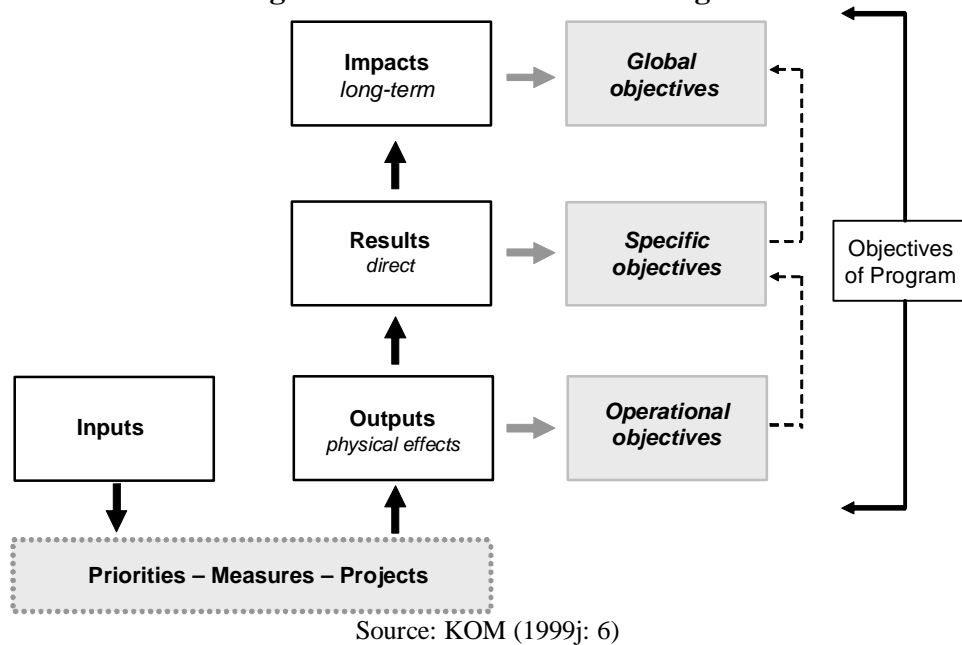
The second criterion group for the analysis of existing evaluation reports has been compiled by the European Union. The demands are valid for every Structural Funds-related intervention and became necessary in order to have at least a small chance to compare the quality and content of the huge amount of regional evaluation reports. The theoretical structure of this EU evaluation concept is oriented at the intervention logic of all Structural Funds related interventions (KOM 2005: 4).

One of the main requirements of the intervention logic is the agreement on different objectives and measures. The dissemination of the resources can be carried out in two different ways: Either from top -down or bottom-up (see Figure 1):

Top-down: All interventions are planned in regard to a *global objective*. This global objective is equivalent to the level of the regional programme or its priorities which are subdivided into measures. Measures themselves are represented through several *specific objectives*. Every specific objective is implemented in numerous different projects, which on their part fulfil *operational objectives* (COM 1999b: 24).

Bottom-up: Every project or measure of a regional programme can only be accomplished with (financial) *input* from the development agency. This input leads to a physical *output* which is under direct control of the manager. The output indicates direct *results* for the beneficiary of the project. These results of an intervention however also have long-term effects and are not limited to the beneficiary of the project. These long-term effects are characterized as *impacts* of an intervention.

Figure 1 shows that operational objectives are expressed through outputs of single projects, specific objectives through results of measures and global objectives through impacts of priorities or the whole programme.

