

**Ewaluacja programów
o charakterze społeczno – gospodarczym
finansowanych z funduszy strukturalnych**

I Konferencja Ewaluacyjna
Warszawa 16 września 2005 r.

Ministerstwo Gospodarki i Pracy
Polska Agencja Rozwoju Przedsiębiorczości

Redakcja i korekta:

Polska Agencja Rozwoju Przedsiębiorczości
Zespół Programowania i Ewaluacji

Agnieszka Haber
Jacek Pokorski
Maciej Szalaj
Rafał Trzciniński

Wydawca:

Ministerstwo Gospodarki i Pracy
Departament Koordynacji Polityki Strukturalnej
Instytucja Zarządzająca Podstawami Wsparcia Wspólnoty
Pl. Trzech Krzyży 3/5
00-507 Warszawa
tel. (+48 22) 693-43-03, fax (+48 22) 693 40 87

Projekt okładki:

Media-Art

Druk:

Epicentrum Sp. z o.o.



Publikacja współfinansowana ze środków
Europejskiego Funduszu Rozwoju Regionalnego
w ramach Programu Operacyjnego Pomoc Techniczna



Spis treści:

1. Elliot Stern - What do we know about the utilization of evaluation?.....	7
2. Rynek ewaluacji w Polsce – diagnoza stanu i perspektywy rozwoju.....	31
Europejskie praktyki i rozwiązania w zakresie ewaluacji programów społeczno-gospodarczych sektora publicznego.....	47
3. David Hegarty - Ewaluacja funduszy strukturalnych w Irlandii.....	49
4. Sandra Speer - Ewaluacja programów społeczno-gospodarczych – doświadczenia niemieckie.....	65
Zarządzanie oparte na wiedzy – wykorzystanie wyników ewaluacji w procesie zarządzania.....	81
5. Jarosław Górniak - Ewaluacja jako czynnik doskonalenia zarządzania strategicznego w administracji publicznej: projekt SAPER.....	83
Budowanie potencjału ewaluacyjnego w Polsce.....	103
6. Olivier Rouland - Budowanie potencjału ewaluacyjnego w Unii Europejskiej: wskazówki dla Polski.....	105
7. Leszek Korporowicz - Tworzenie potencjału ewaluacyjnego.....	113
8. Elżbieta Opałka - Plan działań ewaluacyjnych NPR/PWW na lata 2004-2006.....	127

Szanowni Państwo,

Mamy przyjemność przedstawić publikację poświęconą ewaluacji programów o charakterze społeczno – gospodarczym. Stanowi ona zbiór artykułów przygotowanych przez krajowych i zagranicznych ekspertów – prelegentów I Konferencji Ewaluacyjnej **„Ewaluacja programów o charakterze społeczno – gospodarczym finansowanych z funduszy strukturalnych”**.

W krajach Unii Europejskiej badania ewaluacyjne są uznawane za nieodłączny element wdrażania skutecznej pomocy publicznej. Rosnące znaczenie ewaluacji jako narzędzia planowania, projektowania i zarządzania programami pomocowymi, oraz jako instrumentu wspomagającego efektywne wykorzystanie funduszy pomocowych stało się impulsem dla niniejszego przedsięwzięcia.

Publikacja, którą trzymacie Państwo w rękach, podzielona jest na 3 części, zgodnie z chronologią i tematyką poszczególnych sesji konferencji. Kluczowe zagadnienia w niej poruszane związane są przede wszystkim z wykorzystaniem wyników ewaluacji programów w procesie zarządzania interwencją społeczno-gospodarczą. Główny nacisk został jednak położony na ewaluację programów finansowanych ze środków funduszy strukturalnych.

Chcielibyśmy zwrócić Państwa uwagę na bardzo istotną kwestię potrzeby budowania w Polsce potencjału ewaluacyjnego i szeroko rozumianej kultury ewaluacyjnej, której zadaniem jest podnoszenie skuteczności i efektywności realizowanych programów, finansowanych z funduszy strukturalnych Unii Europejskiej.

