



The use of evaluation in the management of EU programmes in Poland



SPRAWNE PAŃSTWO
PROGRAM ERNST&YOUNG

Martin Ferry
Karol Olejniczak



The use of evaluation in the management of EU programmes in Poland



SPRAWNE PAŃSTWO
PROGRAM *ERNST&YOUNG*

Martin Ferry
Karol Olejniczak

Warsaw 2008

Reviewed by:

Prof. Mark Hallerberg
Dr Renata Hayder

Graphic Designer:

Kotbury

Legal Disclaimer

This report has been prepared by academics to provide readers with general information on topics they may find to be of interest.

The information contained in this report shall in no case constitute provision of services.

Whilst the authors have made every effort to ensure that the information presented in this report is accurate, there is a risk of inaccuracies. Ernst & Young Polska spółka z ograniczoną odpowiedzialnością sp. k. „EY” shall not accept any liability and gives no warranty as to the accuracy and completeness of the information contained in this report. The report may contain references to specific statutes or regulations which are subject to amendments and which should therefore be interpreted solely in specific context in which they are quoted. The information is published without regard to any changes and EY gives no warranties, express or implied, and makes no representations as to its completeness, accuracy and validity.

Furthermore, in so far as it is legally entitled to do so, EY disclaims all warranties, express or implied, including without limitation any implied warranties of merchantability and fitness for a particular purpose. EY, its related companies or partners, agents or employees of EY or its related companies, shall not be held liable to the reader or any third parties for any decision or action based on the information contained in this report nor for any indirect, special or consequential losses or damages, even if information on their likely occurrence is obtained.

Copyright by Ernst & Young Polska Sp. z o.o. sp. k., Poland

Ernst & Young Polska Sp. z o.o. sp. k. holds copyright to the report, unless provided otherwise. All rights reserved. No part of this report, including the text or graphics, may be reproduced or transmitted in any form or by any means without the written permission of EY.



SPRAWNE PAŃSTWO
PROGRAM ERNST&YOUNG

Better Government Programme
Ernst & Young Poland
Rondo ONZ 1
00-124 Warsaw
www.bettergovernment.pl

Media partner::



Contents

Acknowledgements	5
Executive summary	6
Why does evaluation use matter?.....	6
How the problem has been addressed? - Our approach	7
What have we found out? - Our conclusions	9
What's next? - Our recommendations	12
1. Introduction - why does evaluation use matter?	14
1.1. Why does evaluation of EU-funded programmes matter?	15
1.2. Why does evaluation of EU-funded programmes matter in Poland?	21
1.3. What determines evaluation use?	23
2. Our approach - evaluation as a process of knowledge production and use	28
2.1. Our research question - focusing on two determining factors .	28
2.2. Our research strategy - taking a systematic approach	30
2.3. Our analytical model	30
3. Evaluation stories from Poland	41
3.1. How were the case studies selected?	41
3.2. What have we found out? Summary of the case studies	44
4. Conclusions	52
4.1. What knowledge has been used?	52
4.2. What determined use of this knowledge?	54
5. Recommendations	61
5.1. How can we overcome systemic weakness?	61
5.2. How can evaluation processes be improved?	62
Annexes	68
Annex 1: Detailed methodology of the research.....	68
Annex 2: Analytical model	72
Endnotes	74
Bibliography	77

List of Figures

Figure 1:	The functions of evaluation.....	17
Figure 2:	The trade-off between functions of evaluation	21
Figure 3:	Evaluation approaches and their focus	33
Figure 4:	Main evaluation approaches.....	34
Figure 5:	Types of knowledge.....	36
Figure 6:	Use and misuse of evaluation	39
Figure 7:	Evaluation functions and their indicators.....	40
Figure 8:	List of selected case studies.....	42
Figure 9:	Summary of the case studies findings	45
Figure 10:	Matrix for analysis of evaluators' roles	58
Figure 11:	Balancing key logical elements of evaluation	63

Acknowledgements

We would like to thank Mr Bartłomiej Osieka and the Ernst & Young team for their support during our research analysis. In particular, we are grateful to Dr Radosław Zubek for his valuable comments on the research approach. We also appreciate the helpful comments provided by referees Professor Mark Hallerberg and Dr Renata Hayder on an earlier draft of this paper. Of course, the content of this report is entirely our responsibility.

Martin Ferry and Karol Olejniczak

“But it’s not just learning things that’s important. It’s learning what to do with what you learn and learning why you learn things at all that matters.”

Norton Juster, “The Phantom Tollbooth”

Executive summary

Why does evaluation use matter?

The paper focuses on the use of evaluations of EU-funded regional development programmes in Poland. It assesses the extent to which evaluations have been utilized as learning tools in public policy management and the factors that determine this. Despite the growing practice of evaluation the issue of real utility of these studies has not been explored in a scientific manner but, to a large extent, taken for granted.

Evaluation involves systematic inquiry into the worth and merit of public intervention, on the basis of its results, impacts and the needs it aims to satisfy. A basic objective is to contribute to social betterment by supporting democratic institutions in choosing, supervising, assessing and improving public programmes and policies. Over the past three decades, evaluation has become an international standard of modern (performance-oriented) public management.

Why does evaluation of EU-funded programmes matter?

The EU has made evaluation a requirement for interventions co-financed from its budget. Over successive programming periods, the EU Commission has introduced more sophisticated, rigorous and comprehensive approaches to evaluation of its policies and programmes. Since 2000, the evaluation of EU Cohesion policy, and the regional development programmes it funds, has been a legal obligation for Member States and for accession countries.

A review of evaluations of these programmes in different Member States reveals their potential importance. At a basic level, the information generated by evaluations of these programmes can help the Commission and Member States monitor their impact. This, in turn, can inform debates at EU and national levels on the merits and future of Cohesion policy. Used effectively, evaluation offers a range of specific benefits: stronger public sector planning; more efficient deployment of resources; improved management and implementation of programmes or other policy interventions; stronger ownership and partnership amongst actors with a stake in programmes; greater understanding of the factors determining the success of programmes; and broader scope to assess the value and costs of interventions.

Why does evaluation of EU-funded programmes matter in Poland?

Particularly in newer EU Member States, EU requirements have stimulated “evaluation booms” in recent years. The capacity of public administration to organize evaluations has matured and there are now dedicated evaluation units in many of these countries. However, the extent to which evaluation has moved beyond a regulatory requirement to become an active instrument for policy learning is uncertain. In many of these countries, evaluation is a new feature of policy environments that are often in a state of flux. Moreover, knowledge of the factors that determine the use of EU evaluations as tools for policy learning is limited. These processes and issues are apparent in the case of Poland. Here, experience of evaluation has been developing in the Polish public sector since the mid-1990s but its effective use is limited and strongly associated with pre-accession and Structural Funds programmes. In the 2007-13 period, Poland will administer some of the largest sectoral Operational Programmes (OPs) in the history of the EU and also implement a new generation of regional development OPs at the regional level for the first time. In this context, it is clear that evaluation is bound to become increasingly significant in Poland over the next few years. The implementation of these programmes will be scrutinized by the Commission and by Polish authorities from the point of view of ensuring their efficient implementation but also to inform debate on future Cohesion Policy programmes. The key question concerns the extent to which evaluations are being used beyond basic reporting requirements, to improve the quality of on-going and future interventions. How can the scope for evaluation to act as a strategic, learning tool be increased?

How the problem has been addressed? - Our approach

In order to address the issue of real evaluation utilization, it is important to outline the factors that can determine the effective use of evaluation. Based on the extensive literature review we have come up with the list of five basic determining factors.

The first factor is the characteristics of the learner or receiver of the evaluation. This refers broadly to the extent to which administrative systems perceive evaluations as formal, bureaucratic exercises or as integral parts of the policy-making process. More specifically, it concerns the resources Contracting Authorities dedicate to the process, the evaluation knowledge and experience of their staff, the stability of the organization and its capacity to disseminate and respond to evaluation results.

The second factor is the characteristics of the policy being evaluated. This concerns the features of a given policy field, the size and importance of the intervention (in terms of resources, geographical scale and time-scale). The results of evaluation of policies that are seen as particularly crucial, are under debate, or are high on the political agenda for some other reason, will obviously tend to receive particular attention.

The third factor is the timing of evaluations. This can be understood in two ways: in relation to the stage of the programme cycle (ex-ante, on-going, ex-post); and the timing of the research in the overall policy cycle.

The fourth factor is the approach taken to evaluation. This relates to the choice between evaluations that are based on the assessment of programme documents and data by relatively detached experts and approaches that strive to involve different stakeholders in the evaluation process in a more interactive way. Of course, apart from interactivity, the scientific scrutiny of the process also matters here.

The final factor is the quality of the report. This concerns the quality of conclusions, the reliability of evaluation results as well as more technical issues - the 'user-friendliness' or accessibility of the report and the clarity and practicality of its recommendations.

Our research hypothesizes that the two key determining factors are:

- the characteristics of the institution that contracts the evaluation research; and,
- the qualities of the evaluation process itself.

These are vital but under-researched issues in the Polish context, where the evaluation capacity of institutions is still developing from a relatively weak base and where knowledge and experience of evaluation processes is still growing.

Research design and plan

In order to verify these hypotheses, the research applied a hypothetico-deductive method of reasoning. The research strategy was qualitative, based on meta-modeling and case studies. In practice, it consisted of three stages. In the first stage a 'meta-model' was constructed, based on an extensive literature review (covering the last 10 years of research on public management, evaluation periodicals and related literature). The meta-model is constructed in a systemic manner - as a process of knowledge production and utilization - with a range of inputs, activities, products and effects (see: Annex 2). The model also presents, in

one logical chain, the interrelations between the factors determining evaluation use, types of evaluation utilization (and even misuse) and the functions of evaluation in the public sector.

The second stage involved analysis of selected case studies, using the meta-model as a roadmap. Cases were selected through a complex procedure, controlling for those factors of evaluation use that were not included in the hypotheses. Six Polish evaluation studies were selected. These were conducted in the field of regional policy between 2000 and 2006. Three were on-going studies, three ex-post. Each of these evaluations was contracted by different public institutions. For each case study, an in-depth analysis was carried out, including report analysis, interviews with Contracting Authorities and evaluators, and extensive documentary review.

In the third and final stage, conclusions were developed, based on comparative analysis. The conclusions referred both to the knowledge that has been used in the evaluations and the factors that determined this use in particular cases. On the basis of this, two sets of recommendations were generated: one set addressing systemic weaknesses of Polish public sector in the use of evaluation; and, a more technical set containing suggestions that could improve the process of evaluation (from the perspectives of evaluators and Contracting Authorities).

At this point, it has to be noted that we are aware of the limitation of a case study approach in terms of supporting theories and developing generalized observations. However, we used this approach to test the synthetic theory (parts of our meta-model) developed on the basis of earlier theories and concepts. In this way, the case study approach is justified. Moreover, by providing detailed analyses on cases, this approach gives practical, in-depth knowledge in an under-researched field and in a Polish context where research into evaluation use has, up to now, been absent.

What have we found out? - Our conclusions

Generally, technical rather than strategic use of evaluation

The research produced a series of conclusions on how the knowledge produced in these evaluations was used and the key determinants of this use. The case studies confirm existing theories on the types of knowledge produced and applied by different types of evaluation. On-going evaluations generated specific technical knowledge that was potentially of use to programmers for operational purposes.

They served to confirm the opinions of programmers rather than generate new knowledge (although they provided programmers with greater awareness of beneficiaries' perspectives, which is important in EU-funded programmes that stress the principle of partnership). The extent to which evaluations generated knowledge that could be immediately utilized in the programmes concerned depended to a great extent on the clarity and practicality of the requirements set by the Contracting Authority in the Terms of Reference. Overall, the generation and application of operational knowledge in these on-going evaluations suggests that Poland's performance in terms of the technical absorption of Structural Funds (i.e. ability to spend funds on time and in line with procedures) is progressing.

However, analysis of the case studies suggests that evaluation in Poland still faces a significant challenge in generating strategic or conceptual knowledge that can inform policy choices and influence other interventions. This contributes to a situation where strategic absorption of EU-funded programmes (i.e. the ability to use funds in a strategic, development-oriented way) is limited. It is difficult to trace the impact of these evaluations beyond the specific confines of the programme concerned. This includes ex-post evaluations that can, in theory, incorporate strategic and conceptual assessments of the impact of programmes that are more broadly applicable. The tension between conducting evaluations that focus on generating knowledge about process issues and evaluations that focus on generating knowledge about the effects of programmes was also clearly demonstrated. Cases that attempted to combine these objectives encountered serious problems.

Hypothesis 1 confirmed - the importance of learners characteristic

Institutions contracting and receiving the evaluation played a vital role. The absence of strong, consistent 'patronage' or support from an evaluation team in the Contracting Authority (for instance, as a result of staff turnover) clearly undermined effective utilization. Conversely, the presence in the Contracting Authority of an individual or a team with good experience and knowledge of evaluation, awareness of the benefits it can provide as a management tool and enthusiasm for the process was important in determining the strategic use of evaluation. In such instances, however, it is important to divide evaluation responsibilities between Contracting Authorities and evaluators clearly. Sustained patronage of evaluation by these authorities can guarantee that studies are used in a more strategic way in the longer term.

Institutional stability in the Contracting Authority can also be important. The institutional flux and frequent staff rotation apparent in Polish

public administration over a number of years has obstructed the accumulation of knowledge and experience (“institutional memory”) and had a negative impact on efficient evaluation use. In these circumstances, it is impossible to expect evaluation to go beyond basic, bureaucratic exercises to serve any strategic use. In the Polish context, accession to the EU played a part in this. A clear line was drawn between pre and post-accession evaluation structures and approaches and the evaluation knowledge accumulated before 2004 was largely lost. It is also worth noting that on-going changes to the system for implementing Structural Funds in Poland can also threaten the accumulation of evaluation expertise.

Hypothesis 2 confirmed - right choice and interactivity matters

Our second determining factor referred to the characteristics of the evaluation process. This includes the issue of research design. Mistakes made at this stage can result in serious tensions and problems at later stages in the research. The case studies clearly showed that there is a trade-off between focusing an evaluation on processes (partnership building, management improvement) and focusing on effects (accountability, cognitive function). This trade-off is matched by dichotomies in evaluation approaches (participative approaches rooted in constructionist philosophy vs. causal designs grounded in positivist paradigm) and the roles played by participants (evaluators as involved advisors and stakeholders as partners vs. evaluators as controllers - mentors and participants as assessed learners). Efforts to combine them in one research exercise can result in serious problems that can hamper the ultimate utility of the research.

Furthermore analysis of our case studies suggests that interactive evaluation processes are crucial to their efficient use. Interactivity is important for learning and improvement. People and institutions learn faster when they are involved in the process. They also internalize knowledge faster when they are involved in creating that knowledge.

Unexpected findings - quality of data and role of evaluators

The unexpected factor that played an important but negative role for all reviewed ex post studies was the quality of the data available for evaluation. Evaluations were hampered by weaknesses in their programmes’ information base (we could call it a short “programme memory” or, in fact, a lack of it). In the absence of consistent, quality statistical information on the progress of projects, it is impossible to assess the impact of a programme, draw robust, overall conclusions and produce new conceptual insights. Put simply, in these circumstances there is no reference point to measure impact against.

On the other hand, the important positive factor that occurred in the analysis was a proper role of evaluators. All of the conducted case studies emphasizes the key role an evaluator plays in determining the utilization of the evaluation. Evaluators can perform a range of functions depending on the orientation of evaluation: as a relatively detached consultant; as a scientific expert; as an animator or facilitator of stakeholder participation; and as a mentor or educator, encouraging a process of learning amongst stakeholders. There is no optimal model for the evaluator's role. Several hybrids can exist within this broad categorisation. However, if we accept that a current priority in the Polish context is building evaluation capacity and an evaluation culture, then it seems logical that evaluators should follow the role of animators (where studies focus on processes) and educators (where studies focus on the assessment of effects or impacts). Educating and facilitating learning as part of the research process (by making the process interactive) is a responsibility of an evaluator.

What's next? - Our recommendations

Anchoring the evaluation knowledge and experience in institutions

An important recommendation here is to strengthen the evidence base in order to increase the credibility, utility and status of evaluations in the public sphere. In order to strengthen the initial data baseline, ex-ante evaluations and on-going studies should be used to create a reference point for future strategic ex-post analysis. These evaluations should combine analysis of the programming document with empirical studies designed with the information needs of the future ex-post in mind.

Institutional flux and frequent staff rotation is inevitable in administrative systems under construction or in transition. The challenge is to anchor at least some of the knowledge and expertise gained in evaluation processes even when the people directly involved move on. One solution would be to encourage broad dissemination of evaluation experience within institutions so that knowledge is not limited to selected actors. Moreover, establishing public databases of all evaluation studies (including full evaluation reports, methodologies and basic data on the evaluation context) would allow a wider circle of people to learn about specific research constraints and characteristics of the process. Also, a clear division between operational and systemic conclusions and recommendations in evaluations would provide a clear point of reference and at least partial memory of the past programme, its problems, the solutions that have been found, etc.

Thus, the systemic use of evaluation requires a change in thinking. Policy-makers and civil servants have to think in a longer-term perspective and in a more systemic way - beyond their tenure. The processes they initiate will only become apparent later on in the policy cycle, when other institutions and actors may be responsible.

Improving evaluation processes

In order to ensure strong evaluation processes and maximum scope for their effective utilization, it is important that Contracting Authorities (CA) design robust Terms of Reference. It is when Contracting Authorities are designing ToR that important choices must be made, balancing or even choosing between different components of the evaluation process. This includes: the choice between focus on process issues or effects; the degree of coherence between research methodology and expected functions of the evaluation; and, the extent to which the design is realistic, given the level of resources and the availability of data. Beyond this, the utility of evaluations will always be boosted where there is a strong, supportive relationship between the CA and the evaluator throughout all the stages of the evaluation.

Turning to the role of evaluators, it is clear that they have an important role in developing an appropriate research methodology, within the framework set by the CA. This involves a similar process of balancing different considerations (evaluation objectives and scope, resource issues, data constraints) faced by the CA but at a more detailed level. Evaluators must also choose the most appropriate role to play in the research process. It is arguable that one of the key factors for successful evaluation utilization is high interactivity between evaluators and stakeholders. Within this, evaluations of process issues within programmes (that are often organized on an on-going basis) benefit when evaluators animate and facilitate broad participation and discussion amongst stakeholders. For evaluations of programmes' impact (usually, but not exclusively organized as ex-post exercises), evaluators can play the role of educator or mentor. This encourages an interactive approach with stakeholders but ensures that the evaluator maintains the distance to provide objective reflections and insights.

As can be seen from the above, developing an evaluation culture in public administration is a challenging process, requiring long-term planning and sustained commitment from policy-makers. However, the potential benefits in boosting the performance of public interventions - both EU and domestic - are significant.

1. Introduction - why does evaluation use matter?

The paper focuses on the use of evaluations of EU-funded regional development programmes, with particular reference to the case of Poland. Evaluation is a standard practice of performance-oriented, public administrations. Since 2000, evaluation has been a legal requirement for interventions co-financed from the EU budget in Member States and Candidate Countries. Particularly in newer EU Member States, these requirements have stimulated “evaluation booms”. However, the extent to which evaluation has moved beyond a regulatory requirement to become an active instrument for policy learning is uncertain. In many of these countries, evaluation is a new feature of policy environments that are often in a state of flux. Moreover, knowledge of the factors that determine the use of EU evaluations as tools for policy learning is limited. This paper explores the factors that can determine why some evaluations are used more effectively in public sector management than others. It focuses on the case of EU-funded regional development programmes in Poland. Here, driven by EU programming requirements, evaluation is a young but increasingly important component of the policy process. The paper assesses the extent to which evaluations have been utilized as learning tools and the factors that have determined this.