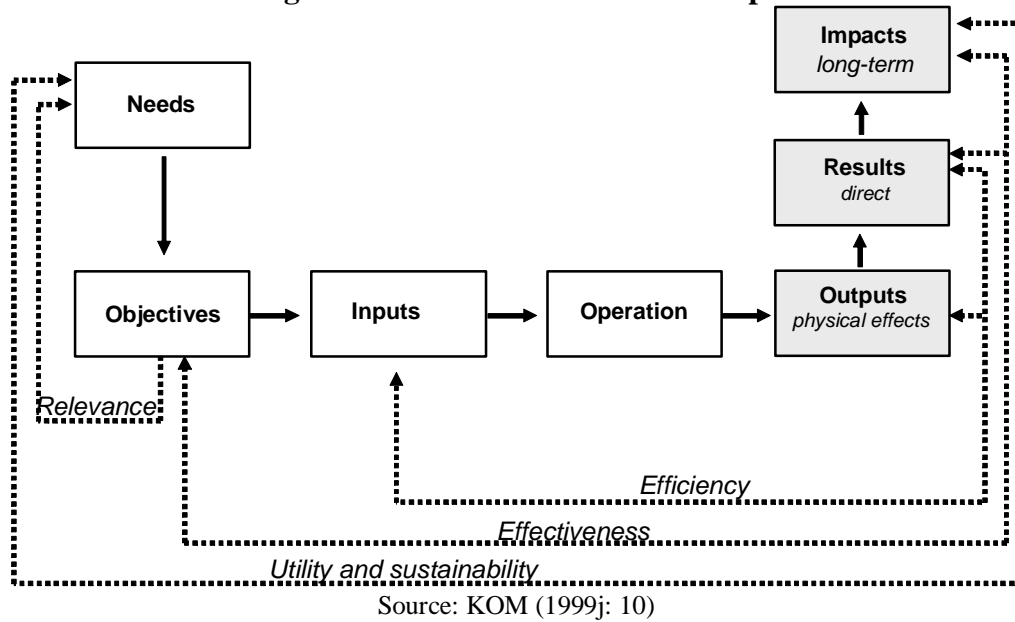
Figure 1: The EU-Intervention logic

The described intervention logic shows, assigned to the evaluation of EU-Structural Funds intervention, that defined operational, specific and global objectives have to be verified in terms of their attainment. The real measurement of effects from operational to specific and global objectives is done by measuring output, results and impacts.

Besides the general intervention logic and structure of evaluations it is considerably important *what* should be evaluated or what general criteria is applicable. The European Commission has disposed five evaluation criteria, which should help to measure the effects of interventions: relevance, efficiency, effectiveness, utility and sustainability. Their assignment in the intervention logic is presented in Figure 2 (COM 1999j: 9-12; COM 1999a: 70-73).

The evaluation criteria *relevance* is based on the objectives of the programme. Before an intervention starts it has to be checked if the intervention is relevant in respect to the needs, problems and issues identified in the target group. This audit has to be repeated some time after the initial implementation. It has to be assessed, if the strategy remains relevant, given the possible evolution of the situation.

A major element in judging the success is to assess the *effectiveness* in terms of the progress made towards the attainment of pre-determined objectives. It is checked to what extent the effects that are induced by the intervention correspond with its objectives as they are outlined in the intervention strategy. The assessment of the effectiveness can as well ask for the adequacy of the measures and the programme for the determined objectives.

Figure 2: The EU-Evaluation Concept

In terms of an evaluation one speaks of *efficiency* whenever it is a question if the resources used were converted into effects: Is the ratio good? Could the same result or impact have been achieved with a lower input? Could a higher result or impact have been achieved?

The criterion of *utility* compares the effects of an intervention with the wider needs of the target group in the funded regions. Over and above the effects that correspond with the stated objectives of an intervention, other effects may occur that are either negative or positive. An assessment of these effects provides the basis for a broader assessment of performance on the basis of an intervention's utility. Further, the criterion of utility is rather abstract and is sometimes referred to as "goal-free" evaluation, as the criterion can be defined by the evaluator.

The criterion *sustainability* is in a way similar to the criterion utility. It does refer to the original objectives of the programme and tries to give answers if the effects of the programme are lasting changes within a target group, geographical zone and so on. An assessment of these effects provides the basis of the sustainability of an intervention's effect. The use of the term "sustainability" is in a way confusing, because it mainly stresses the permanence of the intervention's effects, whereas the term originally suggests the future prospects.

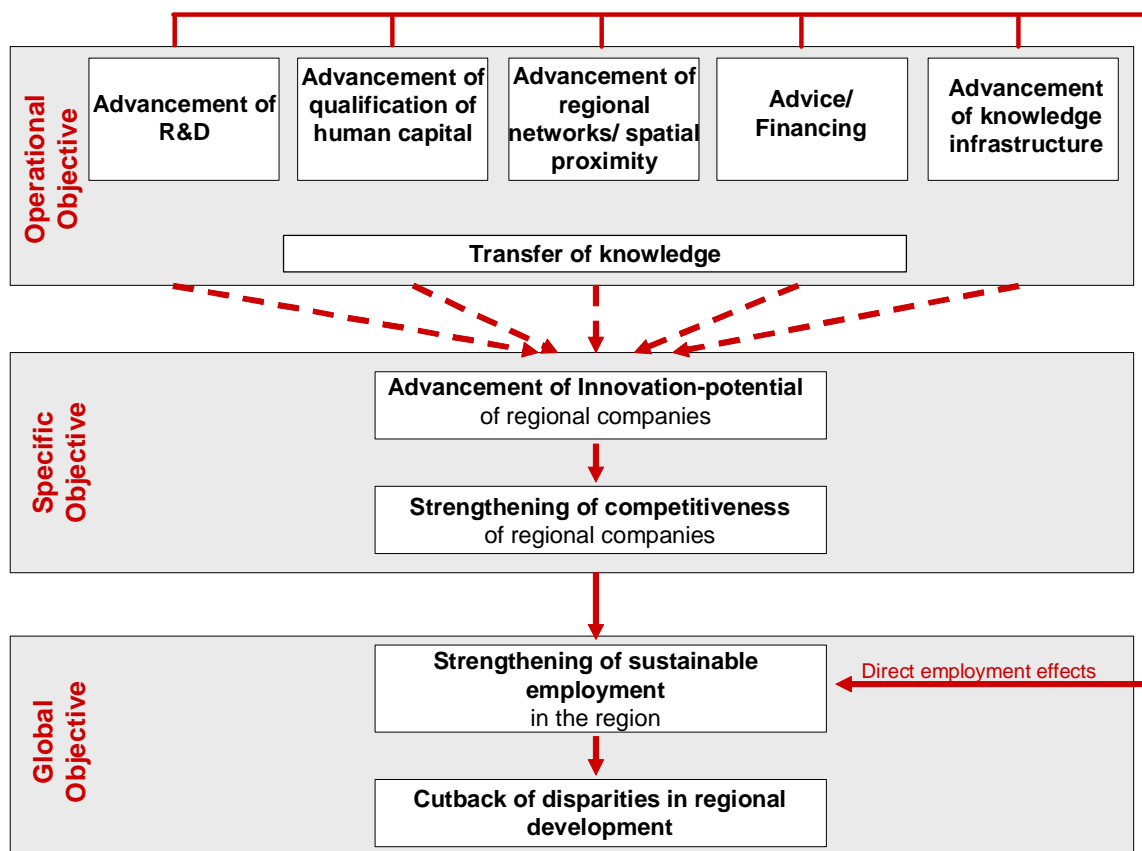
All five evaluation criteria are obligatory for all evaluations of EU-Structural Funds. For voluntary evaluations there are obviously no such commitments. The analysis in Chapter 4 reviews if the evaluation criteria and the different layers of objectives are considered in the evaluation reports. The analysis itself is realised by a simple yes/no-decision. For those as-