Mamy nadzieję, że zarówno tematyka publikacji, jak i sposób prezentacji poruszanych w niej zagadnień znajdzie Państwa uznanie. Jednocześnie chcielibyśmy wyrazić wdzięczność za ewentualne uwagi i komentarze, które będą dla nas cennym źródłem inspiracji przy redagowaniu jej kolejnych edycji.

Organizatorzy I Konferencji Ewaluacyjnej

Ministerstwo Gospodarki i Pracy
Polska Agencja Rozwoju Przedsiębiorczości

Elliot Stern

What do we know about the utilization of evaluation?¹

Based on the available literature on evaluation use²

Introduction

What follows might be better regarded as a literature mapping exercise rather than a literature review. However, sufficient material has been considered to begin to summarise what is known about evaluation use from available published sources.

We have drawn together material from three main sources:

- Academic and scholarly work in evaluation and related research.
- Professional guidance and advice on what evaluators and commissioners of evaluation should do.
- Examples drawn from the world of practice across different countries and institutions.

The sources reviewed fall into a number of broad categories, these are:

- Evaluation literature on the utilisation of evaluation.
- Policy studies about the institutionalisation of evaluation within governments and public administrations.
- Guidelines on evaluation practice and standards.
- Examples of specific evaluation arrangements that are relevant for use in different public administrations – national and international.

Literature on the utilisation of evaluation

There has been debate about the use of evaluation for over twenty five years. This is primarily associated with two main scholars: Carol Weiss and Michael Quinn Patton. In the mid-1970's Carol Weiss began to focus on evaluation use, criticising simplistic notions of

¹ Copyright reserved Elliot Stern 2003 and 2005: Can be cited if excerpts and quotations are acknowledged.

² Part of this review was supported by European Commission, DG Budget. Their support is gratefully acknowledged. All interpretations and conclusions are the responsibility of the author.

instrumental use. Based on empirical studies on how policy makers use evaluation, Weiss has been associated with a complex understanding of decision making and policy making in which evaluation findings are internalised, selectively used and rarely lead directly to specific decisions or changes in policy. This theory, which is sometimes labelled 'an enlightenment view of evaluation', takes a long term incremental view about the way evaluation findings feed through to policy making. In this regard she is close to Peter Rossi's understanding of the conceptual use of evaluation i.e. 'the use of evaluations to influence thinking about issues in a general way' (Rossi & Freeman 1985). She challenges the rational model of decision making and bases her conclusions on studies of how organisations actually work. She effectively argues for cumulative learning across many evaluations rather than direct use from particular evaluations. In some of her more recent work (Weiss, 1999) she associates herself with arguments of Sabatier and Jenkins-Smith (1993) about the importance of policy forums that bring together academics and policy makers and can act as a vehicle for the absorption of evaluation findings. She nonetheless remains true to her earlier arguments that evaluators should never expect that their inputs can override political agendas and administrative necessities which may push decisions in quite different directions from the recommendations of evaluators.

Michael Quinn Patton is more committed to the instrumental purposes of evaluation. 'By utilization I mean intended use by intended users' (Patton 1986, 1997). He places considerable emphasis on understanding the priorities of decision makers, engaging with them, encouraging them to own evaluations and their results in order to enhance use. Furthermore, Patton tends to emphasise the example of individual decision makers, particular people and not the decision itself. Indeed he is interested in individual 'decision maker's cognitive style and logic'. Unlike Weiss he is less interested in organisational and administrative processes and more with the act of decision and the particular decision maker.

It has been commonly observed (see for example Alkin 1990) that an important determinant of the difference in perspective between Weiss and Patton is their respective fields of practice and study. Patton has worked mainly in local community and voluntary organisations, and where he has worked at public administrations they have also been at a local level. Weiss has been more occupied with large-scale national programmes (beginning for example with the 1970's War on Poverty in the US). This probably goes a long way to explaining the one's preoccupation with complex organisational processes and the other's preoccupation with individual decision makers and their decisions.