The report is structured into five sections. Based on an extensive review of evaluation literature, the first section establishes the research context. Drawing on examples from different EU Member States, it highlights the potential benefits of evaluation to public management, particularly in the context of EU-funded programmes. It also emphasizes the particular benefits offered by evaluation in the case of Poland, while noting the challenges evaluation faces in this context. Finally, from a review of evaluation theory, it introduces the basic factors which can determine the efficiency of evaluation use in public management. The second section briefly outlines the methodology adopted for the assessment of these determining factors in the Polish context. It focuses on two factors that are deemed to be particularly significant in the Polish case: the characteristics of the institution that contracts the evaluation research; and, the qualities of the evaluation process itself. Section 3 presents case study analysis of six evaluations of EU-funded programmes carried out in Poland in the period 2000-06, with particular focus on how the characteristics of the Contracting Authority for the evaluation and the features of the process itself determined the effective use of the evaluation. Section 4 synthesizes the results of the case study research and draws a number of overarching conclusions.

The final section makes some recommendations for public authorities and evaluators on how the use of evaluations can be supported in Poland.

1.1. Why does evaluation of EU-funded programmes matter?

Evaluation is defined as systematic inquiry for judgment of the worth and merit of public intervention, on the basis of its results, impacts and needs it aims to satisfy (European Commission, 2004 a; Mathison, 2005; Rossi et al., 2004).¹ Evaluation has been widely recognized and accepted as a learning tool for improving public policies and programmes.² Over the past 25 years, evaluation studies have become a standard, fundamental feature of public management models internationally. (Furubo et al., 2002; Heinrich, 2003; OECD, 1999). In basic terms, the role of evaluation is to contribute to social betterment by supporting democratic institutions in choosing, supervising, valuing and improving public programmes and policies (Mark et al., 2000:3). Some authors link evaluation with the wider context of requirements of modern democratic systems (transparency of public action, informing public opinion, effectiveness of public institutions and learning (Henry, 2001). On a more technical level, evaluation aims at “improving the quality of the present as well as future interventions” (Chelimsky, 1985:7; Rist, 1990:7). This broad objective is usually translated into a list of functions (also named “objectives of evaluation” - compare: Chelimsky (1997), Mark et al. (2000:13)).

The particular focus of this research is EU-funded regional development programmes. The Cohesion policy of the European Union (EU) supports the Structural Funds and Cohesion Fund, which are the main EU-funded regional policy instruments in Member States.³ However, countries preparing for membership also have access to EU regional funds (known as pre-accession funds), designed to mirror the Structural Funds system and prepare newcomers for the management of Cohesion Policy after accession. Therefore, in our analysis we accept the extended interpretation of Cohesion (both Structural Funds and pre-accession funds). Cohesion policy has become one of the most evaluated policies in Europe (Bachtler, Wren, 2006). Moreover, the European Commission’s requirements for the evaluation of development programmes supported under its Cohesion policy are seen as an important driver of broader evaluation practice. Systematic requirements for evaluation were first introduced under the 1988 reforms of the Structural Funds (Council Of The European Economic Community, 1988a:art.6; 1988b:art.25-26). This was driven by several factors, including the doubling of funding, the shift from

The definition and role of evaluation

Evaluation in EU-funded programmes

a project-based approach to multi-annual programming, the need to verify the value added by EU-funded programmes and ongoing concerns about the effectiveness of interventions (European Commission, 1989). Over successive programming periods, the Commission has introduced more sophisticated approaches to evaluation of policy and programmes. Evaluation requirements have become more rigorous and comprehensive with a clearer articulation of the respective responsibilities of the Commission and Member States in ex-ante, mid-term or interim and ex-post evaluation of all programmes.

The use of EU evaluations at different levels

At the European level, the Commission uses the information generated by evaluations to monitor the efficiency and effectiveness of programming and to find out whether the promised targets are being achieved. These results feed into the Commission's broader plans for its future funding programmes. At the national level, evaluation results can also generate information for national policy debates and positions on the future of the Funds. A basic aim can be to compile an overview of the achievements of programmes so that their added value can be assessed (Raines, Taylor, 2002). At the programme level, the basic functions of evaluation, noted above, are applicable to evaluations of EU-funded programmes. Evaluations can inform and enable the process of adjusting the programmes, including the channeling of funds between priorities and measures, based on a changing economic context and progress to date in terms of implementation. Beyond meeting the Commission's regulatory requirements, they can ensure effective and efficient programme management and address any identified weaknesses in structures, processes and procedures.

Potential functions of evaluations in EU-funded programmes

Analyses of evaluations of EU-funded programmes have highlighted aspects that are particularly significant. Notably, Structural Fund evaluation has been widely interpreted as a learning opportunity rather than simply a reporting exercise. Part of the rationale for this is that Structural Fund programmes are designed and delivered on a partnership basis, relying on the contribution and expertise of a wide range of vertical and horizontal partnerships. Increasing emphasis is being placed on knowledge production - the evaluation should test and verify the theories and assumptions behind the policies or programmes Bachtler (2001:46-48). However, in practice, Batterbury (2006:181-182) concludes evaluations currently fulfill only three functions for Structural Funds programmes: accountability & legitimacy, improving planning, improving quality and performance. Other functions, such as building capacity, learning, empowerment, developing a sense of ownership between participants in the programmes, although important, are much less apparent. Nevertheless, a review of evaluation literature generates five potential functions of evaluation in reference

to EU-funded programmes. These are presented in Figure 1, below. The sections below explore these functions with reference to cases drawn from evaluations of EU Structural Funds programmes.

Figure 1. The functions of evaluation.

Function	Description
Improving planning	<ul style="list-style-type: none"> • ensuring rationality & justification of the interventions by verifying its internal coherence, relevance to the needs • ensuring efficient deployment of the resources by verifying the logic of the intervention, rationality of the objectives, external coherence of the strategy
Improving performance & quality	<ul style="list-style-type: none"> • improving programme management - its structure and procedures • improving quality of delivery by verifying to what extent the intervention fulfills the planned objectives, what is the quality of the products
Improving partnership & ownership	<ul style="list-style-type: none"> • strengthening relations, interactions, networks between participants of the particular programme • activating stakeholders & involving them in discussion and operations • empowering marginalized or excluded stakeholders
Improving cognition	<ul style="list-style-type: none"> • explaining factors and processes that determines intervention success or failure (including side effects) • verifying theories and assumptions of policies, programmes concerned • drawing conclusions for others and promoting good practices • building evaluation capacity
Improving accountability	<ul style="list-style-type: none"> • demonstrating to the public to what extent objectives were fulfilled • demonstrating how well public resources were used

Source: (Bachtler, 2001; Batterbury, 2006; Chelimsky, 1997; Tavistock Institute et al., 2003).

First, evaluation can improve planning. Evaluation can provide an opportunity for policy-makers to check the rationality and justification of the interventions by verifying their internal coherence and relevance to existing needs. Evaluations of programmes can be particularly important in assessing whether resources are being efficiently deployed and whether adjustment of initial programme plans is necessary. "Planning" also includes a priori verification of structures and procedures designed for implementation. Monitoring assessment (what data should be gathered and how should it be collected) constitutes a core of this assessment. The planning function can apply

Evaluation and improved
planning

to the interim evaluation of programmes that are still underway or the use of evaluations to inform planning for future programmes. For instance, a mid-term review of the Objective 2 Programme in North Rhine Westphalia 1994-99 presented a number of results, including proposals for the orientation and prioritization of measures, the need for more systematic and comprehensive monitoring, and the institutional setting in which the responsible managing authorities were operating. The results of this were evident in the reprogramming of the Objective 2 programme in 1997-1999. An update of the mid-term evaluation and various thematic case studies were commissioned. On the basis of this, evaluators put forward proposals for the programming period 2000-2006. Following debate amongst economic and social partners, local authorities, and implementing bodies, several of these proposals became part of the NRW Objective 2 programme for 2000-06. Major changes in this programme included: a re-orientation from infrastructure to business and innovation-related measures; the implementation of a systematic, computer-assisted project selection and monitoring scheme; and, the establishment of a technical secretariat to assist the managing authority in areas such as project selection, monitoring, reporting, and information. These were all strongly influenced by the preceding evaluation activities (Jakoby, 2006).

Evaluation and improved
performance

Second, evaluations can be used to improve the performance and quality of programmes. This can refer to assessments of programme management structures and procedures or to the quality of products. Again, such uses can apply to the evaluated programme itself or to plans for future programmes. For instance, in Finland an updated mid-term evaluation (MTE) exercise covering all of its 2000-06 OPs was carried out in 2005. The use of these evaluations to improve the performance and quality of the 2000-6 OPs was limited: the results of the original MTEs were still valid and the OPs were moving towards completion anyway. However, the updated MTEs made several recommendations that were more long-term. Hence, they became a useful point of reference for improving programme management structures and procedures for the 2007-13 period. For instance, the updated evaluations recommended the simplification of the management system and the integration of administrative procedures across different government departments. As a result, for the 2007-13 period the administrative structure for OP management has been rationalised. This has involved the integration of certifying tasks under one authority, the creation of a joint monitoring system for all Structural Fund activity, and the unification of EU funds and national co-finance for OPs in the state budgets of the central ministries which act as Managing Authorities.

Third, evaluation can be used to improve partnership and ownership amongst those participating in a particular programme. This can relate to consolidating links or networks for cooperation between participants. In Sweden, for instance, national authorities undertook overarching reviews of all of the regional mid-term evaluations for the 2000-06 period. The aim was to draw comparative lessons and disseminate these to programme managers and more widely in the policymaking community. The national agency NUTEK prepared a brochure containing an executive summary highlighting the key conclusions across all regional evaluation reports. On the basis of this, a national conference on Structural Funds experiences and regional development was organized. This was followed by regional seminars and conferences, which provided a platform for discussions between regions and the national level on the main conclusions of the evaluations. Evaluations can also empower participants and give them a greater sense of involvement and ownership in the programme. This applies particularly to those organizations that have previously played only a peripheral or marginalized role in the programme. For instance, one of the objectives of the interim evaluation of the Rural Development Plan 2000-06 in the Basque Country was to include as broad as possible participation from stakeholders. Not only was this deemed to have increased utilization of evaluation results, it also encouraged more transparent, less hierarchical links between all of the groups involved in the programme (Izquierdo, 2006).

Evaluation and improved
partnership

Fourth, evaluations can improve policy-makers understanding of different programme-related issues. As noted above, evaluations can help to explain factors and processes (such as strategic relevance and programme implementation) that determine the success or failure of an intervention. In addition, evaluations can test the theories and assumptions on which programmes are based and draw lessons for other programmes through the promotion of good practice. For instance, Scotland's Objective 2 programmes 2000-06 are notable for the emphasis placed on sustainable development as an integrating or 'horizontal' theme. With the prospect of the mid-term evaluations of the 2000-6 period, it was decided to incorporate benchmarking and international comparisons (with Nordrhein-Westfalen) into evaluations of sustainable development in the Scottish programmes. The aim was to achieve a deeper understanding of issues in both countries through collaboration. This involved the organisation of workshops as part of a mutual evaluation exercise for evaluators and steering groups in both programmes. It also included the development of a common methodology to enable comparison between the different programmes, drawing out the strengths and weaknesses of the approaches taken to the horizontal themes in each OP. Improved understanding can also

Evaluation and improved
understanding

refer to appreciation of the role of evaluation itself. In some cases, particularly in newer EU Member States, evaluation requirements for Structural Funds programmes have helped to strengthen overall administrative capacity for conducting evaluations. For example, in Austria, evaluation capacity building activities became more established upon the country joined the EU in 1995. At this time, the Austrian Federal Chancellery (BKA) established the 'Checkpoint EVA' platform (run by the Austrian Spatial Planning Institute, ÖIR) which had the objective to facilitate exchange of experience and learning in regional policy evaluation. The network included evaluators and national as well as regional policy makers (Holzinger, 2001:29). The platform concept was developed further for 2000-06 as part of the MTE, when the coordination and work platform KAP-EVA was established by the Austrian Conference on Spatial Planning (ÖROK). KAP-EVA represents one of the most active examples of institutionalisation of capacity building in an international context (Raines, Taylor, 2002; Strohmeier, Holzinger, 2006). During the period 2002-06, it organized a series of workshops dealing with a variety of evaluation-related issues and involving a range of national and Land-level actors, as well as evaluators from different research institutes. These workshops have been perceived as useful tools for stimulating discussions, setting common standards and adapting EU requirements to domestic needs.

Evaluation and improved
accountability

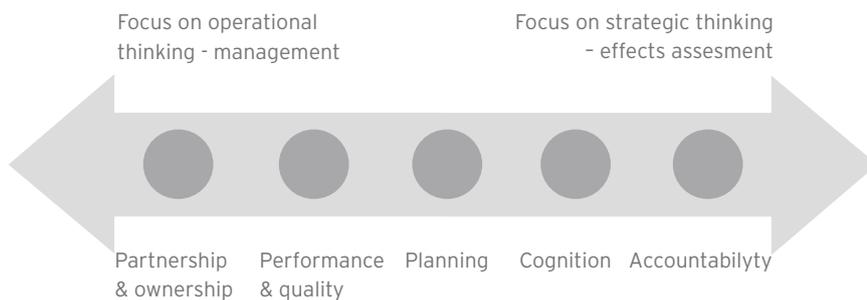
Fifth, evaluations can improve accountability by demonstrating to the public the extent to which the objectives of the programme were fulfilled. It is important to note that there can be a tension in stressing the use of evaluations to provide accountability or transparency. An inherent danger is that the prominence of accountability functions - demonstrating what has been achieved - can lead to the evaluation being seen as simply a report for the benefit of higher authorities, notably the European Commission. It can constrain open debate for fear of exposure before these authorities, thereby diminishing the potential for alternative uses of the evaluation for learning at the programme level. One way to minimise this tension is to generate different evaluation outputs for different audiences and purposes. For instance, evaluations of the Graz Urban programme 2000-06 (in Steiermark) - drew a clear distinction between its learning and reporting functions. It was decided that open discussion could take place within an evaluation Steering Group - which in itself is a form of output - but that the only outcomes which would be reported formally to the Commission would be those which were of key importance to satisfying Commission requirements.

Trade-offs and tensions in
evaluation functions

In practice, one evaluation study can fulfill more than one function. However, the empirical research shows that there is a clear trade-off between some of the functions (Eser, Nussmueller, 2006:256;

Olejniczak, 2007:223; Stern, 2005). Assessing the effects of the programmes is always related to the external audience (citizens) and external control (assessing programme operators effectiveness from the point of view of society as a whole), while improvement of day-to-day management is directed inwards (towards stakeholders) and relies heavily on open, self-critical discussion between programme operators (with the help of the evaluator). Concerning evaluation functions, two basic foci can be identified: focus on process or management on the one hand and focus on impacts or effects on the other. This dichotomy fits the recent distinction made by the EU between operational and strategic characteristics of evaluations in Cohesion Policy (Council of The European Union, 2006:art.47.2). It also corresponds with Hausner's (2007) distinction between technical absorption (ability to spend funds on time and in line with procedures) and structural absorption (ability to use funds in a strategic, development-oriented way). These tensions are summarized in Figure 2, below.

Figure 2. The trade-off between functions of evaluation.



Source: (Olejniczak, 2008)

1.2. Why does evaluation of EU-funded programmes matter in Poland?

As noted above, evaluation has been a commonly recognized standard and a useful learning tool in modern public administrations. If Polish public administration aspires to modern performance-oriented and effective public sector, it has to accept and integrate evaluation as a fundamental feature. So far, the impulse for changing approaches to evaluation in Poland has been external. Commission requirements for evaluating EU-funded programmes have stimulated an "evaluation boom" in recent years and evaluation is a rapidly growing field in the country. The capacity of public administration to organize evaluations has matured: there are now dedicated evaluation units and a general widening of analytical approaches to evaluation is also underway. Initially, evaluations were dominated by documentary analysis but

The recent evaluation
'boom' in Poland

more sophisticated approaches are emerging (e.g. different types of interview, socio-economic simulation processes etc.). There has also been positive interaction between international and Polish groups involved in the evaluation process, through twinning and through cooperation between Polish administrators and international consultants. More generally, a positive model of participatory evaluation has gradually been taking shape within managing authorities, creating the scope for increased interaction between administrators, experts and various partners.

Weaknesses of
evaluation culture

On the other hand, it is important to note that these processes are starting from a traditionally weak base. Experience of evaluation has been developing in the Polish public sector since the mid-1990s but its effective use (or the emergence of what could be termed an "evaluation culture") is limited and strongly associated with pre-accession and Structural Funds programmes (Krajowa Jednostka Oceny, 2005; Olejniczak, 2002). Furthermore, there is limited domestic academic research and literature on evaluation. There is a tendency to conflate evaluation theory and practice with Commission regulations and goals. Additionally, although approaches are evolving, there can still be a need to ensure that the function of evaluation is properly understood and that appropriate human and institutional resources are dedicated to it in a public administration that is often in a state of flux. There can still be perceptions of evaluation as a form of audit and there can still be a tendency to produce large, unwieldy evaluation outputs that are unlikely to be efficiently utilised. Though expansion is underway, there are still only a limited number of firms able to conduct evaluations. The quality of evaluations can be uneven as new consultancy companies develop evaluation experience. Ensuring sufficient institutional capacity and human resources dedicated to evaluation is an on-going challenge. Beyond this, a fundamental weakness for the majority of Polish evaluations has been the limited measurement of different socio-economic trends, the absence of quality statistical data (particularly over longer time periods) and, thus, difficulties gauging the actual impact of programmes (Szlachta, 2006). Thus, although evaluation is an increasingly important part of the policy environment, the extent to which evaluation has moved beyond a regulatory requirement to become an active instrument for policy learning is uncertain.

Polish evaluation
culture in the future

In the 2007-13 period, Poland is administering some of the largest sectoral OPs in the history of the EU and Polish regions are managing their own regional programmes for the first time. In this context, it is clear that evaluation is bound to become increasingly significant in Poland over the next few years. The implementation and impact of these programmes will be scrutinized by the Commission as it

considers Cohesion policy reforms for the post-2013 period. These programmes will also be scrutinized closely in Poland, from the point of view of ensuring their efficient implementation but also as a means of supporting national strategic development and strengthening Poland's role in Cohesion Policy in the future. For all of these reasons, Poland is investing considerable time and financial resources in evaluation. The key question concerns the real impact of this process. In other words, is the opportunity to learn being taken? Is the opportunity being taken to draw conclusions and collect evidence on the positive effects of EU support for Poland's strategic development? Alternatively, is the goal merely to meet Commission regulations and procedures? Up to now, there has been no research dedicated to evaluation utilization in Poland.⁴ There were no scientific, empirical considerations on such issues as:

- What specific factors decide how efficiently an evaluation is used?
- How do we increase the scope for evaluation to act as a learning tool in this administrative environment?

The following research is the first attempt to address these crucial questions.

1.3. What determines evaluation use?

Having outlined the functions and potential benefits of evaluations of EU Cohesion policy programmes, particularly in the Polish context, it is important to outline the factors that can determine their effective use.