pects of the EU-evaluation concept which are not considered in the evaluation reports (a) the reasons are checked as well as (b) the question is asked if this is leading to an advancement or deterioration of the evaluation's quality.

Criterion group 3: Consideration of potential objectives of Innovation policy

The third criterion group – the consideration of potential objectives of innovation policy – gives an insight how intensely evaluation reports deal with innovation policy, e.g. how consistently they assign effects of innovation-related funding to different objectives.

Figure 3: Objectives of Innovation-policy within the EU-Structural funds



From the innovation-related literature³ determinants of innovation can be derived. The EU-Structural Funds are also oriented at these innovation determinants in general. They can be compared with operational objectives that are aimed at by innovation-related projects (see criterion group 2). Figure 3 hereupon shows the operational objectives „research & develop-

³ For details see e.g. Brockhoff 1999; OECD 2002; Nonaka/Takeuchi 1997; Crevoisier/Maillat 1991; Camagni 1995; Freeman 1991; Genosko 1999, Koschatzky 2001; Freeman/Soete 1997

ment“, „qualification of human capital“, „regional networks“, “advice and financing of entrepreneurs for innovation-related projects” and “strengthening of knowledge-infrastructure”.

The transfer of knowledge – which is a key determinant in the innovation process – (AUDRETSCH/FELDMAN 1996; GRUPP 1997) plays a major roll because it is an effect of a successful funding of the operational objectives. Beyond that, knowledge transfer can achieve a strengthening of the regional innovation energy, which is a specific objective. The global objective of EU-Structural funding is the cohesion of regional disparities or the allocation of sustainable employment (KOM 1999j: 6).

The analysis of the third criterion is separated into two different steps. In the first step the consistency of the indicators used in the evaluation reports and the operational objectives are checked. This means it has to be tested, if they can measure the effects of every single innovation-related objective of the programme. In a second step it has to be checked, if the operational objectives of the evaluation report are identical with the above mentioned operational objectives to make sure that all possible operational objectives are paid attention to.

4 Findings of the analysis of innovation-related evaluations

All three criterion groups mentioned in Chapter 3 are the basis for the following analysis. The aim of the analysis is to examine if the existing general evaluation concept is adequate for evaluating innovation-related activities and if the concept is feasible. The analysis covers two evaluation reports:

1. ADE/ZENIT/ENTERPRISE PLC (1999): Evaluation of Research, Technological Development and innovation-related actions under Structural Funds (Objective 2),
2. HIGGINS/TSIPOURI/VAN DER LANDE (1999): Thematic Evaluation of the impacts of Structural Funds (1994/1999) on Research, Technology Development and Innovation (RTDI) in Objective 1 and 6 Regions.

There are four reasons, why both evaluation reports are sound case studies for the description and analysis of current evaluation concepts for innovation-related activities within the EU-Structural funds: (1) Both case studies explicitly deal with innovation-related activities within the EU-Structural Funds and therefore work towards the objective of this paper. (2) Both evaluation studies are meta-evaluations of obligatory overall-evaluations of all funded regions in the programme period 1994-1999. Therefore, the results of the regional evaluations are

enclosed in the meta-evaluation and it can be judged, which efforts for regional evaluations have been made and where their strengths and weaknesses are. (3) Both evaluation reports are generally comparable as they deal with the same topic. Beyond, they both have the same customer (the European Commission), so budget and complexity of the evaluation are comparable as well as the interests, the Commission is safeguarding. (4) It can be assumed, that the European Commission has adapted her own demands on evaluation to both job descriptions. In this way, both analysed studies portray the actual discussion about evaluations and evaluation concepts within the European Commission at the end of the last programming period.