Although the above summarises, in a simplified way, the main debate between Weiss and Patton, both have contributed far more to our understanding of evaluation use. In particular Patton is probably the originator of the term 'process use' i.e. 'changes...that occur among those involved in evaluation as a result of learning that occurs during the evaluation process' (Patton 1997). Considerations of the process aspects of evaluation open up discussions also of different, more action orientated, views of evaluation purpose. For example, Patton also discusses 'intervention orientated evaluation', 'participatory evaluation' and 'empowerment evaluation'. These kinds of evaluation approaches are less relevant to the immediate concerns of this particular study, but do raise interesting methodological questions about how in the course of evaluations one can increase commitment and ownership by attending to the details of the evaluation process.

Weiss, especially in her early work, opens up another aspect of evaluation use, the importance of sound methodology and reliable data (Weiss & Bucuvalas 1980). Indeed in some of this early work Weiss anticipates some very contemporary discussions about evidence based policy making and the use of experimental methods. It should be acknowledged, however, that in later years as she investigated more carefully, the actual uses made of evaluation she shifted towards the position outlined above that we had to understand more about use in an organisational context. In her concern for sound methods and valid research, Weiss is more 'hands on' than many of her contemporaries thus Donald Campbell for example sees use for being essentially part of politics and of little concern for evaluators who should be driven by 'truth'.

In many ways we can perceive a pendulum swing between organisational and implementation concerns on the one hand and methodological and quality assurance concerns on the other, back and forth over the last twenty five years. The current interest in evidence based policy, the rediscovery of 'the gold standard' of experimental methods and randomised control trials (RCTs) and the 'realistic critique' of such trials and experiments, represents a renewed movement toward the methodological end of the pendulum. Thus the 'What Works' school (see Davies, Nutley and Smith, eds. (2000)) is pre-eminently concerned with the nature of research evidence and the need for rigorous methodology. The adherents of this school often regard RCTs as the most reliable form of primary evidence and systematic reviews and meta-analysis – which make quantitative estimates of the overall effects of interventions by aggregating primary studies - as the most reliable approach for secondary evidence. However, already in Davies, Nutley and Smith (op cit) the editors of this collection themselves begin to address questions of better understanding the policy process. This is taken further by Huw Davies and Sandra Nutley

(2002) when they move beyond data and findings to consider how evidence is disseminated and incorporated into practice. Thus these authors also follow the pendulum between methodological and implementation concerns.

However, the central components of the evidence based policy debate remain methodological. For the purposes of this literature mapping we would highlight, in particular, the debate around meta-analysis and systematic reviews (Gallo (1978), Pawson (2002)). The core of this debate is the importance of drawing together evidence from previous related initiatives before embarking on new initiatives. For some, narrative reviews that are essentially structured qualitative literature reviews drawing together lessons, is their preferred approach. For others, they look to quantifying precise effects by looking to quantitative and experimental studies along the lines of RCTs in medical drug trials. Of course, the possibility of undertaking such quantitative studies is limited to a subset of policy areas usually in the social welfare, social benefit and welfare to work area. Despite the methodological disagreements among the proponents of different approaches to such secondary analysis, there is a widespread consensus that brings together evidence from across many evaluations is an important guarantor of the reliability of findings.

The most recent attempt to revisit the issue of evaluation use in a comprehensive fashion, is a publication of the American Evaluation Association – *New Directions for Evaluation* ('The Expanding Scope of Evaluation Use', Caracelli & Preskill, eds (2000)). The initial chapter attempts a reconceptualisation of evaluation use in terms of 'an integrated theory of influence'. This largely brings together much of the work referred to above distinguishing for example between process use and results use, intended and unintended use and a policy cycle time dimension, i.e. immediate end of cycle and long term use. Although this chapter brings together various previous attempts to address evaluation use within a single framework, it does not radically change the broad parameters of the debate outlined previously.