1.3.1. What are the determining factors?

Despite the extensive practice of contracting evaluation studies, information on how the knowledge produced by evaluations of Cohesion Policy is utilized remains limited. Surprisingly, in Europe there is a significant gap between the number and scale of evaluations and research-based theories that investigate the knowledge they produce, its use in public management and decision making, and the factors that determine effective utilization. Existing academic literature on evaluation utilization, as well as empirical research, has been largely developed in the United States. These studies usually focus on the instrumental, immediate use of evaluation (i.e. the implementation of recommendations) and present little analysis of the conceptual use and longer term impact of evaluation results beyond the life cycle of the evaluated programme. Moreover, they often adopt the evaluator's perspective, based on their assessment and perception of how evaluations were utilized. The institutional context of evaluation use is rarely analyzed (Leviton, 2003; Shulha, Cousins, 1997). Research in

Limited research on
evaluation use

the context of European Union programmes has a similar perspective. Studies are usually limited to the short-term, instrumental use of evaluation by a narrow group of actors. So far, there have been two empirical studies of this kind. Both investigated the use of evaluation within the European Commission (EC). As they were contracted by the EC, they omitted scientific consideration and focused instead on effectiveness analysis (compare: EPEC, 2005; Williams et al., 2002).

Five factors that determine evaluation use

Nevertheless a review of evaluation research and literature produces a basic list of five general factors that can play a part in determining evaluation use:

- the learners'/receivers' capacity
- characteristics of the policy being evaluated
- research timing
- evaluation approach taken, and
- quality of the report.

We discuss these briefly below.

The importance of institutional capacity

The capacity of the learner or receiver of the evaluation is a crucial factor. The extent to which Member States can generate and utilize knowledge as a result of evaluation depends on the resources it can dedicate to the process. This can concern the quality of human resources of the institution (that is knowledge on evaluation and experience of staff involved in contracting and supervising the research). Put simply, once staff have experience in evaluation they better understand the specific benefits of evaluation and they are aware of how it can be used (Boyle, Lemaire, 1998). The stability of an institution can also be important. Once an institution is stable and well-established, with good experience in dealing with interventions similar to the evaluated case (i.e. once it has an 'institutional memory') then it accumulates knowledge from evaluations and refers to this in new situations. Finally, the position of the institution in the policy system can be important (its ability to allocate or channel resources, and its power to promote results and make changes can have an obvious impact on the extent to which an evaluation is utilized).

Variable role of the wider administrative culture

The learners characteristics also include the issue of wider, national culture of public management. The degree to which authorities contracting evaluation have the necessary knowledge and skills, and the extent to which institutional frameworks are in place to manage evaluation and promote evaluation practices varies across EU Member States (and even across regions within EU Member States). This is of course related to the place of evaluation in a broader public

administration culture. Toulemonde divides countries into two groups: those where evaluation is bureaucratic (2000) exercise and those where it is a part of democratic functioning. More specifically for European regional policy, research has identified a 'north-south' divide (Bachtler, Wren, 2006; Tavistock Institute et al., 2003). In countries such as the Netherlands, UK, Germany, France, the Nordic countries and, recently, Ireland the evaluation of regional policy is well established. Regular evaluations of national regional policy interventions take place and the knowledge produced by evaluation is an integral part of the policy-making process. By contrast, in Italy, Spain, Portugal and Greece policy evaluation has not been as embedded within the public administrative culture. Evaluations have been seen primarily as a periodic reporting exercise and have tended to be used intermittently and in an ad hoc way (Casavola, Tagle, 2003). A further group of countries consists of newer Member States, such as Poland, where the evaluation of regional policies is very limited and driven largely by Structural Funds requirements. Managing and implementing EU-funded programmes is a complex field, involving a range of actors and organizations at European, national and sub-national levels, and including participants from public and private spheres (EU support has to be co-financed from national public or private funding sources). Evaluations, therefore, must strive to meet the aims of different organizations, some of which may have different or competing interests and perspectives on how the evaluation should proceed and what its outputs should be. There is wide variation in the administrative arrangements for management and implementation of the programmes, including arrangements for their evaluation. Broadly, Member States' systems vary according to the extent to which programme administration is centralized or decentralized approaches and according to the extent to which systems for administering the programmes are subsumed within or differentiated from institutional arrangements for domestic development interventions. All of these issues must be taken into account when considering the functions and use of evaluations.

Characteristics of the policy being evaluated (i.e. the evaluand) can have a direct impact on how an evaluation is used. The features of a given policy field, the size and importance of the intervention (in terms of resources, geographical scope and time-scale) can all be influential. Evaluations of interventions covering fields that are high on the political agenda are clearly of greater potential interest to managers and policy-makers as well as the wider public. Similarly, programmes with a high level of investment (or perceived as critical in particular field) will draw particular attention and as a result, there is greater likelihood that evaluations will be used. The focus of our research, regional development programmes funded by EU

Regional policy
and cohesion policy

Cohesion policy, are currently regarded as important interventions, particularly in newer Member States such as Poland. However, these programmes are challenging to evaluate as they can involve a wide range of interventions, including physical infrastructure, economic infrastructure, business development, human resources, research, technological development and innovation, environmental improvement, tourism and community development.

Timing of the research

The timing of evaluation research can be influential in its use. This can be understood in two ways. First, the timing of the research in relation to the stage of the programme cycle can be crucial. Evaluations of EU-funded programmes are organized at different stages of the programme cycle: before the programme operation (ex-ante), during programme operation (on-going) or after its closure (ex-post). This can determine the scope of the evaluation, the functions it addresses and its utilization. It is widely recognized that ex-ante evaluations focus on planning functions. This concerns strategic issues such as the a priori assessment of a plan's rationality. It also relates to operational considerations, notably the design of the implementation system. Ongoing evaluation addresses operational issues, including programme management, partnership-working and the adoption of 'good practice'.. In relatively advanced programmes, on-going valuations can also include an initial assessment of some strategic issues such as effects and impacts and accountability. Ex-post research focuses on strategic topics - accountability and learning theories (European Commission, 1999:vol.1; 2004a:11-13; 2004b:46-50). Second, the timing of the research in the overall policy cycle can be important. If a policy field, such as regional policy, is in transition between old and new approaches or undergoing radical reforms, or is high on the political agenda for some other reason, related evaluation studies may receive particular attention.

Research approach

The approach taken to the evaluation process and the quality of the process is another determinant of evaluation use. Here it can be useful to distinguish between two general modes: 'expert-oriented' evaluation and 'interactive' evaluation. In the former case, evaluations are based on expert analysis of programming documents and statistical data. Partners provide information to the evaluators but are passive in the process. The evaluator interprets evidence and draws conclusions, issuing a report about findings. Here the scientific scrutiny of this analysis (mainly its validity and reliability) is crucial.⁵ It is up to the programme to respond to these conclusions. In the latter case, a more interactive approach is taken. Programme partners are encouraged by the evaluator to participate in discussions and reflect on the programme. Their views are important in developing the findings,

conclusions and recommendations of the evaluation report. Under this view of evaluation, the processes of interaction, reflection and learning are seen as equal in importance to the final evaluation report, and should facilitate self-discovery and self-diagnosis among programme partners (Owen, Rogers, 1999:65; Patton, 1997). Current Cohesion policy evaluation methodologies range from the 'bottom-up', survey-based assessments of project and beneficiary outcomes to the 'top-down', input-output models of aggregate programme impact and many combine aspect of both approaches. However, a basic contention of our research is that if the process of evaluation is more interactive then evaluation use will be more effective. That is because involvement of the stakeholders in the evaluation work helps to build trust during the process, creates ownership of the results, encourages a consensus and coalition for change (i.e. "buy-in" of the research) and ensures that the research fits information needs (King, 1998).

Finally, the quality of the evaluation report can have a bearing on its utilization. This refers to the quality of conclusions, and the extent to which they are supported by reliable evidence and sound analysis. Furthermore the quality of the report's recommendations can also have an impact on its use. The proposals contained in the report must be relevant to real needs and realistic or operational.⁶ Overall, the report has to be structured and presented in an accessible and user-friendly way if its conclusions and recommendations are going to be utilized effectively (Morris et al., 1987). For Cohesion policy, the Commission lays out quality criteria for evaluation reports. However, as with other aspects of the evaluation process, interpretations of this framework vary across Member States.

The quality and accessibility
of evaluation reports

2. Our approach - evaluation as a process of knowledge production and use

The previous section's assessment of public sector practice and evidence from some empirical studies suggests that evaluation use is a complex process, with a list of determining factors interacting in a range of ways to produce variable results. Our research addresses this complexity in two ways. First, it focuses on two determining factors from the list provided at the end of the previous section (namely, the characteristics of the receiver of the evaluation and the characteristics of the evaluation process itself). Second, the research approaches the issue of causality and the determinants of effective evaluation use in a systematic and integrated way. For this purpose, it follows a meta-model that integrates factors of use, evaluation stages and types of evaluation use into one process map. The following section explains our approach in more detail.

2.1. Our research question - focusing on two determining factors

As noted above, despite an "evaluation boom" (Pollitt, 1998) and extensive contracting of evaluation studies, knowledge on the mechanisms of learning from evaluation is limited. This prompts a basic empirical question: "Why are some evaluations executed for EU funded programmes used more effectively than others in public sector management?" The previous section listed several factors that can play a role in determining how evaluations are used: the timing of the research; the condition of the intervention being evaluated; the quality of the evaluation report; the learners'/receivers' capacity; and, the quality of the evaluation process. All of these variables can interact in complex ways and in order to address this, our research focuses on the last two. The first three of these variables are relatively clear-cut and their importance to the impact of evaluation can be demonstrated in a straightforward way. However, the last two are potentially vital to the effective use of evaluation. At the same time, they are more challenging to assess than other variables and, as yet, they have not been explored in depth in the context of Structural Funds. Therefore, our research focuses on these two variables: the characteristics of the receivers and the quality of the evaluation process. In our opinion, understanding the dynamics of these two variables is an under-researched but crucial area in exploring and improving the use of evaluation. This is particularly the case in the Polish context, where the evaluation capacity of institutions

is still developing from a relatively weak base and where knowledge and experience of evaluation processes is still growing.

This approach generates two hypotheses, with related questions or indicators:

The research hypotheses

Hypothesis 1: effective use of evaluation is determined by the learners'/ receivers' characteristics (quality of human resources, institutional memory, the position or status of the institution).

- What evaluation experience does the learner/receiver have? If the human resources of the institution includes specific experience in evaluation then it is more likely that evaluations will be utilized in a more effective way. That is because there is an understanding of the specificity of evaluation research (e.g. the difference between evaluation and audit, positive power of constructive critique), and awareness of its utility but also of its limitations.
- How well-established or stable is the institution? If the institution is stable and well-established then evaluation use (especially conceptual and long-term use) will be more effective. This is because an institution with strong continuity and institutional memory accumulates knowledge from evaluations and refers to this in new situations.
- Finally, what position does the institution have in the broader public administrative system? if the institution has a strong position in the inter-institutional system, in terms of authority, financial and human resources, then the utilization of evaluation will be more effective. This is because the institution has the power to monitor and put pressure on other actors in implementing recommendations. Furthermore, it has the resources and authority to disseminate and promote the results.

Hypothesis 2: effective use of evaluation is determined by the level of interactions during the evaluation process, the level of stakeholders' involvement.

- To what extent is the evaluation conducted in a 'top down' or 'bottom up' way and what was the balance between quantitative and qualitative methods? If the process of evaluation is interactive then evaluation use will be more effective. That is because involvement of the stakeholders in the evaluation work helps to build trust during the process, creates ownership of the results, encourages a consensus and coalition for change (also called "buy-in" of the research) and ensures that the research fits information needs.

A qualitative approach

2.2. Our research strategy - taking a systematic approach

The research explores evaluation use in a systemic way - as a process of knowledge production and utilization. It follows an extended CIPP model (Context, Input, Process, Products), commonly used in the implementation of most public interventions.⁷ Based on a review of evaluation literature, a model was generated to serve as a 'roadmap' for assessing the relationships between independent variables (learners' characteristics, evaluation processes) and the dependent variable (functions and use of evaluation). In applying the model, a qualitative research strategy was taken, with case studies consisting of structured interviews triangulated by actors and sources. In our opinion that is an optimal toolbox for gaining an in-depth understanding of processes and mechanisms (a detailed description and discussion of the research methodology can be found in Annex 1).

Analysing evaluation as a knowledge production process

2.3. Our analytical model

The proposed model of evaluation utilization is presented in Annex 2. It perceives evaluation as a knowledge production process. It is distinctive in four ways. First, it explores the role of independent variables (institutional characteristics and quality of the process) that have been identified in the literature as important for evaluation use, but have not been addressed so far in empirical research in Europe. Second, it applies a systematic approach. The model presents, in a holistic way and in one logical chain, the interrelations between factors determining evaluation use, types of evaluation utilization and the functions of evaluation in the public sector.⁸ Third, a process approach has been applied. Evaluation is presented as a process of knowledge production with a range of inputs, activities, products and effects. Thus, it mirrors the logic of public intervention design and implementation.⁹ Lastly, the model incorporates different time perspectives (short term use versus longer-term influence) and the different attitudes of users (positive, neutral use, misuse of evaluation). The model consists of 5 main segments, relating to stages in the process of evaluation knowledge production and utilization. The first three segments (evaluation input, process, outputs) deal with the independent variables or determining factors, including the two factors that we are interested in. The two remaining segments (utilization and functions) are the dependent variables. These segments are discussed below by addressing four basic questions:

1. What do we mean by institutional context and input?
[refers to Segment I of the model]

2. What can an evaluation process consist of?
[refers to Segment P of the model]
3. What types of knowledge can be produced?
[refers to Segment O of the model]
4. How can the knowledge produced be used?
[refers to Segments U and F]

It also has to be noted that the discussion on each case study presented in Section 3 follows exactly these four questions.

2.3.1. What do we mean by institutional context and input?

Institutional context and institutional input into the evaluation process includes several aspects. As noted in section one, administrative culture can be a determining factor in evaluation use. It can be defined as the commitment to evaluation in a given administrative system and the extent to which it is used in policy-making. This factor is presented on the model in Part I1 (see: Annex 2). For our analysis, it is controlled by two means: a) choosing case studies only from Poland, b) choosing studies performed in the similar time period - 2000-2006. These were years of intensive preparation in Poland for EU accession and also includes the first years of Polish membership.

The characteristic of the evaluand (the researched intervention) is another potentially important factor in determining evaluation use. This factor can be defined by reference to several issues. It refers to the importance of the policy field to which a particular intervention is related. It can also be defined in terms of the size of the intervention (its resources, geographical scale and duration). Moreover, it can be defined in relation to the importance of the intervention for the policy field. The evaluand's characteristics are structured in Part I2 of the analytical model. For this research this factor has been controlled by three means. First, the research has focused on regional development as a policy field with the longest history of EU assistance, occupying a relatively high position on the policy agenda. Second, case studies were selected on the basis of an initial assessment of their size and importance in the regional policy field. It should be stated that in the Polish context every intervention directed at regional development was important because EU resources and programmes were in fact the only regional development interventions taking place.

As a determining factor, research timing has been defined according to the timing of the research within the programme cycle (ex-ante, on-going or ex-post); and its timing in the regional policy cycle (the transition period between old and new policy, a time of radical reforms, high system turbulence, the overall pressure for change of

Administrative culture

The characteristic
of the evaluand

Research timing

the policy etc.). The discussed factors were included in the model as context description (Part I3 - see: Annex 2). Timing with regard to the programme cycle was controlled by narrowing the focus of the study to on-going and ex-post evaluations. This focus on types of evaluation that are likely to be carried out over the next few years also strengthens the relevance of our research findings to current policy-makers.¹⁰

Timing in terms of the policy cycle was controlled to a limited extent by choosing a similar time-frame for the executed evaluations (years 2000-2006). In fact, every empirical analysis done in these years had potential importance because this period witnessed intense discussions on domestic regional policy directions and on arrangements for the implementation of Structural Funds.

Characteristics of learners

The characteristic of learners is also a potentially important determining factor in evaluation use. It is defined by three elements: the quality of human resources of the institution (that is knowledge on evaluation and experience of staff involved in the contracting and supervising the research); the institutional stability (how long it has been dealing with the interventions similar to the evaluated case); and the institution's position in the policy system (its ability to allocate resources, its power to make changes). As we stated in Hypothesis 1, this factor determines utilization in several ways. First, when institutions and their staff have experience in evaluation they better understand the functions and utility of evaluation. Second, if institutions are stable and well-established (and there is a strong 'institutional memory') then they accumulate knowledge from evaluations and draw on it in new situations. Third, when institutions have a strong position in the public policy system they can promote evaluation results and monitor the implementation of evaluation recommendations. The factor of learners' characteristics is presented on the model in Segment I, part I.4 (see: Annex 2). As the core of the hypothesis 1 it is one of the main subjects of our research.

Technical settings

'Technical settings' are understood as the resources assigned to an evaluation (size of the budget and allocated time) in proportion to the expected scope and sophistication of the research scope (as stated in the Terms of Reference of an evaluation contract). The balance (or imbalance) between assigned resources and the requirements of the contract, has a direct impact on the quality of results and, ultimately, their utility. The combination of very limited resources (or short deadlines) and a very ambitious, or lengthy list of research questions (often set by inexperienced Contracting Authorities) hampers the robustness of the evaluation. This can undermine the reliability and utility of the final report for potential users. Technical settings are outlined in the first segment of a model (Part I4). This

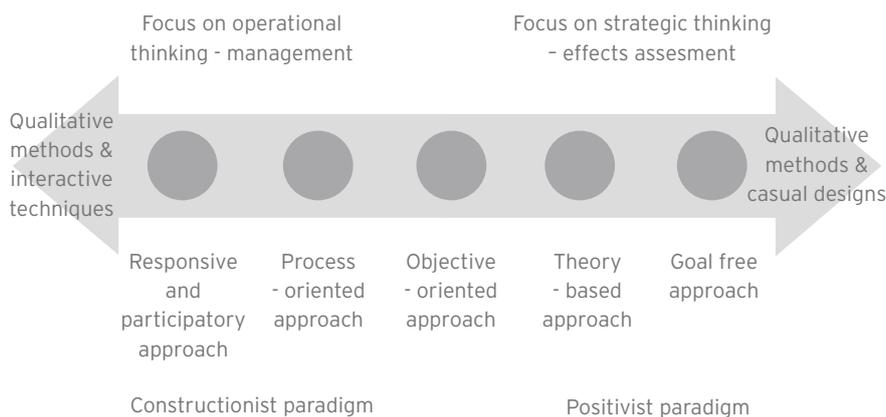
factor of balance (or rationality) between contract input and contract requirements has been taken into account during the study selection process. All selected cases have been completed in accordance with the ToR, which indicated that resources assigned to the contract were rational and in balance with objectives and methodology.

2.3.2. What can an evaluation process consist of?

The evaluation process provides the basis for Hypothesis 2 of our research as a determining factor of evaluation use. It is presented in the model in Segment P (see: Annex 2). The process factor includes three related issues: the choice of research approach; the role of participants in the process; and, the quality of the process. The following section discusses these separately, before proposing an integrated approach to their analysis.

We define an evaluation approach as a choice of research focus, strategy, and design and the selection of an appropriate methodological toolbox. On the basis of literature, five basic evaluation approaches can be distinguished. The basic dichotomy between evaluations that focus on programme effects or impacts and those that focus on programme management also applies here. This dichotomy is underpinned by a more fundamental divide between qualitative and quantitative research strategies (see Figure 3).¹¹

Figure 3. Evaluation approaches and their focus.



Source: (Olejniczak, 2008)

These related dichotomies determine technical choices on evaluation research design and methodology. Those evaluations focused on effects must apply a causal design in order to describe programme impacts and assess their value. This obviously requires a comparative approach.

Quasi-experimental designs and more quantitative methods (such as surveys, structured observations on sample, also in-depth case studies, etc.) can be applied here.¹² On the other hand, evaluating programme management can involve canvassing opinions and perspectives on programme implementation and bottlenecks. Here logic-process models, case studies, interviews, focus groups and workshops may be more appropriate.¹³ The main approaches are summarized in Figure 4, below.