Table 3: Review of existing evaluation reports

Review of existing evaluation reports
Evaluation standards are adhered in general
Evaluation methods are broad
Difficulties with the availability of data
Difficulties with the comparability of data
Unclear presentation of the evaluation and its results
Conclusions are often not deducted from the analysis
è Intransparent setup of the reports as well as the deduction of conclusions
EU-specific evaluation standards adhered only partly
Orientation at the EU-Intervention logic, but no description of the programme's results
Impacts of the funding only adhered abstractly
Hardly observation of evaluation criteria
Availability of data makes the aggregation of effects almost impossible
No feedback to objectives of the funding
è No evaluation in the sense of meaning
Operational objectives not adhered consequently
All-embracing understanding of innovation; operational objectives are also important in the analysed reports
But: Evaluation of operational objectives only with inputs, not with the effects (output, result, impact)
è Conclusions about the success of measures hardly possible è no precise deduction of conclusions

The analysis of the two evaluation reports discussed above shows divergent results for all three criterion groups. Both evaluation reports show similar strengths and weaknesses, so a separate description of the analytical result seems to be dispensable. As shown in Table 3, the *general demands on evaluation* are met in principle. One main positive aspect is the comprehensive evaluation method used in both evaluation reports, which makes use of both qualitative and quantitative data. At the same time the used methods and ascertained results are not presented clearly. In addition the evaluation report's conclusions are not completely drawn off

their analysis results. One main reason for this weakness seems to be the poor availability and comparability of the data.

While the *general demands on evaluation* are met in principle, the *EU-specific evaluation demands* are only met for the minor part of the evaluation reports. Certainly the evaluations are based on the EU-Intervention logic, but the effects are only portrayed with the output of the funding and a very short and abstract description of the impacts – an analysis of the results of the innovation-related actions is missing. The evaluation criteria *relevance, effectiveness, efficiency, utility* and *sustainability* are hardly paid attention to. Beyond, there is no reference to the objectives of the funding, and thus both evaluation reports remain at the level of an analysis – a real evaluation in its sense does not happen.

The understanding of innovation in both evaluation reports as well as their analysis of the funding's input makes clear that the *operational objectives* mentioned in criterion group 3 are regarded to be relevant for the achievement of innovation within the reports. However, the effects of innovation-related funding are only mentioned on the whole and don't apply to the operational objectives. This missing consideration of operational objectives results in the absence of statements about the effect(s) of single innovation-related projects or measures. Consequently, it is impossible to give policy advice for the future funding of innovation.

As a whole the analysis shows that a separated identification of effects (output, result, impact) is useful to portray the different layers of the programme's objectives. Furthermore the analysis demonstrates that the examination of the evaluation criteria *relevance, effectiveness, efficiency, sustainability* and *utility* is absolute for good evaluations, which are aimed to go beyond a simple analysis of effects. Although the EU mainly claims quantitative indicators, the evaluation reports show that the use of quantitative as well as qualitative indicators and methods does improve the deepness of the evaluation's results.

A general problem of both evaluation reports is the high variability of the available innovation-data of each funded region. This makes an aggregation of the effects of all funded regions – in order to draw conclusions for the whole EU or to compare regions – hardly possible. Independently from the analysis it is clear that – at the same time – the complexity of innovations (and regional innovation strategies) can not be portrayed with standardised indicators alone. Mainly because of the poor data availability and the missing comparability of indicators both evaluation reports do not focus sufficiently on the cause analysis of a good or

bad performance of the innovation-related funding and the deduction of policy advice for future funding.

With reference to the EU-evaluation concept it turns out that some main demands on high-quality evaluations for innovation-related actions are already integrated in the concept. At the same time it is distinct that these demands – especially those of the EU –strongly aim at the analysis of effects. Both evaluation reports only describe the effects partly, because of the poor quality of data, so that the deduction of policy advice is presently not sufficient or transparent. Even though good evaluations contain these aspects, the current evaluation concept seduces from paying less or no attention to these central issues of evaluation.

The weaknesses of the present evaluation concept indicate that there have to be made some changes in order to be applicable for innovation-related activities and to strengthen the policy advisory function and the quality of evaluations.