There are two additional chapters in this collections which should be referred to for the purposes of this mapping exercise. In Chapter 2, Preskill and Torres consider 'the Learning Dimension of Evaluation Use'. These authors highlight, in particular, learning at an organisational level that occurs as part of the evaluation process. They draw on 'constructivist' theories of learning, in which learners themselves engage in an active process of interpretation and dialogue in order to construct their own meanings. They link this perspective with Lave and Wenger's (1991) notions of 'communities of practice' i.e.

groups of individuals working together, who are interdependent and whose tacit knowledge and problem solving capacities are integrated into their social and professional life. The authors suggest that learning from evaluations 'will most likely occur when evaluation is collaborative, is grounded in constructivist and transformational learning theories, and builds communities of evaluation practice.' The implications of this argument for evaluation use is to reinforce the importance of developing communities of both evaluators and users within organisations, if evaluation is to become part of an active learning process. It also steers us away from the narrow view of intended use that certainly informed the earlier work of Patton. Given that communities of practice will interpret and construct their own meanings using data and findings that they bring into their own context, the use of evaluation may not be as evaluators or the commissioners of evaluation originally intended. However, transforming such evaluative outputs and processes into their own organisational context still constitutes evaluation use.

Policy studies on the institutionalisation of evaluation

There is an extensive literature around the role of evaluation in policy systems. This literature is largely based in policy analysis and public administration rather than in evaluation per se. The preoccupations of this body of literature include organisational learning in the public sector, how to build evaluation into public administrations, the nature of effective evaluation capacity, institutionalising and professionalising evaluation and analyses of the policy process itself. This latter set of preoccupations is the closest 'cousin' to the earlier discussion of evaluation use. However, within this literature the starting point is the nature of the policy process which then moves on to considerations of evaluation practice whereas within the earlier cited literature, the starting point is usually evaluation practice which then moves on to considerations of policy making.

Organisational learning has become a common metaphor within studies of various kinds of organisations, although most of this work has centred on the private sector. Leeuw, Rist & Sonnichsen, eds. (1994) bring together a range of experiences from the public sector across Canada, Sweden, the Netherlands, Norway and the United States that seeks to apply the concept of organisational learning through evaluation to public sector institutions. Among the themes highlighted in this study is the role of internal evaluations within public bodies (see especially chapters by Sonnichsen & by Mayne). The authors highlight the important role that can be performed by internal evaluation offices and how these can influence the overall organisations' evaluation practice and learning. They

suggest that developing a 'double-loop learning' process that reports not only to programme managers but also to top management within government departments is an important part of the contribution that evaluation units can make³. However, these studies also suggest that whilst internal units can make an important contribution to organisational learning, they depend ultimately on what Sonnichsen calls 'a disposition towards critical self-examination'. His notion that 'self-reflection is crucial before organizations begin to learn' highlights the importance of creating a general evaluation culture within an organisation.

In the concluding chapter of this collection, Ray Rist identifies two sets of preconditions for learning from evaluation. The first set arises from the importance of fitting in with the policy cycle. This recognises that organisations need 'information at different phases of the policy cycle'. Synchronising evaluation outputs with different policy needs over time is seen as an important means of encouraging learning. The second set of preconditions for learning emphasises how information is transmitted and filtered within public organisations. Thus studies within this framework suggest that 'governmental organizations appear more receptive to information produced internally than that which comes from external sources'. This is especially so when the news is bad! Bad news is easier to receive from internal rather than external evaluators. Another precondition appears to be the credibility of the sources of information. This may on occasions favour internally generated evaluation findings, but may also depend on who are the sponsors or gatekeepers who bring information in to a public body. This can be seen as related to wider issues of relationship building and trust. The existence of such organisational attributes seems an important preconditions for receptivity to evaluation findings and to organisational learning within the Public Sector.