Figure 4. Main evaluation approaches.

Approach	Description
Goal-free approach	Analysis of real programme effects (also side-effects), investigation of what programme has done (rather than what it was trying to do); analysis of causal relations between programme operations and changes in the target area or group as well as the broader environment in reference to initial needs (not objectives); dominance of the quantitative methods, use of causal designs.
Theory-based approach	Analysis of theories, assumptions and mechanisms of programme in terms of its effects; mixture of qualitative and quantitative methods, use of models and literature reviews
Objective-oriented approach	Description of effects and analysis of programme objectives in relation to its products, results and impact; analysis of changes in the target area / group with reference to set objectives; dominance of quantitative methods, use of causal designs.
Process-oriented approach	Analysis of programme processes and on-going performance with a focus on process improvement and addressing bottlenecks solving; dominance of qualitative methods, use of logic and process models.
Responsive & participatory approach	Analysis of programme processes and on-going performance in a collaborative manner - strong involvement of stakeholders and animation of their dialogue (participants make decisions, together with evaluator, throughout the whole process - on focus, design, data collection, programme assessment); dominance of qualitative methods, use of logic mapping and animation techniques.

Source: developed by authors on the basis of Alkin Et Al., 2006; Dahler-larsen, 2005; Mathison, 2005.

It is important to bear in mind that approaches to evaluations are bound to vary. Every programme, its policies, its institutional and political context and the actors involved, are distinctive and there is no standardised, optimal solution to the design and execution of these studies.

The role of participants concerns the number of evaluation participants (i.e. the number of stakeholders that are involved in the process) and the intensity of interactions between them (how often and in what way evaluators interact with stakeholders). Several studies have shown that securing the involvement of a broad range of actors beyond primary stakeholders and ensuring their interactive participation in all aspects of the evaluation can have positive impacts. This refers both to the utilization of the results of the evaluation and to other benefits gained from participating in the process (Cousins, Whitmore, 1998; Forss Et Al., 2002). The participation of a range of stakeholders in the preparation and management of an evaluation can ensure a stronger sense of 'ownership' of the process and a greater level of commitment to the results. The development of networks or organisational links through the evaluation process can ensure on-going dissemination and discussion of the results as the evaluation progresses rather than a 'one-off' presentation of conclusions in a formal, final report. This process can produce evaluation findings that are relevant and actionable and, thus, likely to be utilised by stakeholders. Beyond this, proponents of "participatory approaches" to evaluation emphasise how they can prompt a process of 'institutional learning' when doing things together and sharing responsibilities. Making actors work together in a critical reflection of systems, behaviours and relationships can consolidate an evaluation culture amongst the organisations concerned. Further, this can contribute to the development of more transparent, less hierarchical and more communicative organisational relationships (Patton, 1997). At the same time, it is important to note that organising participative evaluation processes can be time-consuming and demanding for evaluators and their success depends on the willingness of stakeholders to become actively involved.¹⁵

A final aspect of the research process refers to its quality. Here, our research focuses on two issues. The quality of input data made available for evaluators is obviously important. Poor data results in poor findings and weak conclusions (according to the simple rule: "garbage in - garbage out"), which in turn undermine the report reliability and use by audience. This is dependent mainly on programme operators and the quality of the programme monitoring system). The level of scrutiny of data collection, analysis and assessment is also important. This is dependent on evaluators' research proficiency and their adherence to scientific standards.

The role of participants of the evaluation research

Quality of the research process

2.3.3. What knowledge could be produced?

Reports quality & user friendliness

The next section of our analytical model deals with the types of knowledge that can be produced by evaluation research. Given varying programme features, institutional and political contexts and stakeholder agendas, approaches to the evaluation process are bound to vary. There is no single solution to the choice of approach, role of participants and components of the research exercise. Various approaches can produce high or low quality outputs. However, the quality of the final report depends on the extent to which these components of the evaluation process are harmonious and consistent. In other words, in order to produce a high quality report, it is vital that the focus of the research and its strategic objectives match the chosen methodology.

The quality of the report

Report quality can be defined in a subjective way (from the stakeholders perspective), in line with basic EC assessment template used for both external aid and Structural Funds programmes. The main criteria of quality are: meeting contract requirements, relevant scope, defensible design, reliable data, sound analysis, credible results, impartial conclusions and clear report (European Commission, 1999: vol1 p.180). A report's 'user friendliness' is also important. Taking into account time constraints and the myriad policy documents produced as part of current public administration processes, a complicated, lengthy and unwieldy report can easily discourage target audiences from reading even the most revealing findings. Simplicity, clarity and coherence are they key for gaining the attention of policy stakeholders. Report quality is dealt with in Part O1 of the model (see: Annex 2) In our research, it is controlled through the selection of evaluation reports that have been accepted by Contracting Authorities and EC (thus, in theory at least, they passed the assessment of quality).

Types of knowledge produced by evaluation study

The types of knowledge produced by evaluations can obviously have an impact on their use. Based on the review of literature we distinguish six types of knowledge, grouped in two segments: diagnostic and prescriptive (see Figure 5).

Figure 5. Types of knowledge.

Types of knowledge	Description
Diagnostic knowledge	- explains the things, shows bigger picture, indentify issues
know about	explains context in which programme operates, the opportunities and threats, also the socio-economic processes & mechanism related to the programme

know what	explains what is works and what doesn't in the programme
know why	explains why the things work or not in the programme, it also include reflection on the relevance of theories that underpin the mechanism of intervention
Prescriptive knowledge - gives solutions, explain how to cope with the things	
know how	shows how to improve things, change the underperforming issues
know who	points out who should do the job, who should be involved, which actors should interact
know when	points out when operation should be undertaken

Source: Developed by authors on the basis of Lundvall, Johnson, 1994; Nutley et al., 2003.

Types of knowledge are included in the model in segment O2. This factor is controlled to a certain extent because, based on theory, each completed and accepted evaluation should incorporate both categories of knowledge (diagnostic and prescriptive). They should identify issues and make suggestions.¹⁶

Dissemination is also a key stage in determining the utilization of evaluation (Mathison, 2005:118). The dissemination strategy of an evaluation can vary, depending for example on the balance between 'expert' and 'participative' approaches to evaluation. In participative evaluations, a single report will not be sufficient if learning is being promoted in a continuous manner within the programme. Dissemination strategies that go beyond the basic goal of making the final report available to those stakeholders directly effected can contribute significantly to the understanding and utilisation of the evaluation. Creating different fora or events for the presentation and discussion of evaluation results can not only bring the report to wider audiences but help to highlight the key messages and themes. More generally, these platforms also provide an opportunity to outline the broader benefits of evaluation as a tool for programme management, increasing the chances that reports will be utilised effectively. The dissemination approach is presented in our model in O3. This factor is controlled by choosing reports that have been made available in public domain.

2.3.4. How can evaluation knowledge be used?

At this point, we move our discussion to the dependent variables in our research, i.e. evaluation utilization and functions.

In terms of evaluation use, our definition refers to four components. The first concerns the aspects of the evaluation that are actually used. The findings of an evaluation (i.e. conclusions, results and

Different dissemination strategies

recommendations from the evaluation report) can obviously be drawn on by programmers. However, knowledge gained from the evaluation process itself can also be utilized. Interactions and discussions that take place between evaluators and stakeholders during evaluation research can be used (Alkin, Taut, 2003:6; Kirkhart, 2000). A second aspect refers to who uses the evaluation. Evaluations can be used by individuals, groups or institutions (Cousins, 2003). Users can also be differentiated according to their relationship with the evaluated programme. Adapting the Weiss et al. typology (2005:295) to the context of EU-funded programmes, the following potential users can be identified:

- the personnel and operators directly involved in the implementation of the particular intervention
- stakeholders of the programme - project applicants, direct and final beneficiaries
- wider European public - that is personnel that implements analogous interventions in EU
- interest groups - associations, organizations, networks of companies, regions, communities related to the evaluated type of intervention or field of the policy
- thematic and issue networks - experts, research institutions, think tanks and consultancies that follow the development in the particular field or policy
- main stakeholders of the policy - institutions that shape the particular policy, sets of interventions (heads of the related DGs in EC, national ministries)
- Mass-media and wider public opinion

The form and timing of evaluation use

The third identified aspect in this definition of evaluation use concerns the form of evaluation utilization. In the literature, there is a well established distinction between instrumental, conceptual and symbolic use.¹⁷ In addition, from the 1980s on, a number of authors have highlighted the potential for misuse or even abuse of evaluation. The final aspect in this definition is the time perspective of evaluation use. The effects of the evaluation can take place during or at the end of the research process. Changes in the examined intervention and the institutions directly involved in the programme can appear immediately or after some delay. Influence can occur in a cumulative way in the next editions of the programme or in similar interventions (Alkin, 2005; Kirkhart, 2000:7; Patton, 1997:20). For the purposes of our research we have developed a simple typology of evaluation utilization (see Figure 6). These are presented in "Segment U" of the model.

Figure 6. Use and misuse of evaluation.

Positive use	
Evaluation use	Occurs when evaluation study is used by programme actors for the purposes of the present programme (the evaluand).
Evaluation influence	Occurs when evaluation study is used by other actors or policy players for similar interventions to the analyzed programme.
No use	
No use of high quality report	Takes place when evaluation research is of high quality but due to radical and quick changes of the programme or contexts are not utilized. Its findings and recommendations, although of high quality, became obsolete and are not valid anymore.
No use of poor quality report	Takes place when report has low quality and merit. Its shallow findings and simplified recommendations could in fact be harmful to the programme
Negative use of evaluation	
Pseudo-use	Occurs when evaluation research has been done only to fulfill formal requirements (the sole objective is to produce a report - symbolic use) or when an evaluation is carried out to formally support a decision that has already been taken.
Misuse	Occurs under one of three circumstances: when poor evaluation findings are used for decision-making; when only positive conclusions are presented while critical findings are hidden (spin); or, when the report is manipulated, conclusions are cut out of context or the whole report is held back.

Source: Developed by authors on the basis of: Alkin, Coyle, 1988; Alkin, Taut, 2003; Cousins, 2004; Leviton, 2003.

Finally, in the model we have distinguished 5 functions that evaluation study can play in public management. These have already been explored in Section 1, but in order to clarify our definition, Figure 7, lists indicators that can be used to investigate particular functions.

The functions of evaluation

Figure 7. Evaluation functions and their indicators.

Function	How will I know it when I see it?
Partnership & ownership	<ul style="list-style-type: none"> • evaluation involved new programme stakeholders • evaluation involved the marginalized stakeholders • evaluation increased the level of interaction between stakeholders • evaluation established connections outside of programme
Performance & quality	<ul style="list-style-type: none"> • procedures have been changed in the evaluated programme • institutional competencies have been changed in the evaluated programme • procedures or institutional competencies designed for the new programmes have been based on evaluation findings
Planning	<ul style="list-style-type: none"> • strategy of evaluated programme has been changed • money has been reallocated within the evaluated programme • evaluation findings have been used in building new programmes' strategies
Cognition (understanding)	<ul style="list-style-type: none"> • study explained the context and mechanism of programme success / failure • study discussed the theories & assumptions behind the evaluated programme • study presented good practices • research increased an experience for new evaluation contracting / execution (understanding of evaluation importance and possibilities)
Accountability	<ul style="list-style-type: none"> • programme effects have been presented to stakeholders • programme effects have been presented to EC and key decision-makers • programme effects have been presented to public opinion

Source: Developed by authors.

Having set out our analytical model and provided definitions of the key factors involved in our research, the following section applies our research framework to evaluation case studies from Poland.

3. Evaluation stories from Poland

Section 3 presents analyses of six evaluations of EU-funded programmes carried out in Poland in the period 2000-06. There is particular focus on how the characteristics of the Contracting Authority for the evaluation and the features of the process itself determined the effective use of the evaluation. First we present the method of selection and then findings from our six case studies.

All six of them had the same structure of analysis and they follow the same pattern of presentation. Thus, for each case we addressed 5 questions:

1. What was the institutional context and input?
2. What was evaluation process?
3. What knowledge was produced?
4. How was the knowledge used?
5. What determined evaluation use?

First four questions follows exactly the discussion from the chapter 2.3 and the segments of the analytical model. Answers to the last question give the conclusion from the analysis of particular case study.

3.1. How were the case studies selected?

We are aware of the fact that the case study design has obvious limitations in terms of supporting theories and developing generalized observations (Bryman, 2004:33-55). However, following Flyvbjerg (2006), we use the case study approach to test the synthetic theory (parts of our meta-model) developed on the basis of earlier theories and concepts. In this way, the case study approach is justified. Moreover, by providing detailed analyses of cases, this approach gives practical, in-depth knowledge in an under-researched field and in a country context where research into evaluation use has, up to now, been absent. In fact, research shows that policy-makers tend to rely on quantitative studies when they want to make their decision reliable and convincing but they reach for qualitative studies when they want to learn about the inside, people, processes and mechanisms (Christie, 2007).

Our procedure of case studies selection had three objectives:

- To allow a controlling of the variables that are not tested in the research,
- To test the variables that form the basis of our hypotheses, i.e. to analyze the role of the learners' characteristics and the evaluation process in the effective use of evaluation
- To allow comparative, structured analysis between cases.

These have been obtained by applying a four step procedure.

The first step involved the collection of the full population of evaluation reports carried out in Poland for EU programmes between 2000 and 2006. A total of 180 evaluation studies were identified, out of which 118 were contracted in the chosen period. This step allowed us to control the factor of culture (all studies have been done in Poland, for Polish institutions and by mostly Polish evaluator teams). The time that has elapsed between the execution of these evaluations and our research gave the possibility of making the eventual effects visible.

The second step involved the selection of regional policy as our field of interest. This has been a crucial policy field in Poland for a number of years. This allowed us to control the factor of evaluation importance (all studies done for regional policy were of the interest to policy-makers, regional authorities and wider public). This stage narrowed our population to 30 reports.

In the third step, we reviewed this population of evaluation studies according to their timing. We have excluded ex-ante evaluation reports because we contended that their learning value is low since the next generation of ex-ante research will be done only after 2012. Instead, the focus is on the on-going and ex-post studies. Through this, we also controlled for the factor of timing. This left a population of 25 reports (7 ex-post and 18 on-going).

In the fourth and last step we chose 6 reports out of 25. The basis for the choice were Contracting Authorities. We wanted to choose evaluations that were contracted by six different institutions. That allowed us to explore hypothesis 1 on the influence of the characteristics of Contracting Authorities.

As a result following case studies have been selected (see Figure 8, below).

Figure 8. List of selected case studies.

Case number	Name of the programme under evaluation	Evaluation type
Case 1	PHARE STRUDER 2	Ex post
Case 2	PHARE Initiative 2	Ex post
Case 3	PHARE CBC Poland -Germany 1999-2001	Ex post
Case 4	SAPARD 2000-2003	Mid-term
Case 5	PHARE-Transition Facility (regional projects)	Interim
Case 6	ZPORR - Priorities 1, 2 (partly) and 3	On-going

Source: Developed by authors

The first case study concerns the ex-post evaluation of the PHARE STRUDER 2 programme. The programme ran from 1999-2000 as a continuation of STRUDER 1 which ran between 1993 and 1999. STRUDER 2 was a complex programme that focused on a relatively new theme in the Polish regional development agenda: support for SMEs and entrepreneurship through the provision of advice to businesses and also small infrastructure projects. The programme was implemented by the Polish Agency for Regional Development (PAR) and it contracted the consultancy firm Indygo to carry out the evaluation in 2002.

Our second case explores the ex-post evaluation of PHARE Initiative - PL9811 (INDYGO, 2004). The programme ran between 1999 and 2001 in the framework of Accession Partnership. The evaluation study was contracted out by the evaluation unit in PARP (Polish Agency for Enterprise Development). It was executed in 2004, by Polish consultancy Profile sp. z o.o. and Profile-Dialog sp. z o.o.

Third case study analyse the ex-post evaluation of the PHARE Cross-Border Cooperation programme between Poland and Germany - PHARE CBC (EUROREG et. al 2006). The programme ran from 1999-2001 and the evaluation was contracted by the Implementing Authority for European Programmes (Władza Wdrażająca Programy Europejskie - WWPE), formerly the Implementing Authority for the PHARE Cross Border Cooperation Programme (Władza Wdrażająca Program Współpracy Przygranicznej PHARE - WWPWPP). The study was contracted to a consortium led by PricewaterhouseCoopers, but evaluation tasks were carried out exclusively by the Center for European Regional and Local Studies (EUROREG), Warsaw University.

Case four - the interim evaluation of PHARE regional development projects in 2006 was part of a broader evaluation of PHARE projects in Poland (EMS, 2004). All active Phare programmes were evaluated on a sectoral basis. The Interim Evaluation Report covered PHARE support to projects in the regional development sector from 2001 to 2003. The assistance provided under the evaluated regional development projects totaled €134 million. The projects focused on the development and enhancement of infrastructure in the regions to increase their attractiveness for inward investment and to facilitate the development of SMEs. The Contracting Authority was the Office for European Integration (UKIE) and the evaluator was the consultancy company ECORYS.

Our fifth case study - the mid-term evaluation of the SAPARD 2000-2003 Programme was contracted out in 2003 by the Ministry of Agriculture, in an EU-wide tender procedure (AGROTEC et al., 2003).

The evaluation was conducted formally by International Consultancy - Agrotec SpA in cooperation with Polish institutes. In fact the team of experts was mostly recruited from a major public research institute (Instytut Ekonomiki Rolnictwa i Gospodarki Żywnościowej - IERiGŻ) that specialized in the field of agriculture development. The study was conducted within four months. Although SAPARD is commonly related to agriculture, it included a strong focus on regional development for authorities in rural territories as well as for mid-size farm production companies.

Our last case is "The study of the differences in payment transfers in ZPORR, at the level of voivodships" (Klimczak Tomasz et al., 2006). This was an on-going evaluation done at the end of 2006, for the Integrated Program of Regional Development 2004-2006. The Contracting Authority was the Managing Authority for ZPORR - a unit located within Ministry of Regional Development. Evaluation was conducted by a consultancy firm - WYG International (although some team members were academics).

3.2. What have we found out? Summary of the case studies

The findings from our case studies are summarized in the table below. The column headings follow our five analytical questions.

Figure 9. Summary of the case studies findings.