5 Alternative concept for evaluating innovation-related actions

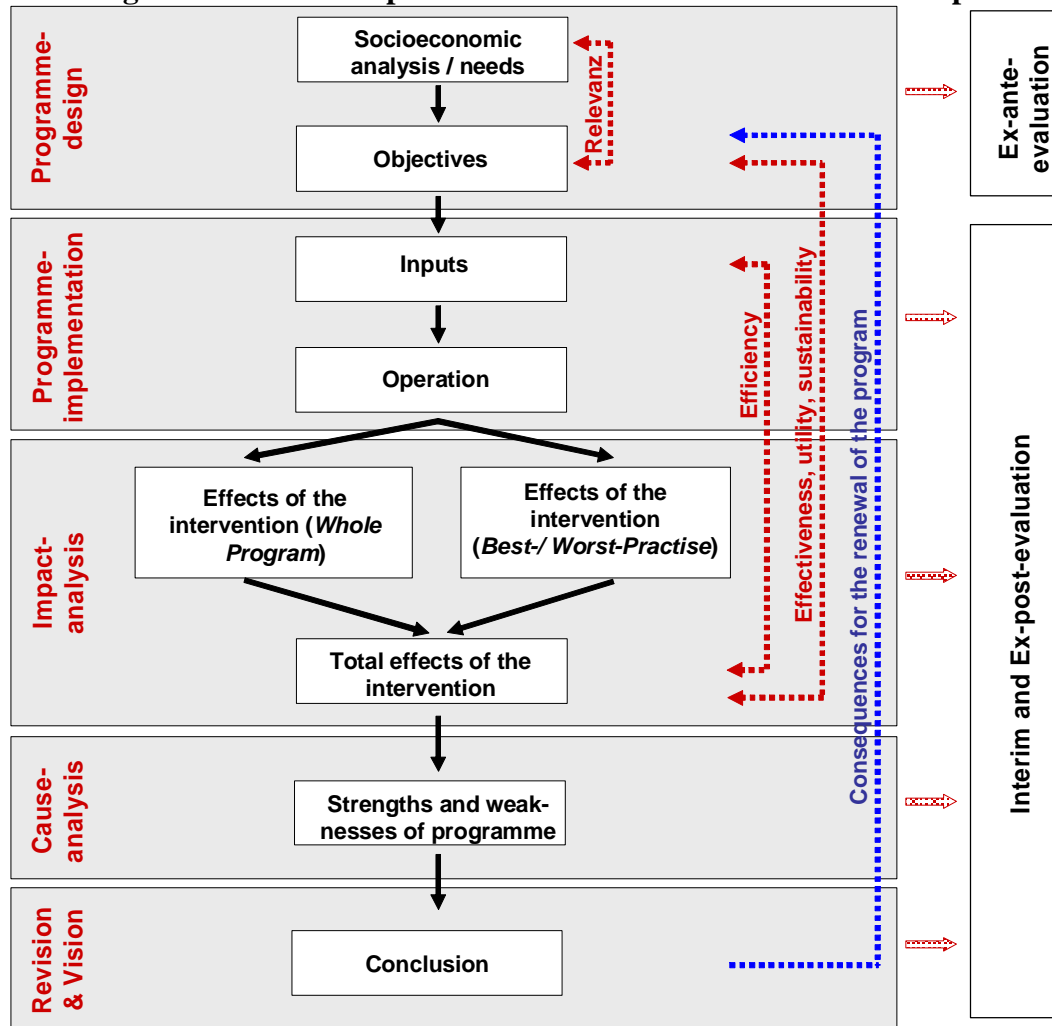
The alternative evaluation concept for innovation-related actions starts from the existing strengths and weaknesses of the present general evaluation concept and its implementation, respectively. It is structured into five evaluation phases (Figure 4):

1. Programme conception

Within the programme conception the aim of the evaluation is to evaluate the results of the socioeconomic analysis as well as the identified call for action (evaluation of relevance in the ex-ante analysis). The phase of programme conception is essentially oriented at the present evaluation concept.

2. Programme implementation

The evaluation is not directly included in the implementation of the regional programme because this implementation is a task of the responsible regional authorities. However, the funding input has to be evaluated as a basis for the evaluation in order to estimate the relation between the input and the effects (evaluation of efficiency). The phase of programme implementation is also oriented at the present evaluation practice.

Figure 4: Evaluation phases of the alternative evaluation concept

3. Impact analysis

The phase of impact analysis is of central concern for the assessment of the intervention's effects. Thus, an impact analysis will be introduced that is structured into two parts as shown in Figure 4. The effects are categorised as output, result and impact in order to portray different levels of objectives. The two parts are the following:

Ø *Partly standardisation of the assessment of the intervention's effects (entire programme):*

For the analysis of effects on the programme level, quantifiable indicators will be determined that shall be assessed EU-wide for any intervention with innovation relevance, independently from the actual (financial) amount of the funding.

Apart from the interregional comparability of indicators, a determination of standardised core indicators may further lead to an improvement of the implementation process as well as the quality of evaluations. This is due to the fact that standardised indicators already

provide a general evaluation framework and therefore all essential demands are easier to fulfil (REID 2001: 408). This also means that standardised indicators may present all areas of objectives (operational, specific, global objectives) and in this way conclusions can be drawn on the success of the selected strategy.

Moreover, the standardisation leads to a lower extent of time and resources for the derivation of indicators during the phase of the programme conception (REID 2001: 408). Instead, the collection of data and their analysis shall be intensified in order to achieve satisfying evaluation results.

Beyond the core indicators, any region may analyse as many indicators as it wants that are not determined. As core indicators alone do not provide a comprehensive overview of the impacts, the derivation of additional indicators in general will be necessary. Due to the fact that the European Commission will examine the completeness of indicators in the future it can be assumed that additional indicators will most likely be considered. Similar to the actual evaluation practice the number of indicators shall not be prescribed. The aim should rather be to assess the impacts comprehensively, regardless of the number of indicators.

Ø *Open assessment of the intervention's effects (best- und worst practice projects):*

In contrast to the partly standardised evaluation of the entire impacts, the analysis of best and worst practice projects is made in a deliberately open process. This relates to the selection of projects as well as to the way of analysis of their results (e.g. quantitative, qualitative, evaluation method).

Depending on the structure of the regional programme and on the proceeding of the programme implementation all those regional projects are selected that might be – according to the regional actors – a basis for deriving generalised statements about the entire programme or the innovation policy as a whole. It is important though, that the selection of projects includes successful (good practice) as well as unsuccessful projects (worst practice) of the funded region because both groups help to improve the knowledge (e.g. AUTIO 1998: 137; DIEZ 2001: 915). Due to the expert knowledge of the regional actors the selection of projects does not only portray regional specifications but also guarantees the usefulness of the results for the continuation of the funding or the conception of a new funding activity.

The evaluation concept should not be the frame where textual or methodological objectives for the analysis of best and worst practice examples are determined. Such a restriction is in contrast to the general condition that regional specifications of a programme shall also be portrayed in the evaluation. Besides the implementation of the programme, the analysis of best and worst practice examples by the responsible authority could additionally deal with qualitative effects of the innovation policy that can hardly be shown by the core indicators.

The entire effects of the innovation policy can be identified by both parts of the impact analysis. Thus, conclusions on the efficiency of the funding can be drawn regarding the used resources but also regarding the analysis of impacts, benefits and the sustainability of the funding concerning the objectives. Altogether, the twofold character of the impact analysis combines the advantages of standardised as well as open evaluations and thus can provide an improved measurability of the effects: The standardisation of evaluations on the one hand guarantees their comparability and a high quality. On the other hand only an open and flexible concept leads to results that in the end are useful for the persons responsible for the programme (FRITZ/HUTSCHENREITER/STURN 1997: 2).

4. Cause analysis

In the phase of cause analysis the results of the previous evaluation phases are gathered. Besides the results of the impact analysis, also the identified input belongs to this phase. The cause analysis shall identify the reasons for a particularly successful or – in contrast – an insufficient programme implementation. The cause analysis phase is not a part of the present evaluation concept. In practice, the reasons of the existing effects are often not (or not sufficiently) taken into consideration.

5. Revision & vision

In terms of the user's utility of an evaluation, the evaluation phase "revision & vision" is extremely important. In addition to the real effects of an intervention and the analysed causes the evaluators draw off policy advice for future funding. The stronger emphasis of this evaluation phase is a central issue for the evaluation, because the analysis of available evaluation reports has shown, that policy advice is only partly drawn off the evaluation results.

In addition, the evaluation phase revision & vision introduces a supplementary feedback loop from the conclusions of an evaluation back to the objectives of a possible new funding period.

The already existing feedback loops *relevance*, *efficiency*, *effectiveness* as well as *utility* and *sustainability* of funding also remain in the evaluation concept. They guarantee that not only the quantitative or qualitative effects of the funding are measured, but that they are as well assessed in the background of the initial situation and the outlined objectives.

It is true that this feedback loop is even currently demanded in general evaluation standards (e.g. the evaluation criteria user's utility), but in real-life this demand is not reliably covered as KUHLMANN (1997: 17) stresses. Therefore an integration of this feedback loop into the alternative evaluation concept can more easily and insistently stress its' requirement and the circular flow of an evaluation (TAYLOR/BACHTLER/POLVERARI 2001: 352).

6 Main findings and conclusions

In summary, it shall be stated that the evaluation concept for the evaluation of innovation-related activities is mainly based on the current general evaluation concept of the EU. Therefore the evaluation criteria *relevance*, *efficiency*, *effectiveness*, *utility* and *sustainability* are taken over from the evaluation as well as the separated analysis of the funding's effects in output, result and impact. Additionally, it is still possible to have a free choice of evaluation methods.

Main changes of the innovation-related evaluation concept towards the existing general one occur for example regarding the stronger emphasis of single evaluation phases, in particular the phases of cause analysis and revision & vision. Together with the feedback of the evaluation results to necessary changes of future funding, these evaluation phases should contribute to the stronger emphasis of the policy advisory function of evaluation.

The introduction of a twofold impact analysis (partly standardised and quantitative measurement of the entire programme effects as well as open and qualitative measurement of best- and worst practise) should ensure that the effects of the funding are not only measured quantitatively for the entire programme, but as well qualitatively and deepened for important single fields (projects or measures) of the programme. In addition, the twofold impact analysis allows to compare the effects of different regions to each other as well as to describe specific features of the region's innovation strategy and its impacts.