Case	Institutional context & input	Evaluation process	Knowledge produced	Knowledge used	Determinants of use
Case 1: Ex post evaluation of PHARE-STRUDER II	<p>The programme was important as it focused on entrepreneurship as a new regional development theme in Poland. It included new, regionalized approaches to business support.</p> <p>The programme experienced a significant cut in programme funding prior to the evaluation. This had serious implications for its status and the scope of the evaluation.</p> <p>Contracting Authority (CA) had experience of EU programmes, a core of staff with knowledge of the EU evaluation model and a network of regional and local authorities and agencies. It provided strong support and patronage for the evaluation.</p>	<p>Reduced funding limited the scope of the evaluation. A small sample of projects was used and there was narrow territorial coverage.</p> <p>CA experience adapted process to maximize evaluation utility. Evaluators were required to carry out in-depth study of small number of 'exemplar' projects. The aim was to develop 'best practice' case studies for dissemination.</p> <p>Small project sample allowed evaluators to take an interactive approach, based on CA regional and local network.</p>	<p>Knowledge produced was technical, programme specific and predictable for CA.</p> <p>Narrow scope of the study meant that dissemination of results was limited.</p> <p>'Good practice' case studies were produced and these featured in publications and presentations by CA and its successor organization</p>	<p>Use of knowledge limited: programme had ended and the CA had been liquidated.</p> <p>Case studies contributed to conceptual knowledge of local business support instruments in Poland.</p>	<p>Characteristics of CA vital. Positively, core of experienced staff adapted evaluation approach to maximize utility despite reduced funding.</p> <p>Negatively, institutional flux constrained use, particularly in terms of the accumulation of evaluation knowledge and experience</p>

Evaluation stories from Poland

Case	Institutional context & input	Evaluation process	Knowledge produced	Knowledge used	Determinants of use
Case 2: Ex post evaluation of PHARE-Initiative	<p>Liquidation and replacement of the CA after the evaluation significantly limited the use of the study for the development of knowledge and capacity.</p> <p>Relatively small Programme but of high political importance (restructuring of steel industry). Context totally changed during the study execution (steel prosperity made plans for re-training irrelevant).</p> <p>No strategic continuity between 2 editions of the programme. Evaluators with little knowledge in evaluation but high experience in social studies.</p> <p>Initial Contracting Authority (CA) team with high</p>	<p>In the initial phase the discovery of lack of data - no context data, no baseline data made the causal, quantitative analysis impossible to conduct.</p> <p>There was no consideration of the availability of data to answer evaluation research questions.</p> <p>Two years elapsed between the end of programme and the evaluation, making it difficult to trace beneficiaries.</p> <p>The quantitative strategy set out in the ToR was irrelevant for small, people-oriented intervention. Instead, the limited qualitative tools applied would be most revealing.</p>	<p>Quality of report relatively good.</p> <p>Knowledge produced not very revealing, some interesting findings on management but irrelevant in the context of new programme.</p> <p>Changes in the programme made it impossible to produce any valuable technical knowledge.</p> <p>Limited data and flaws in design made findings on conceptual issues not reliable.</p> <p>Interesting conceptual knowledge on how to</p>	<p>No immediate use of evaluation.</p> <p>No clear evaluation influence.</p> <p>None of the five evaluation functions were fulfilled.</p> <p>No planning or performance improvement because next edition started before evaluation results were received.</p> <p>No partnership because no interactions.</p> <p>Some weak signs of cognition function appeared (lessons on what is needed for effects</p>	<p>Radical change in the programme context: discontinuity between editions of the programme.</p> <p>High rotation in CA, no ownership of the evaluation and its results; simple execution of bureaucratic ritual.</p> <p>Lack of baseline and contextual data combine with irrational push towards quantitative approach.</p>

Case	Institutional context & input	Evaluation process	Knowledge produced	Knowledge used	Determinants of use
<p>Case 3: Ex post evaluation of PHARE-CBC PL-DE 1999 - 2001</p>	<p>evaluation experience but high staff turnover during the contract (no one from the original team lasted until evaluation closure). Strong pressure to quantitative approach; good balance of budget and timing, rigid requirements on evaluation criteria.</p>	<p>Extremely low interactivity - evaluators were left alone, there was nobody to discuss with them the progress and problems of the research process. Limited dissemination; no feedback received.</p>	<p>measure beneficiaries preparation for programme and importance of baseline. However this knowledge was lost due to CA rotation. Relatively good diagnostic knowledge. Prescriptive knowledge very limited. Broad and not operational recommendations.</p>	<p>assessment) - lost during shifts in CA. No accountability because weak causal proof and no public interest in the programme.</p>	
<p>Case 3: Ex post evaluation of PHARE-CBC PL-DE 1999 - 2001</p>	<p>Perceived as an important programme: EU prioritizes cross-border cooperation and high level of funding relative to other PHARE programmes. CA was stable for the duration of the evaluation and experienced in EU programmes and evaluation. However, there was subsequent discontinuity between PHARE/Structural Funds implementation structures and staff.</p>	<p>Challenge of wide range of project themes and types answered by developing three levels of analysis: programme, thematic/sectoral, and project. Comprehensive methodology combined clear CA requirements with flexibility for evaluators in designing methodology. Design incorporated quantitative (statistical analysis) and qualitative (interviews, study visits) approaches.</p>	<p>Well structured and accessible outputs produced, including final report, summary leaflet and best practice output for Neighbourhood Programme. Included different types of knowledge: conceptual (explaining mechanisms of cross-border cooperation), technical (related to project</p>	<p>As an ex post, no direct influence on programme. Evaluation informed other programmes of CA (e.g relating to indicator sets). Process broadened and deepened evaluation culture in CA. Cognitive knowledge of CBC improved and disseminated by evaluation team.</p>	<p>Institutional context crucial. Experience and stability of CA and evaluator produced comprehensive and innovative methodology. On the other hand - discontinuity between PHARE/Structural Funds and CA role limited the transfer of knowledge.</p>

Evaluation stories from Poland

Case	Institutional context & input	Evaluation process	Knowledge produced	Knowledge used	Determinants of use
	<p>Evaluators also experienced. Involved in evaluation of previous CBC programme.</p> <p>Evaluation motivated by need to assess impact of programme but also to transfer know-how and facilitate preparation of new Structural Funds cross border programme on Poland's eastern border.</p> <p>Significant time and funding allocated to evaluation.</p> <p>General ToR allowed scope for evaluators to develop methodology; previous collaboration between CA and evaluators created high level of trust.</p>	<p>It incorporated a future-oriented perspective as it looked to proven insights for future programmes in eastern Poland. Future stakeholders from eastern Poland incorporated into the research.</p> <p>Weaknesses in monitoring database and baseline data limited quantitative components. Evaluators attempted to fill some data gaps - time-consuming and demanding for beneficiaries.</p> <p>Participation of beneficiaries in qualitative aspect varied, particularly in eastern regions.</p>	<p>generation, indicators), and best practice examples (for eastern regions).</p> <p>Results (especially technical) were anticipated by CA.</p> <p>Dissemination a priority, particularly in eastern Poland. Series of seminars and publicity events organized.</p>	<p>Process strengthened links between PHARE CBC and INTERREG programming units in CA.</p> <p>Evaluation cited by CA as good practice model.</p> <p>Transfer of knowledge to eastern regions limited by uneven participation in the study and perception that the PHARE and Neighbourhood Programme contexts were very different.</p>	<p>Evaluation process also vital: multi-level analysis, qualitative and quantitative components, future-oriented dimension.</p> <p>Use limited by weakness of database on which some conclusions were built.</p> <p>Limits to the participative evaluation process undermined transfer of knowledge, particularly to eastern regions.</p>
Case 4: Interim evaluation of PHARE development projects 2002 - 2003	<p>CA experience of EU programmes strong. Accession process boosted CA capacity as it assumed important responsibilities in developing Polish evaluation system.</p>	<p>CA played strong role: the process was guided in detail by a new Polish evaluation handbook, created to facilitate stronger role for Polish authorities in evaluation after EU accession.</p>	<p>CA not surprised by results. Knowledge produced was technical and programme specific.</p> <p>Scope for conceptual knowledge constrained by</p>	<p>Evaluation use was limited: changes to programme were minor.</p> <p>Initial experience of testing the handbook was valuable.</p>	<p>Increased capacity and responsibility of CA determined evaluation use: study as an important stage in the evolution of Polish evaluation culture.</p>

Case	Institutional context & input	Evaluation process	Knowledge produced	Knowledge used	Determinants of use
	<p>CA contracted external consultants to contribute to the development of the new system, including Polish evaluation handbook.</p> <p>Evaluation handbook informed the PHARE study.</p> <p>Significant discontinuity between structures and staff involved in evaluation of PHARE and SF programmes. CA staff had limited involvement with post-accession evaluation system.</p>	<p>Process attempted to boost learning function (in order to boost Polish evaluation culture) with generation of programme-specific knowledge.</p> <p>Structured series of meetings between CA, evaluators and stakeholders in the programme throughout the evaluation process.</p> <p>Focus on learning produced a didactic rather than interactive approach to meetings.</p> <p>This illustrated tension between using the process of evaluation as an instrument of accountability/reporting and using it as a learning tool.</p>	<p>detailed requirements of evaluation handbook.</p> <p>Handbook requirements produced unwieldy report, limiting accessibility.</p> <p>CA only made the Executive Summary available.</p>	<p>Handbook now seen as basic guide for evaluation.</p> <p>Approaches first adopted here taken on in other evaluations.</p>	<p>Strong cognitive element, through learning and development of handbook.</p> <p>Handbook provided strong model but detailed guidance constrained the development of innovative or conceptual knowledge.</p> <p>Experience dispersed through institutional flux and staff turnover during shift from pre-accession to SF systems.</p> <p>Tensions in combining learning function with assessment of projects limited the impact of participative approach.</p>
<p>Case 5: On-going evaluation of SAPARD</p>	<p>Large programme with high policy status. The first in Poland to support progressive rural development.</p> <p>Important as preparation for Agriculture Operational</p>	<p>EC requirements were extensive and demanding. Focus on effects was flawed given delayed start to programme. Also tensions between objective-oriented and process-oriented perspectives.</p>	<p>Report was technical in nature. It targeted the main stakeholders but was not very accessible to beneficiaries.</p>	<p>Initial use limited to minor changes of implementation system but experience informed ex post evaluation.</p>	<p>Flawed research design (pushed by EC standard requirements) limited usefulness.</p> <p>Collaborative approach by evaluators, stakeholders</p>

Evaluation stories from Poland

Case	Institutional context & input	Evaluation process	Knowledge produced	Knowledge used	Determinants of use
<p>Programme 2004-6. However, perceived relevance declined after 2004-6 programme was launched.</p> <p>Complex implementation procedures delayed start from 2000 to 2002.</p> <p>Delay had impact on mid-term evaluation required by EC: evaluation focused on impacts but none had been achieved yet.</p> <p>CA had no practical experience of evaluation but some EC training.</p> <p>CA had high institutional stability and low staff turnover.</p> <p>Evaluators specialized in agricultural sector.</p> <p>Although general EC evaluation requirements focused on impacts, CA included some Polish-specific questions and required regular contact with evaluators.</p>	<p>CA and evaluators moved emphasis to process issues, also asking beneficiaries to forecast or assume likely impacts.</p> <p>CA struggled with small population of beneficiaries. Employed mixture of quantitative and qualitative methods on sample.</p> <p>The process was very interactive, particularly between evaluators and CA. Included strategic, technical and ad hoc meetings, though the process was not formalized. Helped to redirect study.</p> <p>Dissemination was in line with EU practice. It was available online and results were presented in meetings and seminars.</p>	<p>Knowledge produced was not surprising to CA but the inclusion of beneficiaries' views broadened CA perspective.</p> <p>Also produced knowledge for CA on how evaluation should be planned and implemented from the outset as a management tool.</p> <p>Broader lessons on the need for more regionally and horizontally variegated programmes to meet varied needs.</p> <p>This informed future interventions.</p> <p>Conceptual and technical knowledge on EU evaluation approaches produced.</p>	<p>Ideas from the evaluation informed implementation systems (e.g. monitoring) for 2004-6 and 2007-13 programmes.</p>	<p>and CA redirected evaluation to real needs to a certain extent.</p> <p>CA continuity and stability facilitated lesson learning.</p>	

Case	Institutional context & input	Evaluation process	Knowledge produced	Knowledge used	Determinants of use
<p>Case 6: On-going evaluation of ZPORR in 3 voivodships</p>	<p>Significant programme in terms of financial scale, content (regional development) and implementation (decentralized). Unsettled implementation process: frequent revisions, uneven pace of implementation in regions. Evaluation motivated by need to solve implementation problems and improve absorption of EU funds. Evaluators relatively experienced and highly motivated. CA inexperienced but had received training and were motivated. CA initially strong and stable but erosion of responsibilities over time. Limited budget for evaluation: short, intense perspective, well focused thematic and territorial scope. ToR required interactive approach, qualitative methods. It set out minimal but clear guidelines for final reporting, with a focus on process issues.</p>	<p>Focus on technical, process issues, combining simple monitoring analysis with diagnostic, qualitative component. Interactive approach: regular meetings involving wide range of stakeholders. This created a 'problem solving, network for the evaluation. Strong patronage by CA; took responsibility for: organizing meetings, implementing recommendations, dissemination. Strong CA patronage limited evaluation focus on its role. Comprehensive dissemination in regions and across Ministries.</p>	<p>Good quality report with clear, practical recommendations. Again, output decided by CA. Knowledge produced confirmed expectations. However, value in broadening perspective among central and regional partners. Knowledge was purely technical. Strong diagnostic component created robust basis for developing prescriptive knowledge. CA made clear distinction between systemic and process-related issues.</p>	<p>Detailed plan produced for implementation of recommendations. CA monitored implementation. Immediate use: simplification of payment procedures and implementation of recommendations in regions. Network created by evaluation process continued after end of study. Raised awareness of evaluation functions and provided model for other studies. Longer-term transfer of knowledge limited by discontinuity: implementation systems and structures changed completely for 2007-13.</p>	<p>Good short term use determined by participate evaluation approach. Active role of CA throughout evaluation stages ensured utility (although issue of separation of functions with evaluators). Quality of report and interactive approach guaranteed utilization of study beyond confines of programme. Institutional discontinuity and staff turnover prevented further influence.</p>

4. Conclusions

A synthesis and review of the research undertaken produces some important conclusions regarding the efficient use of evaluations in these programmes. The following section sets out these conclusions, looking at how the knowledge produced in these evaluations was used and assessing the factors that determined this use.

4.1. What knowledge has been used?

At a basic level, the case studies confirm existing theories on the types of knowledge produced and applied by different types of evaluation. In our case studies dealing with on-going evaluations, their role in generating technical knowledge on the management and implementation of programmes (i.e. process issues) is clear. This knowledge has immediately been used for operational purposes in the programmes concerned. On-going evaluations are usually not very revealing for stakeholders. Short-term evaluation studies (on-going evaluations are usually very intensive, 2-3 months long) can generate little innovative or surprising knowledge on the implementation system for people who have run the systems for months or even years and already know well how they operate. The knowledge provided in our case studies by on-going evaluation tended to be of a technical, programme-specific nature. All of these on-going evaluations provided specific technical knowledge that was potentially of use to programmers for operational purposes.

The extent to which our case studies contained knowledge that could be immediately utilized in the programmes concerned depended on how clearly the Contracting Authority stated their information needs. This referred particularly to demands in the Terms of Reference for prescriptive as well as descriptive information. This was definitely the case for the evaluation of PHARE projects (Case 4) and ZPORR (Case 6). In the former case, UKIE and the Commission developed detailed requirements for the evaluation (through the Polish Evaluation Manual) and this produced descriptive and prescriptive knowledge that was technical and programme specific. In the latter case, the Managing Authority of ZPORR clearly specified the basic objective and *raison d'être* of the study and required the evaluator to use a template for the evaluation recommendations (how, when and who should implement changes). On the other hand, Case 5 (SAPARD) illustrates that more ambiguous and unrealistic information requirements imposed by the European Commission resulted in the generation of very little valuable knowledge. Nevertheless, the generation and application of operational knowledge in these on-going evaluations suggests that

Focus on technical,
programme specific issues

Utility of technical
knowledge

Poland's performance in terms of the technical absorption of Structural Funds (i.e. ability to spent funds on time and in line with procedures) is progressing.

The added value provided by the evaluator is in giving an external perspective and combining stakeholders' points of view into a synthesized overview. All three of our on-going evaluation case studies confirmed this. In each case, the most revealing knowledge for Contracting Authorities concerned greater awareness of beneficiaries' perspectives. In EU-funded programmes, where the role of partnership in programme management and implementation is emphasized, this type of knowledge should not be underestimated. Some elements of conceptual knowledge appeared when the evaluation study included initial assessments of the actual impact of programmes (as in Case 5). Some conceptual knowledge was also produced in an indirect way when these studies revealed constraints and limitations in information gathering, prompting Contracting Authorities to reflect on the strength of existing databases. However, the main focus in these studies related to specific management and implementation issues and the provision of technical support (this was most visible in Case 6).

A basic contention of this paper is that efficient use of evaluations should contribute to the generation of strategic or conceptual knowledge that can inform policy choices and influence other interventions. In this respect, the picture in Poland is less straightforward. It is difficult to trace the impact of these evaluations beyond the specific confines of the programme concerned. In theory, this type of impact is associated with evaluations that included conceptual considerations. This type of knowledge can be provided by well executed ex-post evaluations which move beyond assessment of the process of implementing a programme to its actual impact. This applies to a certain extent to Case 3 (PHARE CBC). This evaluation involved some of the most comprehensive research in our cases. It was, therefore, able to generate some strong conclusions and interesting insights on cross-border projects (i.e. that projects were taking place along and not across borders). These insights influenced thinking in this policy. Case 3 also demonstrates that ex-post evaluations, though dealing with a closed programme, can provide technical knowledge that can be utilized in similar programmes. Generally, however, the production and application of this type of strategic or conceptual knowledge was very limited. This suggests that Poland still faces a significant challenge in strengthening its strategic absorption of EU-funded programmes (i.e. its ability to use funds in a strategic, development-oriented way).

Added value

Limited strategic and conceptual knowledge

Tension between process and effects oriented evaluations

It is also worth noting that the tension between conducting evaluations that focus on generating knowledge about process issues and evaluations that focus on generating knowledge about the effects of programmes was clearly shown in our cases studies. Both case studies that attempted to combine operational (process focused) with strategic (effects focused) considerations for the same programme had serious problems at the stage of analysis and with the roles that evaluators were supposed to perform (see Case 4 and Case 5). The role of our selected determining factors (the characteristics of the evaluation receiver and the research process) in producing this overall picture is discussed below.

4.2. What determined the use of this knowledge?

Our research focused on two factors that can be regarded as crucial in determining how efficiently evaluations are used: the characteristics of the institutions contracting and receiving the evaluation report and the characteristics of the evaluation research process.

4.2.1. Characteristics of institutions contracting the evaluation

Our first determining factor referred to the characteristics of the institution contracting the evaluation. A review of the case studies from this perspective produces several key points:

- Institutional patronage

Support from the Contracting Authority

Responsibility for the efficient utilization of an evaluation lies not only on the side of the evaluator but also on the side of the Contracting Authority (CA). The presence in the CA of an individual or a team with good experience and knowledge of evaluation, awareness of the benefits it can provide as a management tool, and enthusiasm for the process is essential. Case 6 illustrates this in the clearest positive way, as a dedicated team in the ZPORR evaluation unit was determined to play a strong role in the evaluation process from the initiation to the conclusion. The team was relatively inexperienced in terms of evaluation but this was compensated for by their openness and their willingness to interact and learn. However, two potential drawbacks can be observed here. First, the CA became so involved that it fulfilled some of the evaluator's tasks (final list of simplified recommendations, animating meetings with stakeholders). Second, the strong involvement of the CA to an extent excluded it from analysis as part of the evaluation. At the other end of the spectrum is Case 2, where lack of patronage (as a result of staff rotation in CA during contract execution) resulted in no use. Continuous patronage can also guarantee

that studies of limited use in short term can be used in a longer term (see Case 1 and Case 5).

- Institutional stability

A second characteristic of the CA that can have an impact on the use of evaluation is institutional stability. In this respect our hypothesis has been confirmed, usually by negative cases. The institutional flux and frequent staff rotation apparent in Polish public administration over a number of years has obstructed the accumulation of evaluation knowledge and experience and had a negative impact on efficient evaluation use. This is illustrated in several of our cases. In Case 1, institutional flux, linked to national elections, meant that the evaluation team was broken up, the Contracting Authority was disbanded and the evaluated programme itself was discontinued. In these circumstances, it is impossible to expect the evaluation to serve any meaningful use. In Case 2, frequent staff rotation in the evaluation unit in the CA also undermined the eventual utility of the report: without consistent guidance and support from the CA, the evaluator simply carried out a basic, bureaucratic exercise. Another cause of institutional instability was Poland's accession to the EU. A clear line was drawn between pre and post-accession evaluation structures and approaches. For instance, the evaluation knowledge accumulated under PHARE programmes in Case 4 was largely lost when the Ministry of Regional Development took over from UKIE and developed its own approach using new staff. It is also worth noting that on-going changes to the system for implementing Structural Funds in Poland can also threaten the accumulation of evaluation knowledge (see the uncertainty surrounding the evaluation team responsible for ZPORR - Case 6). In terms of positive examples, two case studies can be referred to. In Case 5, the benefit of maintaining and using evaluation knowledge before and after accession was demonstrated. Case 5 provides a positive example of retaining knowledge that seemed irrelevant at the beginning. Case 3 is interesting as the relative stability of the CA and its long-term relationship with the evaluators was a key determinant in ensuring good evaluation outputs and use.