In summary it may be said that the alternative evaluation concept only means little additional work for most of the funded regions. For those regions that already provide high-quality

evaluations, the determination of core indicators could even be labour-saving. Beyond, it can be estimated that the results of the evaluations are more consistent in terms of the assessment of funding effects on the innovation potential and the innovation activity of regions. This holds out the prospect that evaluation of innovation-related actions will less be interpreted as a nuisance but as a possibility to recognise weaknesses in the own political actions and put improvements into them.

7 References

- [ADE/Zenit/Enterprise 1999] ADE; Zenit GmbH; Enterprise plc (1999): Evaluation of Research, Technological Development and Innovation-related actions under Structural funds (Objective 2); Final Synthesis Report; European Commission (Hrsg.); Brussels.
- [Autio 1998] Autio, Erko (1998): Evaluation of RTD in Regional Systems of Innovation; in: European Planning Studies, Vol. 6, No. 2; S. 131-140.
- [Beywl/Joas 2000] Beywl, Wolfgang; Joas, Steffen (2000): Evaluation ist unnatürlich! Eine Einführung in die nutzenfokussierte Evaluation entlang eines Seminars von Michael Q. Patton; in: Clemens, Wolfgang; Strübing, Jörg (Hrsg.): Empirische Sozialforschung und gesellschaftliche Praxis; Opladen; S. 83-100.
- [Beywl/Taut 2000] Beywl, Wolfgang; Taut, Sandy (2000): Standards: Aktuelle Strategie zur Qualitätsentwicklung in der Evaluation; in: Vierteljahreshefte zur Wirtschaftsforschung, Vol. 69, Heft 3; S. 358-370.
- [Brockhoff 1999] Brockhoff, Klaus (1999): Forschung und Entwicklung; Planung und Kontrolle; München, Wien, Oldenbourg.
- [Camagni 1995] Camagni, Roberto (1995): The Concept of Innovative Milieu and its Relevance for Public Policies in European Lagging Regions; in: Papers in Regional Science, Vol. 74, Heft 4; S. 317-340.
- [COM 1999a] European Commission (1999): Evaluating socio-economic programme: Evaluation design and management; Volume 1; Luxemburg.
- [COM 1999b] European Commission (1999): Selection and use of indicators for monitoring and evaluation; Volume 2; Luxemburg.
- [COM 1999j] Europäische Kommission (1999): Der neue Programmzeitraum 2000-2006: Methodische Arbeitspapiere; Indikatoren für die Begleitung und Bewertung: Eine Indikative Methode; Arbeitspapier 3; Luxemburg.
- [COM 2000/1051] European Commission (2000): Focus on Results: Strengthening evaluation of Commission activities; Communication to the Commission from Mrs. Schreyer in agreement with Mr. Kinnock and the President; SEC(2000)1051 from 26/07/2000.
- [COM 2003/646] Europäische Kommission (2003): Bericht der Kommission; 14. Jahresbericht über die Strukturfonds (2002); Brüssel.
- [COM 2004] Europäische Kommission (2004): Eine neue Partnerschaft für die Kohäsion: Konvergenz Wettbewerbsfähigkeit Kooperation; Dritter Bericht über den wirtschaftlichen und sozialen Zusammenhalt; Arbeitspapier der Bediensteten der Kommission, SEK (2002) 1349 vom 09.12.2002; Luxemburg.
- [COM 2005] European Commission (Hrsg) (2005): Evaluating EU Activities; An introduction; Luxemburg.
- [Crevoisier; Maillat 1991] Crevoisier, Olivier; Maillat, Denis (1991): Milieu, industrial organization and territorial production system: towards a new theory of spatial development; in: Camagni, Roberto (Hrsg.): Innovation networks: spatial perspectives; London, New York; S. 13-33.
- [DeGEval 2002] Deutsche Gesellschaft für Evaluation e.V. (2002): Standards für Evaluation; Köln.
- [Diez 2001] Diez, Maria Angeles (2001): The Evaluation of Regional Innovation and Cluster Policies: Towards a participatory Approach; in: European Planning Studies, Vol. 9, Nr. 7; S. 907-923.
- [Dosi 1988a] Dosi, Giovanni (1988): Sources, Procedures, and Microeconomic effects of Innovation; in: Journal of Economic Literature, Vol. XXVI; S. 1120-1171.
- [Erdmenger/Ziegler 2004] Erdmenger, Katharina; Ziegler, Astrid (2004): Strukturpolitik nach 2006 - Der Vorschlag der Europäischen Kommission; in: WSI Mitteilungen 6/2004; S. 325-331.