- Quality of the data

A final aspect related to the characteristics of the Contracting Authority concerns the quality of the data available for evaluation. Authorities implementing EU-funded programmes have a responsibility to maintain systems for gathering and aggregating information on their progress. However, all of our case studies were hampered in the generation of knowledge by weaknesses in their programmes' information base. Where there is an absence of consistent, quality statistical information on the progress of projects, it is impossible to assess the impact of

Organisational
flux

Weaknesses in
data bases

a programme, draw robust, overall conclusions and produce new conceptual insights. For evaluations focused on operational issues, the problem of data availability may be less important. They are usually based on primary data (usually qualitative) that can be collected 'on the spot' with no comparative analysis required. However for ex-post evaluations that seek to assess a programme's impact, lack of baseline and contextual data hamper attempts of investigating causal relations, and as result strategic reflection on the programmes utility and mechanisms of development is lacking. A point of reference is needed to measure programme effects. Thus, regular data collection and monitoring are required. Without it there can be no real assessment of effects (that is going beyond mere listing of products).

In the absence of baseline data, one option for evaluators is to apply a constructivist approach instead of comparative, positivist approach. This would involve asking stakeholders if they regard the programme to be successful or it could mean comparing the programme to similar interventions. However, neither of these approaches produces robust conclusions. It is also worth noting that where there is no coherent system of data collection and monitoring, evaluation studies often attempt to build their own data sets (see, for example Case 3). This can improve the data base to a certain extent but it is very time-consuming for evaluators. Moreover, it places an additional burden on beneficiaries as, in the absence of central sources, they are most the reliable and basic source of information.

4.2.2. The role of the research process

Our second determining factor referred to the characteristics of the evaluation process. Again, review of the case studies produces some important conclusions.

- Research design

From our research it is clear that the design of the evaluation research is essential to its subsequent use. Important questions in this respect concern evaluation functions, when the methodological choices should be made for the evaluation, and by whom. Looking at the evaluation process in our case studies, it is clear that crucial decisions regarding the design of the research were made at the very outset. European Commission manuals and theoretical models simplify the reality by suggesting that methodological choices on how the evaluation should be structured are taken after contracting out and are the responsibility of evaluators. Our cases show that the framework for these choices is set at the stage of constructing the evaluation Terms of Reference. Mistakes made at this stage can result in serious tensions and problems at later stages in the research. The clearest example is provided by Case

5 (SAPARD), where the European Commission imposed unrealistic, complicated tasks, oriented on effects assessment without taking into account the programme's situation (e.g. the Commission required impact analysis despite the fact that the programme was still at an early stage). As a result, the scope of the study was too extensive and time-consuming. This resulted in a bureaucratic exercise that did not provide the Commission with any useful information. Another negative example is Case 2 PHARE-Inicjatywa, where there was no reflection on whether there was data available to answer the research questions. Again this resulted in frustrating procedures and limited conclusions.

On the other hand, Case 1 (STRUDER 2) provides a positive example of how discussions between the Commission and the CA resulted in a more realistic research design. In Case 6 (ZPORR) the CA set out minimal but clear guidelines (insisting on a participative approach to the evaluation) and also decided to narrow the scope of the study in order to make it more focused and manageable. This resulted in good use. It is interesting to compare Case 3 (PHARE CBC) and Case 4 (PHARE projects) in this respect. Evaluators for PHARE CBC had clear requirements (evaluation objectives) but flexibility in methodological approach. This, combined with experience of the evaluation team and good levels of trust with WWPE, produced good outputs. In PHARE projects the evaluators were given detailed instructions on every stage of the evaluation process. The ToR simply referred them to the new Polish Evaluation Manual. Although this guaranteed that the evaluation met the minimum requirements set by UKIE and the Commission, it lessened the likelihood that interesting or innovative knowledge would be produced. This, in turn, limited the scope for evaluation utilization.

- Interactivity of the process and building trust

Analysis of our case studies suggests that interactive or participative approaches to evaluation are crucial to their efficient use. Interactivity is important for learning and improvement. People and institutions learn faster when they are involved in the process. They also internalize knowledge faster once they are involved in the process of its creation. Critical findings are more likely to be accepted and acted upon. When stakeholders take part in generating this knowledge, they feel co-responsible for the findings and perceive evaluation results more as an internal assessment rather than an external critique. Important in this process are levels of trust and a close links between evaluators and stakeholders. A clear example here is Case 6 (ZPORR) where the difficult topic of financial blockages was analysed on the basis of three voivodships. In this case, although a number of problems were identified in the evaluation, this did not result in a "name and shame exercise". Instead it provided knowledge for all actors and created a problem-

Regular dialogue
is crucial

solving network. Additionally, Cases 5 (SAPARD) and 1 (STRUDER) illustrated that interactions and dialogue between CA and evaluators helped to overcome problems in design and data in the execution of useful report. Case 2 (PHARE Initiative) provides a negative example. The absence of any dialogue between the evaluators and the stakeholders resulted in a flawed design, based on unrealistic assumptions on the quality of data available.

- The role of the evaluator

The points made above illustrate the key role an evaluator has in determining the utilization of the evaluation. In order to grasp the surprisingly wide spectrum of potential evaluator roles identified in the case studies, a matrix was developed (see Figure 9). The vertical dimension shows the level of methodological sophistication adopted by the evaluator. The horizontal dimension indicates the intensity of interactions during the research process.

Figure 10. Matrix for analysis of evaluators' roles.



Source: Developed by authors.

The matrix produces four basic categories with associated roles that evaluators can play:

In the bottom-left square, the toolbox utilized by the evaluator is very limited. Basic methods are employed such as desk research, a limited number of interviews, small expert panels or focus groups or small scale questionnaire research. Scientific scrutiny (e.g. through sample selection, case studies choice, triangulation of sources, representation in interviews) is not incorporated. The evaluator has only limited contact with the Contracting Authority and stakeholders are approached as research subjects rather than active participants in the evaluation. Under this approach, the likelihood that the evaluation will produce knowledge that can be utilized in a strategic or conceptual way is low. Moreover, the scope for evaluators to play a part in encouraging a learning process as part of the evaluation is limited.

The evaluator as a detached consultant

In the top-left square, the evaluator takes a strong scientific approach, aiming to establish objective results. The programme, its activities and systems are scrutinised in a detached way and distance from the stakeholders is maintained (treating them as subjects of the research). The scientific robustness of the research is very high in terms of the research design and methodology and in conducting the analysis. The evaluator interprets evidence and draws conclusions, issuing a report about findings. Findings are communicated, not discussed. It is then up to the programme to digest and follow up on the report, often after the departure of the expert, whose contractual obligation was limited to delivering the study. Given the level of scientific scrutiny, this approach can potentially provide useful conceptual or strategic insights, though the potential for learning is limited by the evaluator's detached approach.

The evaluator as
a scientist or expert

In the bottom-right square, the toolbox utilized by the evaluator consists mainly of qualitative methods and interactive techniques, including: desk research, interviews, expert panel, focus group or small scale questionnaire research, logic models. Scientific scrutiny is limited (e.g. in sample selection, case studies selection) but triangulation of sources is incorporated (gaining different perspectives on the same issues or themes). Here, the evaluator plays the role of animator or facilitator by bringing together the different perspectives of programme stakeholders. On the one hand, the scope for producing conceptual or strategic knowledge that could be used outside of the programme is limited by lack of scientific scrutiny. On the other hand, the potential for learning to take place as part of the process and for stakeholders to be committed to using evaluation results is high. Most stakeholders are drawn into the evaluation process both as research subjects and research partners.

The evaluator as
a facilitator

In the top-right square, the evaluator follows a robust scientific approach, maintaining objectivity in the research. However, he does not avoid interactions with stakeholders. On the contrary, he shares and discusses decisions on design and methodology with stakeholders and involves them in the discussion of results as they appear. Here, the evaluator plays the role of educator rather than detached expert, involving stakeholders in a dialogue about the programme strategy, achievements and processes and providing advice based on knowledge and experience. In this case, the evaluation can produce conceptual or strategic knowledge and also encourage a process of learning amongst stakeholders.

The evaluator
as mentor

Again, it is important to stress that there is no optimal model for the evaluator's role. Several hybrids can exist within this broad

categorisation. However, some points are worth highlighting. First, if we accept that a current priority in the Polish context is building evaluation capacity and an evaluation culture, then it seems logical that evaluators should consider including functions that educate or facilitate learning as part of the research process. Nevertheless, it is important not to underestimate the extent of this challenge. Case 4 (PHARE projects) is interesting in this context. The evaluation attempted to take a participatory approach and this had a certain amount of success in disseminating knowledge about the role and benefits of evaluation. However, in this case, the control and accountability function (for reporting to UKIE and the Commission) was at odds with a learning-oriented process of evaluation. This required programme actors to be open and honest in their responses, and to expose the difficulties and tensions which have been encountered. Evaluators found it very difficult to perform an educational/facilitating role and a reporting role at the same time. Moreover, facilitating stakeholder participation must take into account distinctions that exist between groups in terms of position, authority, confidence to participate and understanding of evaluation. The example of the differing partner contexts on Poland's western and eastern borders in Case 3 (PHARE CBC) is relevant here.

5. Recommendations

“When people realize things are going wrong, there are two questions they can ask. One is, ‘What did we do wrong?’ and the other is, ‘Who did this to us?’. The latter leads to conspiracy theories and paranoia. The first question leads to another line of thinking: ‘How do we put it right?’

(Lewis, 1997:121)

The conclusions listed above raise two main questions. How can we anchor evaluation knowledge and experience in institutions and thus ensure that evaluations contribute to strategic planning? How can evaluation processes be improved to support a process of learning and encourage the effective use of the knowledge that they produce? The final section of the paper makes some recommendations under these headings.

5.1. How can we overcome systemic weakness?

- Strengthen the evidence base

A clear message from our research is the crucial role a strong base of programme information plays in determining the credibility and therefore utility of evaluations. It was also clear that there are limitations in the Polish system in this respect. In order to strengthen the initial database, ex ante evaluations and on-going studies (performed at the beginning of a programme cycle) should be used for creating a base or reference point for future strategic ex post comparative analysis. These evaluations should go beyond mere logical analysis of the programming document structure. They should involve empirical study and be carefully designed, with the information needs of the future ex post in mind.

Obviously, it is already too late for the old generation of Structural Funds programmes (2004-6). Their on-going evaluations could be used in a selective way only to explore impact issues. Since the baseline for comparison have not been created there, it will be extremely challenging (if not impossible) to assess real effects of these funds. However, organizing small research packages on specific parts of programmes that have already been completed (or are about to be completed) could help in three ways. First, it could provide initial indications on programme effects before the full ex post evaluations get underway and in time to inform the planning process of the next generation of programmes. Second, it could test ex post methodologies and approaches on a smaller scale (as a type of ‘pilot ex post’). Third,

Assessments
of impact

The need for a long-term thinking

it could also show needs and challenges that have to be coped with, in order to avoid similar mistakes in the new system (2007-13).

The ex ante studies for the 2007-13 programmes were completed in 2006 and 2007. However, in order to strengthen the evidence base and reference points for assessments of their future impact, on-going evaluations of these programmes in the next year or so should include separate studies - absolutely empirical - that create the baseline for future comparison. Again, for practical reasons, this could involve focusing impact-related studies on smaller, specific parts of programmes (priority axes, actions). The last recommendation is challenging as it relies on an acceptance by policy-makers and civil servants that they initiate processes whose benefits will only become apparent later on in the policy cycle when, given institutional flux in Poland, there is a strong likelihood that other institutions and actors will be responsible. These studies won't have an immediate effect but they will finally create a sound empirical basis for a future impact analysis. This process of accumulation of information and knowledge should start now.

- Strengthen institutional memory

Anchoring evaluation knowledge

Institutional flux and frequent staff rotation is inevitable in administrative systems under construction or in transition. The challenge is to anchor at least some of the knowledge and expertise gained in evaluation processes even when the people directly involved move on. One solution would be to ensure public databases of all evaluation studies are available as a source of information. PARP, and recently KJO have already started this process of collecting and storing all evaluation reports. In order for this to be effective, full evaluation reports should be stored, including methodological sections. Basic data on the evaluation context (time and budget of the research) should also be attached as this would allow people to learn about specific research constraints and characteristics of the process. Finally, the reports stored, particularly on-going evaluations, should have a clear division between operational and systemic conclusions & recommendations. Thus, in cases where staff or institutions have changed, their replacements could have a point of reference and at least partial memory of the past programme, its problems, the solutions that have been found, etc.

5.2. How can evaluation processes be improved?

Recommendations under this heading can be grouped according to those aimed at the Contracting Authority and those aimed at the evaluator.

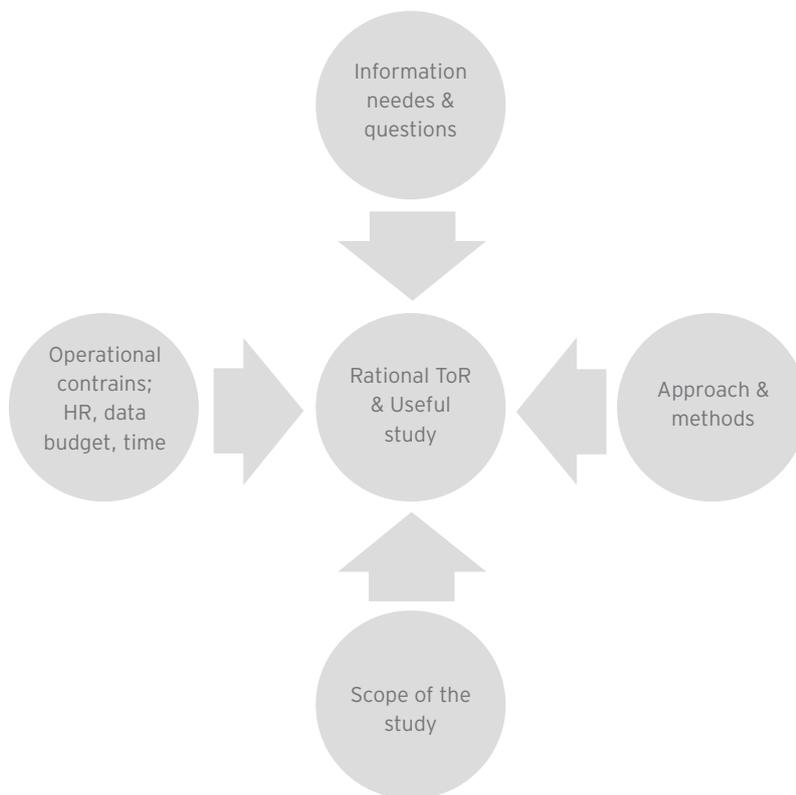
5.2.1. Contracting Authority

- Create clear, appropriate Terms of Reference

Designing robust Terms of Reference is key to improving the practical implementation of evaluation research and increasing the prospects for utilization. Deciding the aims and structure of the evaluation process takes place not after the study is contracted out but before, at the stage of ToR design. Logical mistakes at this stage can hamper the utility of evaluation. Thus, the Contracting Authority is required to make some important decisions, balancing different components of the evaluation process (see Figure 11).

Balancing information needs
and operational constraints

Figure 11. Balancing key logical elements of evaluation.



Source: developed by Authors

Each of these four elements involves decisions that will set the framework for the subsequent evaluation process. The main questions are:

- Should the evaluation study focus on process issues or effects (i.e. should the evaluation look mainly at operational or strategic issues)? This dilemma applies mainly to on-going evaluation. One study cannot combine both aspects and it is important that

the decision on which aspect to cover, fit the real information needs. For instance, focusing on effects or impacts when there are important operational problems to solve will hamper the evaluation and limit its utility (as stakeholders will perceive it as a control mechanism).

- Does the methodology set for the research match the anticipated functions of the evaluation? Approaches and methods for evaluation design are to a large extent determined by its functions and focus (i.e. on the choice between research into process and effects). The objective of assessing and improving programme processes requires interactive methodological approaches, partnership building tools and flow analysis. Assessment of effects emphasizes research on causal designs and theory-based approaches. Ideally, these types of evaluation should not set preconceived goals or narrow objectives (i.e. assessing programme impact against its initial forecasts). Imposing this as part of the research process risks turning the evaluation into a control exercise. Using goal-free approach would allow investigating real effects and reducing a bias of following only programme designers' initial assumptions. Moreover, potentially useful knowledge, for instance on the unanticipated impacts or 'side-effects' could be developed.
- Are the evaluation questions and design realistic, given the level of resources available for its execution? Where there is limited time and funding available, narrowing the scope of the study (e.g. in geographical terms) can ensure greater focus and generate more useful knowledge. This is especially applicable to process analysis where the focus of the evaluation can be on qualitative analysis of specific cases rather than broader, longer-term quantitative research. On the other hand, focus on effects (especially for ex post evaluations) requires a broader, more comprehensive scope which is in turn more time-consuming and costly.
- Finally, it is important to ask whether the design of the research set out in the ToR is realistic given the availability of data. It is important at this early stage to check the quality of the data sets. This applies particularly to evaluations focusing on programme effects or impacts. Filling in extensive background data during the evaluation process (e.g. monitoring data that should be collected as part of programme management and implementation systems) hampers analysis and risks turning the evaluation into a monitoring exercise, generating knowledge that is relevant to technical (products fulfillment) rather than structural absorption (real utility of effects).

There is no optimal solution to the choice of approach, role of participants, components of the research exercise and the functions of an evaluation. Various approaches can produce high or low quality outputs. And the number of possible approaches goes beyond those five presented in this analysis. However, the utility of an evaluation depends on the extent to which the overall research design ensures that these components of the evaluation process are harmonious and consistent. In other words, in order to produce a high level of utility, it is vital that the focus of the research and its strategic objectives match the chosen methodology. Be aware of the choices and make them rational, well grounded in the context.

- Input into the methodology

From our cases studies, it is evident that Contracting Authorities can play different roles in setting the methodology for the evaluation research. On the one hand, the Terms of Reference could set just the main objectives and leave the choice of methods entirely open. Similarly, they could set the methodological approach to be taken in broad terms (e.g. desk analysis of monitoring data and a more detailed analysis of a sample of projects) and leave it up to the evaluator to respond with more detail (eg. sampling methods and the potential coverage of questionnaires). On the other hand, the CA can take a more prescriptive approach, setting out in detail what the research process and outputs should consist of. Our case studies show that although the detailed prescriptive approach guarantees that the CA gets the outputs it is looking for, lack of flexibility narrows the scope for generation of new or innovative knowledge. Where the CA provides a broader frame of reference, specifying information needs while leaving methodological choices in the hands of the evaluator, there is more opportunity for evaluators to bring in their own ideas and propose innovations. It should be noted that, even if the CA does not prescribe its requirements in the Terms of Reference, they need to have thought them through with sufficient clarity and detail to be able to appraise and compare the bids which are submitted.