- [Freeman 1991] Freeman, Christopher (1991): Networks of innovators: A synthesis of research issues; in: Research Policy, Vol. 20; S. 499-514.
- [Freeman/Soete 1997] Freeman, Christopher; Soete, Luc (1997): The Economics of Industrial Innovation; Third Edition; London.
- [Fritz/Hutschenreiter/Sturn 1997] Fritz, Oliver; Hutschenreiter, Gernot; Sturn, Dorothea (1997): Evaluation of RTD-Programs: Best Practice considerations and the Austrian Experience; in: Newsletter 3; Plattform Forschungs- und Technologieevaluierung, S. 2-11; www.fteval.articles/newsletter/Newsletter_03.pdf, zugegriffen am 11.06.2004.
- [Genosko 1999] Genosko, Joachim (1999): Netzwerke in der Regionalpolitik; Marburg, Schüren.
- [Grupp 1997] Grupp, Hariolf (1997): Messung und Erklärung des Technischen Wandels; Grundzüge einer empirischen Innovationsökonomik; Berlin.
- [Higgins; Tsipouri; van der Lande 1999] Higgins, Tom; Tsipouri, Lena; van der Lande, Robert (1999): Impact of Structural Funds 1994-1999 on research, technology development and innovation (RDTI) in Objective 1 and 6 regions; Dublin, Athens, The Hague.
- [Joint Committee 1999] Joint Committee on Standards for Educational Evaluation/ James R. Sanders (Hrsg.) (1999): Handbuch der Evaluationsstandards; Die Standards des "Joint Committee on Standards for Educational Evaluation"; Opladen.
- [Koschatzky 2001] Koschatzky, Knut (2001): Networks in Innovation Research and Innovation Policy - An Introduction; in: Koschatzky, Knut; Kulicke, Marianne; Zenker, Andrea (Hrsg.): Innovation Networks, Concepts and Challenges in the European Perspective; Heidelberg, New York; S. 3-25.
- [Kuhlmann 1997] Kuhlmann, Stefan (1997): Wie bewertet man Forschungs- und Technologiepolitische Programme; in: Newsletter 3; Plattform Forschungs- und Technologieevaluierung, S. 12-17; www.fteval.articles/newsletter/Newsletter_03.pdf, zugegriffen am 11.06.2004.
- [Kuhlmann/Bührer 2000] Kuhlmann, Stefan; Bührer, Susanne (2000): Erfolgskontrolle und Lernmedium: Evaluation von Forschungs- und Innovationspolitik; in: Vierteljahreshefte zur Wirtschaftsforschung, Vol. 69, Heft 3; S. 379-394.
- [Kuhlmann/Holland 1995] Kuhlmann, Stefan; Holland, Doris (1995): Evaluation von Technologiepolitik in Deutschland; Konzepte, Anwendung, Perspektiven; Heidelberg.
- [Nonaka/Takeuchi 1997] Nonaka, Ikujiro; Takeuchi, Hirotaka (1997): Die Organisation des Wissens; Wie japanische Unternehmen eine brachliegende Ressource nutzbar machen; Frankfurt, New York.
- [OECD 2002] Organisation for Economic Co-operation and Development (2002): Frascati-Manual; Proposed standard practice for surveys on research and experimental development; Paris.
- [Perrin 2000] Perrin, Burt (2000): A Venture Capitalist Approach to Evaluation, or How to Evaluate Innovation; Paper prepared for the 4th EES Conference, October 12-14, Lausanne; www.europeanevaluation.org/docs/53_perrin.pdf, zugegriffen am 15.04.2004.
- [Reid 2001] Reid, Alasdair (2001): What can be learnt from thematic evaluations?; in: Informationen zur Raumentwicklung, Heft 6/7.2001; S. 399-408.
- [Rossi/Freeman/Hofmann 1988] Rossi, Peter H.; Freeman, Howard E.; Hofmann, Gerhard (1988): Programm Evaluation: Einführung in die Methoden angewandter Sozialforschung; Stuttgart.
- [Schubert/Schuh 2004] Schubert, Uwe; Schuh, Bernd (2004): Sustainable Development and Evaluation. An Introduction; in: Hacker, Scott; Johansson, Börje; Karlsson, Charlie (Hrsg.): Emerging market economies and European economic integration; Cheltenham; S. 299-334.
- [Taylor/Bachtler/Polverari] Taylor, Sandra; Bachtler, John; Polverari, Laura (2001): Structural Fund evaluation as a programme management tool: comparative assessment and reflections on Germany; in: Informationen zur Raumentwicklung; Evaluation und Qualitätsmanagement der EU-Strukturpolitik, Heft 6/7.2001; S. 341-358.
- [Toepel 2000] Toepel, Kathleen (2000): Evaluation in der Regionalpolitik; in: Vierteljahreshefte zur Wirtschaftsforschung, Vol. 69, Heft 3; S. 395-405.
- [Widmer 2004] Widmer, Thomas (2004): Qualität der Evaluation - Wenn Wissenschaft zur praktischen Kunst wird; in: Stockmann, Reinhard (Hrsg.): Evaluationsforschung. Grundlagen und ausgewählte Forschungsfelder; 2. Auflage; Opladen; S. 83-110.
- [Wiechmann/Beier 2004] Wiechmann, Thorsten; Beier Markus (2004): Evaluationen in der Regionalentwicklung; in: Raumforschung und Raumordnung, Heft 6; S. 387-396.