- Taking an interactive approach and providing patronage

A basic recommendation resulting from our research is that the utility of evaluations will always be boosted where there is a strong, interactive relationship between the CA and the evaluator. The close involvement of the CA at all stages of the evaluation research serves a number of purposes: it provides a 'champion' for the project, ensuring that it continues to be seen as a priority in the institution; it boosts evaluation experience in the CA; it ensures that the evaluation research meets the needs and interests of the institution; it can help to ensure the participation of other stakeholders and beneficiaries in the evaluation

Balancing clarity and flexibility
in research design

Contracting Authorities as active
patron of the studies

process; and it can ensure that evaluation findings are disseminated properly. All of this increases the scope for efficient evaluation use. However, it is important that this level of interactivity does not result in the CA taking on responsibilities of the evaluator, particularly where it is one of the actors being evaluated.

5.2.2. *The evaluator*

- Design clear, appropriate methodology

The issues and choices set out in Figure 11 are also applicable to the role of the evaluator. Within the frameworks set by the CA, the evaluator must become familiar with the programme context, delivery mechanisms and partnership organisation and at each stage of the research process make choices on the approach to be taken. This involves a similar process of balancing different considerations (evaluation objectives and scope, resource issues, data constraints) faced by the CA but at a more detailed level. Is the methodological approach appropriate to the overall objectives of the evaluation, are the research questions appropriate and can they be answered given time constraints, funding levels and quality of the database? As noted above, close, on-going links between the CA and the evaluator throughout the evaluation process can ensure that solutions are found to these issues.

- Look at the scope of the study and decide what role to play

Referring back to Figure 9, it is clear that evaluators can play a range of roles in the evaluation process. From our research it is arguable that one of the key factors of success is high interactivity. Thus, the most positive roles are those two on the right side of Figure 9. We argue that the choice between educator-mentor and animator-facilitator is determined again by the orientation of the study (the dichotomy discussed extensively earlier and presented on figures 2 & 3).

Studies focused on processes (operational issues) have highly utilitarian character: their main aim is to provide current stakeholders (usually programme operators) with quick solutions to pressing problems. Thus, the evaluator should listen carefully to its client and other stakeholders voices but also go beyond this and conduct an assessment (or rather a problem-solving exercise) in a collaborative way. The added value comes from listening to the views of a wide range of stakeholders and combining different perspectives. 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. The role of animator and

The evaluator's role in
process-oriented evaluations

facilitator fits exactly this purpose. In contrast, for several reasons, the educator role is not appropriate in process oriented studies. First, the CA has the main technical knowledge on process issues and it would be pointless for the evaluator to take an 'educator role'. Second, there are usually time constraints in this type of evaluation (bottlenecks have to be identified and solved quickly) and this limits the use of highly sophisticated, scientifically scrutinized methods. Third, the educator-mentor role introduces hierarchical relationships (with evaluator at the top) while problem-solving situations urge for flat, participatory approach and the empowerment of the main actors.

On the other hand, studies that focus on effects (i.e. strategic issues) have a more cognitive, reflective character and are directed not toward one stakeholder group but more pro public bono. In order to understand the bigger strategic picture, test or develop theories, the evaluator should play an educator-mentor role. This involves taking a robust scientific approach and ensuring objectivity in the research. Strong interactive relationships with stakeholders are maintained (by sharing and discussing decisions on design and methodology). Following supervision from the evaluator in the research process, stakeholders can participate in discussion of results and findings as they appear. However, the evaluator retains complete control of decision-making for the study. This allows the evaluator to keep some distance from current programme operators and objectivity in assessments of programme effects. Again, for several reasons, the role of animator-facilitator is not appropriate in this type of evaluation. First, the shortened distance with current operators, close collaborative relationship with them and effort to "get in their shoes", narrows the analysis of effects only to a fragmentary and often short-term perspective. Second, the basic methodology and low scientific scrutiny associated with the animator's role limits the scope to produce scientifically robust conclusions. To be credible and convincing, studies of programme effects have to be sophisticated and well grounded in thorough analysis. It is worth noting that the fact that ex post evaluations come so late in a programme cycle is to be welcomed. Although they cannot be directly used for immediate programme management, these studies give evaluators the opportunity to play the more reflective role of educator. In the absence of pressures from programme managers, they can reflect on the merits of regional development strategies, from a broader strategic and theoretical perspective pro public bono.

The evaluator's role in effects-oriented evaluations

Annexes

Annex 1. Detailed methodology of the research

For the purposes of this research, a hypothetico-deductive method of reasoning has been applied. The research strategy was a qualitative one, based on the meta-modeling followed by case studies. The meta-model had formed the background for the core of the research, which examined the relationships between independent variables (learners' characteristics, evaluation processes) and a dependent variable (functions of evaluation). This analysis was made on the basis of six case studies of Polish evaluations.

As concern Research plan, methods and data sources the proposed research plan were implemented in 5 phases.

- Phase 1 - Development of the meta-model.
- Phase 2 - Selection of the case studies
- Phase 3 - Development of the case studies toolbox
- Phase 4 - Case studies execution
- Phase 5 - Analysis and conclusions.

Below we briefly present each phase in more details - its objective, methods and data sources used.

Phase 1: Development of the meta-model

The objective of constructing the meta-model was threefold:

- to present the hypotheses in a clear form, grounded in the complex system of real-life interrelations;
- to gather, organize and clearly present the definitions, types, and factors of evaluation utilization, based on the literature and on extensive practice;
- to help in mapping and following the tested interrelations.

The developed meta-model has been a road map of the research. It was developed in the form of a one-page scheme. Later on it was a process map of factors of evaluation use and the concepts and stages of evaluation utilization. Each element of the model has been clearly defined and illustrated with examples of evaluation research from the European Union.

Research methods used in this phase were of qualitative nature. The meta-model and sets of definitions on evaluation utilization was built on

the basis of in-depth literature review as well as analysis of international practice. The review covered the period from 1990 (when the systemic analysis of evaluation utilization started) up to present. Some earlier (pre-90s), key publications were also included. The review focused on English language sources. Specific data sources included:

- major scientific periodicals dealing with evaluation and regional policy (American Journal of Evaluation, Evaluation & Program Planning, Evaluation Review, Evaluation Practice, New Directions for Evaluation, Studies in Educational Evaluation, Evaluation, Science Communication, Regional Studies)
- key textbooks on evaluation (mostly from SAGE Publisher), selected evaluation reports carried out for the Structural Funds (mainly in UK, Ireland, Germany, Austria, Sweden)
- EU evaluation manuals, documents and official publications on evaluation
- Evaluation studies from Structural Funds programmes, analyses and studies on EU regional policy and Structural Funds system

Phase 2: Selection of case studies

The objective of case studies selection was, again, threefold:

- to control the variables that are not tested in the research
- to test the variables that form the basis of our hypotheses, i.e. to analyze the role of the learners' characteristics and the evaluation process in the effective use of evaluation
- to allow comparative, structured analysis between cases

The case studies selection was executed in three steps. They are discussed in the main text of the report.

Data sources at this stage included lists of evaluation studies (and sets of evaluation reports) from all Polish institutions that were fulfilling a function of Implementing Authorities for pre-accession and Structural Funds. This includes: Polska Agencja Rozwoju Regionalnego, Polska Agencja Rozwoju Przedsiębiorczości, Władza Wdrażająca Programy Europejskie, Urząd Komitetu Integracji Europejskiej, Krajowa Jednostka Oceny.

Phase 3: Development of the case studies toolbox

The objective of this phase was to develop a toolbox for the analysis of case studies.

The toolbox outlined the scope of the desk research for the case studies, listed the relevant contextual documents to be obtained and analyzed for each case, and set out the range of the interviews and questionnaires, interview scenarios, the initial coding system for qualitative data analysis, etc. The analysis toolbox was structured in line with the elements of the analytical model.

The research methods at this phase included the review of the literature (textbooks on social research methods, new literature on qualitative data analysis) as well as the meta-model (used to develop a tool box to assess the case studies). The structure of interviews and questions mirrored the elements of meta-model in order to cover all of its analytical components. The researchers used to some extent an appreciative inquiry when conducting the interviews. In principle, this approach is based on asking questions in positive terms (e.g. what positive changes have been observed, what has been satisfying, what would be the ideal situation - from the perspective of the interlocutor). This approach has been proved to be more constructive in research of organizational processes and changes (Preskill, Catsambas, 2006). Data sources in phase 3 were based on materials developed in phases 1 and 2.

Phase 4: Case studies execution

The objective of this phase was to collect the data for each case study.

Methods for each case study included such qualitative methodological approaches as: review of programme documents, desk research of earlier studies and relevant expertise, interviews with contracting authorities, interviews with evaluators, subsequent interviews/questionnaires with other key programme stakeholders. It should be noted that each case study followed the same research procedure.

In phase 4 both primary and secondary data sources were used:

- national, administrative statistics and databases of evaluation reports (already reviewed)
- programme documents and working materials obtained from Managing Authorities of the particular interventions
- related articles, public documents and studies obtained through on-line newspaper archives, government and research institutes' web pages, scientific databases
- interviews with evaluators and with Contracting Authorities

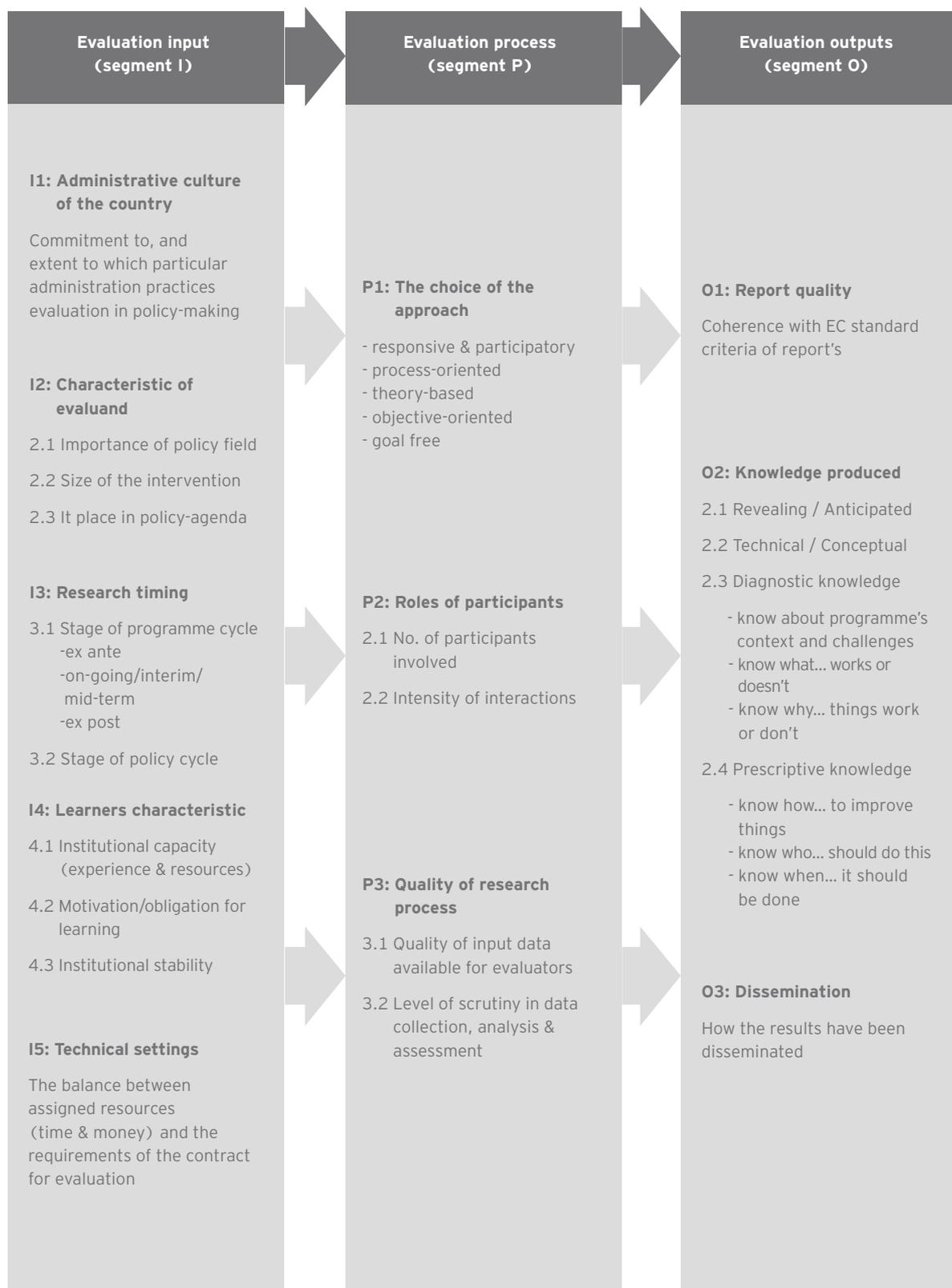
- interviews / questionnaires with key stakeholders of the evaluation process being assessed

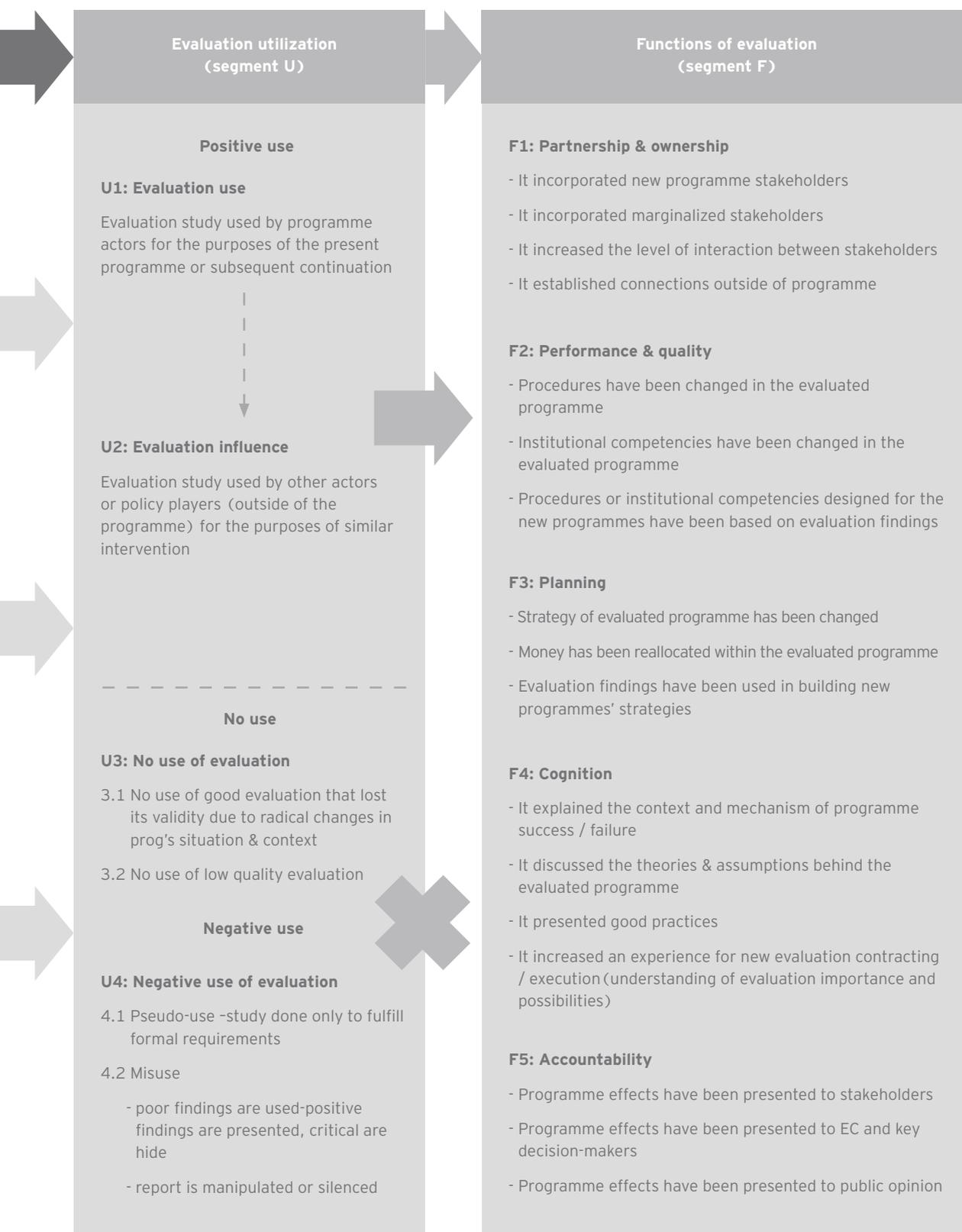
Phase 5: Analysis and conclusions

The objective of the final phase was to produce the analytical report.

The method of analysis followed the structure of the meta-model. For each case, analysis and conclusions were made on the relevance of the model. Furthermore, a comparative analysis between cases was carried out.

Annex 2. Analytical model





Usesource: Based on Olejniczak (2007), Cousins (2003), Mark & Henry (2004), Johnson (1998), Cousins (2004), Alkin & Taut (2003), Kirkhart (2000), Batterbury (2006)

Endnotes

- 1 It is worth noting that, in Polish official documents, evaluation is usually translated as “assessment” (pl: “ocena”) which highlights the controlling aspect of evaluation and its accountability perspective. As a result, most of the beneficiaries, as well as managing institutions, perceive evaluation as yet another form of central “control”.
- 2 Evaluation originated in the United States. In the 1960s, rapid increases in public spending, related to the implementation of Great Society programme, resulted in the urgent question of the real impact and utility of these interventions. Evaluation was expected to address this issue in a scientific manner, using a socio-economic toolbox of reasoning (Cronbach, Associates, 1981). The 1980s push for rationalization of the welfare state and the process of business-oriented reform of the public sector put evaluation into the everyday toolbox of modern public managers. By the mid-1990s, New Public Management became a basic model and global paradigm for public management. Its central philosophy of performance-oriented management strengthened the role of evaluation in rationalizing public activities. Public management models promoted by such international organizations as the OECD, World Bank or USAid further reinforced the position of evaluation. The emerging paradigm of New Public Governance, despite its holistic approach, emphasis on flexible networks and horizontal structures, keeps the core assumption of performance-oriented management (Osborne, 2006; Wolf, 2000). Thus, evaluation maintains a crucial role.
- 3 These are instruments for supporting social and economic restructuring across the EU at both regional and national levels. The Structural Funds have been available in priority ‘Objective’ areas (i.e. Objective 1, 2 and 3) and in support of Community Initiatives (i.e. Equal, Interreg III, Urban II And Leader +).
- 4 The PhD thesis by Olejniczak (2007) investigated the use of evaluation in INTERREG III Programmes (that is outside of Polish public sector), while the SAPER project performed in Małopolskie focused more on indicators for studies’ quality assessment and public institutions abilities to synthesize results from different reports (Górniak, 2005).
- 5 The issues of external, internal validity, construct validity and reliability are usually discussed in the context of quasi-experimental designs. For an extensive debate on these issues see: Bingham, Felbinger, 2002; Shadish et al., 2001
- 6 Weiss & Bucuvalas (1980) state that users on evaluation perform two simple tests. First is the truth test (Can I rely on this research? Does it provide me with strong arguments?). Second is the utility test (Does it show me clear directions? Does it give me recommendations and alternatives?).

- 7 Stern (2008:5) labels this CMO (Context + Mechanism = Outcome) while others (Frechtling, 2007) simply call it "logic models" (that consisted 4 main components: inputs activities outputs outcomes).
- 8 The main sources of inspiration were Cousins & Leithwood (1986) and Johnson (1998).
- 9 In this we draw heavily on the works of Cousins (2003), Mark & Henry (2004) and Owen (2007), adapting them to the context of EU programmes (2000, 2004b).
- 10 Ex post and on-going evaluations will be performed in the upcoming months and years, while next generation of ex-ante studies is predicted in 2012.
- 11 In the distinction between qualitative and quantitative research strategy we follow the interpretation by Bryman (2004:19-21). This distinction is not related to mere differences in research methods. It has three aspects: the role of theory in relation to research (inductive, generation of theory in qualitative strategy vs. deductive, testing theory in quantitative strategy), epistemological orientation (interpretivism in the first case vs. natural science model - positivism in the second case) and ontological orientation (constructionism vs. objectivism)
- 12 For an extensive discussion of these approaches see: (Bingham, Felbinger, 2002; Russ-eft, Preskill, 2001:147-175; Shadish et al., 2001)
- 13 Focus on process does not exclude considerations of causal relations. However, the procedure for identifying e.g. reasons for management bottleneck is less scientifically formal.
- 14 To put it simply, evaluation is not based on programme goals but on what programme actually has done. It is oriented towards needs of the target clients, not on what managers and developers assert to be their intentions (Mathison, 2005:171). This approach has been developed by M.Scriven.
- 15 At this point one important distinction has to be made: between "interactive" and "participatory" terms (Mathison, 2005:292). "Interactivity" is just the term that describes the basic relationship between evaluator and CA. Starting from the mid-70s, there has been a common agreement that in good evaluation (despite the approach used) an evaluator should routinely interact with programme staff and stakeholders while determining their concerns, discussing methods and framing questions. However being "interactive" evaluators still retains complete control of decision-making for the study (regardless of the data-collection methods). Thus, "interactivity" is desired and possible in all types of the above listed approaches (including those focused on effects and using quantitative methods). Contrary to this, the term "participatory" describes only one of the strategies that goes well beyond "interactivity". Here evaluator allows stakeholders to join the decision-making. He takes the role of a partner, facilitator and coach, shares the responsibility or even just steps back and provides support for the judgments made by the stakeholders themselves.

- 16 The step of making suggestions - giving prescriptive knowledge (the utilitarian orientation) is in fact a key distinctive feature between evaluation and scientific research, even applied research.
- 17 "Instrumental use" is a transposition of evaluation findings onto direct decisions, actions of the programme operators (e.g. changes in the programme on the basis of recommendations). "Conceptual use" is an influence of evaluation findings onto the changes in thinking of decision-makers, programme operators, wider public about the programme, similar implemented interventions or social processes in general. Finally, "symbolic use" is use of evaluation as the official justification for the decision already taken on the basis of intuition or political game, social interests, or decision-makers experience (Alkin, Taut, 2003:5; Leviton, Hughes, 1981; Mark, Henry, 2004:36; Rich, 1977; Shulha, Cousins, 1997:197; Weiss et al., 2005:13-14)

Bibliography

- Agrotec, Ierigż, Asa (2003) *Średniookresowa ocena (ewaluacja) realizacji programu SAPRAD w Polsce w latach 2000-2003 (PL-7-05/00)*, Komisja Europejska, Ministerstwo Rolnictwa i Rozwoju Wsi, Warszawa
- Alkin M. C. (2005) Utilization of evaluation, in Mathison S. (red.) *Encyclopedia of evaluation*, ss. 143, 434-6. Sage, Thousand Oaks, Calif. ; London.
- Alkin M. C., Christie c. A., Rose M. (2006) Communicating evaluation, w Shaw I., Greene j. C., Mark M. M. (red.) *Handbook of evaluation: policies, programs and practices*, ss. 384-403. SAGE, London.
- Alkin M. C., Coyle K. (1988) Thoughts on evaluation utilization, misutilization and non-utilization, *Studies in Educational Evaluation* 14(3), 331-40.
- Alkin M. C., Taut S. M. (2003) Unbundling evaluation use, studies in *Educational Evaluation* 29, 1-12.
- Bachtler J. (2001) Quod erat demonstrandum? Ewaluacja polityki regionalnej, *Studia Regionalne i Lokalne* 4(7).
- Bachtler J., Wren C. (2006) Evaluation of European Union Cohesion Policy: Research Questions and Policy Challenges, *Regional Studies* 40(2), pp.143-53.
- Batterbury S. (2006) Principles and purposes of european union cohesion policy evaluation, *Regional Studies* 40(2), pp.179-88.
- Bingham R. D., Felbinger C. L. (2002) *Evaluation in practice: a methodological approach*. Seven Bridges Press, New York ; London.
- Boyle R., Lemaire D. (1998) Building effective evaluation capacity: lessons from practice. Transaction Publishers, New Brunswick, NJ.
- Bryman A. (2004) *Social research methods*. Oxford University Press, Oxford.
- Casavola P., Tagle L. (2003) Building Capacity for evaluation: Lessons from Italy, *Fifth European Conference on Evaluation of the Structural Funds "Challenges for Evaluation in an Enlarged Europe"*. European Evaluation Society, Budapest, 26-27 June 2003.
- Chelimsky E. (1985) *Program evaluation: patterns and directions*. American Society for Public Administration, Washington, D.C.
- Chelimsky E. (1997) The Coming Transformation in Evaluation, w Chelimsky E., Shadish W. R. (red.) *Evaluation for the 21st century: a handbook*. Sage, Thousand Oaks, Calif.; London.
- Christie C. A. (2007) The use of evaluation - latest empirical inside from US, *Seminars in the framework of Academy of Evaluation - Koźmiński School of Management*, Warsaw, December 2007.
- Council of the European Economic Community (1988a) *Council Regulation (EEC) No 2052/88 of 24 June 1988 on the tasks of the Structural Funds and their effectiveness and on coordination of their activities between themselves and with the operations of the European Investment Bank and the other existing financial instruments*, 2052/1988, Official Journal of the European Communities L 185 , 15/07/1988, p.0009-0020
- Council of the European Economic Community (1988b) *Council Regulation (EEC) No 4253/88 of 19 December 1988 laying down provisions for*

Bibliography

- implementing regulation (EEC) no 2052/88 as regards coordination of the activities of the different structural funds between themselves and with the operations of the European Investment Bank and the other existing financial instruments*, 4253/1988, Official Journal of the European Communities L 374, 31/12/1988, p.1.
- Council of the European Union (2006) *Council Regulation (EC) No 1083/2006 of 11 July 2006 laying down General provisions on the European Regional Development Fund, the European Social Fund and the Cohesion Fund*, 1083/2006, OJ L210, 31.7.2006
- Cousins B. J. (2003) Utilization effects of participatory evaluation, w Kellaghan T., Stufflebeam D. L., Wingate L. A. (red.) *International handbook of educational evaluation*, ss. 2 V. (1x, 1061 p.). Kluwer Academic, Dordrecht ; London.
- Cousins B. J. (2004) Commentary: minimizing evaluation misuse as principled practice, *American Journal of Evaluation* 25(3), 391-7.
- Cousins B. J., Leithwood K. A. (1986) Current empirical research in evaluation utilization, *Review of Educational Research* 56(3), 331-64.
- Cousins J. B., Whitmore E. (1998) *Framing participatory evaluation*, *New Directions for Evaluation* 80, 5-23.
- Cronbach L. J., Associates (1981) *Toward reform of program evaluation: Aims, methods, and institutional arrangements*. Jossey-Bass Publications, San Francisco.
- Dahler-Larsen P. (2005) Evaluation in Public Management, w Ferlie E., Lynn L. E., Pollitt C. (red.) *The Oxford handbook of public management*, ss. 615-39. Oxford University Press, Oxford.
- EMS (2004) *Interim Evaluation of the European Union Phare Programme Country: Poland Sector: Aid Co-ordination*, Europeaid, Brussels
- EPEC (2005) *Study on the use of evaluation results in the Commission. Final report*, European Policy Evaluation Consortium (EPEC), Paris
- Euroreg Warsaw University, Pricewaterhousecoppers PL, Irs (2006) *An ex post evaluation of the Phare Cross-Border Co-operation Programme Poland-Germany 1999-2001*, Implementing Authority for Phare-CBC Poland, Warsaw
- Eser T. W., Nussmueller E. (2006) Mid-term Evaluations of Community Initiatives under European Union Structural Funds: a process between accounting and common learning, *Regional Studies* 40(2), pp.249-58.
- European Commission (1989) *Guide to the Reform of the Community Structural Funds*, Office for Official Publications of the European Communities, Luxemburg,
- European commission (1999) *MEANS collection: evaluating socio-economic programmes*. Dg-regio, Office for Official Publications of the European Communities, Luxembourg.
- European Commission (2000) *Methodological Working Paper 3 - Indicators for Monitoring and Evaluation: An indicative methodology*, DG-REGIO, The New Programming Period 2000-2006 Methodological Working Papers Office for

- Official Publications of the European Communities, Luxembourg,
http://ec.europa.eu/regional_policy/sources/docoffic/working/sf2000_en.htm
- European Commission (2004a) *Evaluating EU activities - A Practical Guide for the Commission Services*, DG-BUDGET, Office for Official Publications of the European Communities, Luxembourg,
- European Commission (2004b) *Project Cycle Management Guidelines, vol 1*, Europeaid, Office for Official Publications of the European Communities, Luxembourg,
http://ec.europa.eu/europeaid/evaluation/methodology/index_en.htm
- Flyvbjerg B. (2006) Five Misunderstandings About Case-Study Research, *Qualitative Inquiry* 12 (2), 219-45.
- Forss K., Rebien C. C., Carlsson J. (2002) Process use of evaluations. Types of use that precede lessons learned and feedback, *Evaluation* 8 (1), 29-45.
- Frechtling J. A. (2007) *Logic modeling methods in program evaluation*. Jossey-Bass ; [Chichester : John Wiley, distributor], San Francisco, Calif.
- Furubo J.-E., Rist R. C., Sandahl R. (2002) *International atlas of evaluation*. Transaction Publishers, New Brunswick, U.S.A.
- Górniak J. (2005) Ewaluacja jako czynnik doskonalenia zarządzania strategicznego w administracji publicznej: projekt SAPER, *I konferencja ewaluacyjna - Ewaluacja programów o charakterze społeczno-gospodarczym finansowanych z funduszy strukturalnych*, ss. 83-101. Ministerstwo Gospodarki i Pracy & Polska Agencja Rozwoju Przedsiębiorczości, Warszawa, 16 września 2005.
- Hausner J. (2007) Czy należy pomagać regionom słabo rozwiniętym?, *Symposium Regional Studies Association - Sekcja Polska*, Warszawa, 18.04.2007.
- Heinrich C. J. (2003) Measuring Public Sector Performance, W Peters G., Pierre J. (red.) *Handbook of Public Administration*, ss. 25-37. Sage Publications, London.
- Henry G. T. (2001) How modern democracies are shaping evaluation and the emerging challenges for evaluation, *American Journal of Evaluation* 22 (3), 419-29.
- Holzinger E. (2001) *6 Jahre Programm-Evaluation - eine Bilanz*. ÖIR, Wien.
- Indygo (2004) Evaluation of PHARE PL 9811 Initiative Programme, PARP, Warsaw
- Izquierdo B. (2006) The Use of Evaluation in multi-actor Context: A Participatory Experience in the Basque Country, *Easy-Eco conference improving the Quality of Sustainable Development Projects*. Easy-Eco, Saarland, October 2006.
- Jakoby H. (2006) Evaluation as part of the regional policy life cycle: the example of North Rhine-Westphalia, *Regional Studies* 40 (2), 281-4.
- Johnson B. (1998) Towards a theoretical model of evaluation utilization, *Evaluation and Program Planning* 21, 93-110.
- King J. A. (1998) Making sense of participatory evaluation practice, *New Directions for Evaluation* 80 (Winter), 7-92.

Bibliography

- Kirkhart K. E. (2000) Reconceptualizing Evaluation Use: An Integrated Theory of Influence, *New Directions for Evaluation* 88, 5-25.
- Klimczak Tomasz, Pylak K., Podyma D. (2006) *Badanie przyczyn zróżnicowania w poziomie płatności realizowanych w ramach ZPORR na poziomie województw*, WYG International dla Ministerstwa Rozwoju Regionalnego, Warszawa
- Krajowa Jednostka Oceny (2005) Rynek ewaluacji w Polsce - diagnoza stanu i perspektyw rozwoju, *I konferencja ewaluacyjna - ewaluacja programów o charakterze społeczno-gospodarczym finansowanych z funduszy strukturalnych*, ss. 31-45. Ministerstwo Gospodarki i Pracy & Polska Agencja Rozwoju Przedsiębiorczości, Warszawa, 16 września 2005.
- Leviton L. C. (2003) Evaluation use: advances, challenges and applications, *American Journal of Evaluation* 24(4), 525-35.
- Leviton L. C., Hughes E. F. X. (1981) Research on the utilization of evaluations: a review and synthesis, *Evaluation Review* 5(4), 525-48.
- Lewis B. (1997) The west and the Middle East, *Foreign Affairs* 76(1), 114-30.
- Lundvall B.-Å., Johnson B. (1994) The Learning Economy, *Journal of Industry Studies* 1(2), 23-42.
- Mark M. M., Henry G. T. (2004) The mechanisms and outcomes of evaluation influence, *Evaluation* 10(1), 35-57.
- Mark M. M., Henry G. T., Julnes G. (2000) *Evaluation: an integrated framework for understanding, guiding, and improving policies and programs*. Jossey-Bass, San Francisco.
- Mathison s. (Red.) (2005) *Encyclopedia of evaluation*. Sage, Thousand Oaks, Calif. ; London.
- Morris I. L., Freeman M. E., Fitz-Gibbon c. T. (1987) *How to communicate evaluation findings*. Sage, Newbury Park ; London.
- Nutley S., Walter I., Davies H. T. O. (2003) From Knowing to Doing. A Framework for Understanding the Evidence-Into-Practice Agenda, *Evaluation* 9(2), 125-48.
- OECD (1999) *Improving Evaluation Practices. Best Practice Guidelines for Evaluation and Background Paper*, PUMA, OECD,
- Olejniczak K. (2002) Towards the evaluation culture in Poland - experiences and prospects, *Annual Conference of Regional Studies Association "Evaluation and EU Regional Policy"*. Regional Studies Association, Aix en Province, June 2002.
- Olejniczak K. (2007) Ewaluacja w zarządzaniu programami INTERREG, rozprawa doktorska, Wydział Zarządzania, Wyższa Szkoła Przedsiębiorczości i Zarządzania im. L.Koźmińskiego, Warszawa.
- Olejniczak K. (2008) *Mechanizmy wykorzystania ewaluacji. Studium ewaluacji średniookresowych INTERREG III*. Wydawnictwo Naukowe Scholar, Warszawa.
- Osborne S. S. (2006) The New Public Governance?, *Public management review* 8(3), 377.
- Owen J. M. (2007) *Program evaluation: forms and approaches. 3rd edition*. The Guilford Press, London.
- Owen J. M., Rogers P. J. (1999) *Program evaluation: forms and approaches*. SAGE, London.

- Patton M. Q. (1997) *Utilization-focused evaluation: the new century text*. Sage Publications, Thousand Oaks, Calif.
- Pollitt C. (1998) Evaluation in Europe. Boom or bubble?, *Evaluation* 4(2), 214-24.
- Preskill H. S., Catsambas T. T. (2006) *Reframing evaluation through appreciative inquiry*. SAGE, Thousand Oaks, Calif. ; London.
- Raines P., Taylor S. (2002) Mid-term Evaluation of the 2000-06 Structural Fund Programmes, *IQ-Net Thematic Papers* 11(2).
- Rich R. F. (1977) Uses of social science information by federal bureaucrats: knowledge for action versus knowledge for understanding, w Weiss C. H. (red.) *Using social research in public policy making*, ss. 199-211. D.C. Heath, Lexington, Mass.
- Rist R. C. (1990) *Program evaluation and the management of government: patterns and prospects across eight nations*. Transaction Publishers, New Brunswick, N.J. ; London.
- Rossi P. H., Lipsey M. W., Freeman H. E. (2004) *Evaluation : a systematic approach*. Sage, Thousand Oaks, Calif. ; London.
- Russ-eft D. F., Preskill H. S. (2001) *Evaluation in organizations: a systematic approach to enhancing learning, performance, and change*. Perseus, Reading, Mass.; Oxford.
- Shadish W. R., Cook T. D., Campbell D. T. (2001) *Experimental and quasi-experimental designs for generalized causal inference*. Houghton Mifflin, Boston.
- Shulha L. M., Cousins B. J. (1997) Evaluation Use: Theory, Research, and Practice Since 1986, *Evaluation Practice* 18(3), 195-208.
- Stern E. (2005) What do we know about the utilization of evaluation?, *I konferencja ewaluacyjna - Ewaluacja programów o charakterze społeczno-gospodarczym finansowanych z funduszy strukturalnych*, ss. 7-29. Ministerstwo Gospodarki i Pracy & Polska Agencja Rozwoju Przedsiębiorczości, Warszawa, 16 września 2005.
- Stern E. (2008) Editorial, *Evaluation* 14(1), 5-7.
- Strohmeier G., Holzinger E. (2006) Die Koordinations- und Arbeitsplattform KAP-EVA: Ein Prozess in der Evaluierung der Strukturfonds-Programme in Österreich,, *OIR papers*.
- Szlachta J. (2006) Polskie doświadczenia w zakresie ewaluacji, *II Konferencja Ewaluacyjna - Ewaluacja ex-ante jako narzędzie doskonalenia zarządzania strategicznego w administracji publicznej*. Ministerstwo Rozwoju Regionalnego & Polska Agencja Rozwoju Przedsiębiorczości, Warszawa, 22 listopada 2006.
- Tavistock Institute, GHK, IRS (2003) *The Evaluation of Socio-Economic Development. The Evalsed Guide (1st edition)*, www.evaled.com, downloaded 21.08.2006
- Toulemonde J. (2000) Evaluation culture(s) in Europe: differences and convergence between national practices, *The Danish Evaluation Society Annual Conference*.

Bibliography

- Weiss C. H., Bucuvalas M. J. (1980) Truth Tests and Utility Tests: decision-makers' frame of reference for social science research, *American Journal of Sociological Review* 45 (April), 302-13.
- Weiss C. H., Murphy-graham E., Birkeland S. (2005) An Alternate Route to Policy Influence How Evaluations Affect D.A.R.E., *American Journal of Evaluation* 26(1), 12-30.
- Williams K., De Laat B., Stern E. (2002) *The Use of Evaluation in the Commission Services. Final Report*, Technopolis France, Paris
- Wolf A. (2000) Trends in public administration - a practitioner's view, *International Review of Administrative Sciences* 66, 689-96.



Authors



Martin Ferry

Martin Ferry is a Research Fellow at the European Policies Research Centre, University of Strathclyde. He holds a PhD in East-Central European Politics, University of Glasgow. His main research interests are economic and regional development and policy in Central and Eastern Europe, management and implementation of EU Cohesion policy, evaluation of EU Cohesion policy, regional development in Poland and the United Kingdom and demographic ageing and public policy.

[martin.ferry@strath.ac.uk]



Karol Olejniczak

Karol Olejniczak is a Research Fellow at the Centre for European Regional and Local Studies at University of Warsaw. He holds a PhD in Management obtained at the Leon Koźminski Academy in Warsaw. He also studied at EUROREG - University of Warsaw and European Institute of Public Administration - Maastricht. In his research he focuses on the themes of evaluation of public programmes, public management, regional policy and territorial cooperation.

[k.olejniczak@uw.edu.pl]



SPRAWNE PAŃSTWO
PROGRAM *ERNST & YOUNG*

Better Government Programme
Ernst & Young Poland
Rondo ONZ 1
00-124 Warszawa
www.bettergovernment.pl