Another collection we have considered, focuses on evaluation capacity (Boyle & Lemaire, eds. 1999). Although some of the authors in this volume overlap with the Leeuw et al. collection referred to above, the concerns here are broader. Building evaluation capacity is seen in terms of 'national evaluation systems'. Thus evaluation capacity goes beyond the internal organisation of public bodies to include the location of evaluation in the executive or the legislature (e.g. Parliament) and broader issues of governance and

³ It should be note that this use of 'double-loop learning' is unusual, even though this citation is true to the quoted authors. From Gregory Bateson, who originated many of these concepts in the 1950s and 1960s through to more contemporary authors who have come to be identified with them, such as Donald Schon and Peter Senge, 'double loop learning' is usually about learning to learn, by understanding higher order rules that can be applied across contexts.

institutional arrangements. For the purposes of this literature mapping exercise a number of themes explored in this collection are particularly relevant.

One such theme is the design of evaluation functions and offices. Richard Sonnichsen considers the advantages and disadvantages of a 'centralized' and a 'decentralized' model of evaluation functions. He notes the potential for greater independence and credibility of centralized functions and their ability to develop strategic evaluation plans. This model has been favoured in Canada for example since the late 1970s. The 'downside' of centralized units and functions is also recognised. They are often seen as threatening by other units of administration with attendant resistance to change and potentially strained relationships.

Decentralised evaluation units are most likely to be intended to support decision making and programme effectiveness at a programme level. They are usually aligned with programme management and are hence less threatening. On the other hand issues of independence and bias can arise from their closeness to programme personnel. Another weakness may be lack of methodological skills in evaluation. However, the most important criticism raised by Sonnichsen is the possible lack of power that decentralised units have, especially where decisions about programme and policy futures are still made at a centralised level. Ultimately this debate resolves itself into one of evaluation purpose. Where the purpose is primarily to improve programme and policy implementation there appear to be strong arguments for a decentralised model which will also favour learning at a decentralised level. Where the purpose is primarily to support central strategies and policy making the argument for a centralised model appears to be stronger – here the learning would tend to occur centrally rather than at programme or policy division level. Issues of professional competence and skills acquisition appear to be stronger within a decentralised framework. This issue is further discussed in the same volume by Richard Boyle who discusses the Human Resources aspect of evaluation specialists and the professionalisation of evaluation. He highlights in particular, the importance of appropriate training courses and education curricular, the way expertise is deployed mixing different skills together and the importance of continuous in-service training. From the point of view of this review much of this discussion is generic rather than use focused. However, from a use perspective Boyle makes a strong argument for 'developing evaluation users'. Evaluation users also need to be trained and developed to become 'consumers'. In some ways this can be seen as one of the consequences also of introducing managing for results approaches in public service organisations. This, in Boyle's terms creates the link between the supply and demand sides of evaluation use.

The shift in the institutional practice of evaluation within the public sector from monitoring and management to accountability and performance is widely noted. (See especially the chapter by Per Oyvind Bastoe: *Linking Evaluation with Strategic planning, Budgeting, Monitoring and Auditing.*) One definite tendency that this and other studies demonstrate is the integration of evaluation and monitoring with various approaches to performance management.

Among the elements of this approach (following from an influential OECD paper in 1995) are:

- objective and target setting,
- management responsibility to implement against targets,
- the monitoring of performance,
- the feed-in of such performance into future policy making and programming , and,
- the provision of information to external parliamentary and audit committees for ex post review.

There is also disagreement about the need and benefit of linking audit and evaluation, which Bastoe also recognises, especially when implementing performance management systems. It is, for example, important also to consider 'how learning actually takes place in organizations'. This attention to learning is also the pre-occupation of other authors in this field of the institutionalisation of evaluation in the public sector.

For example, a quite different approach to analysing evaluation use in policy settings is suggested by Peter Van der Knaap in his article 'Policy Evaluation and Learning: Feedback, Enlightenment or Argumentation?' (1995). He challenges the traditional rational – objectivist model of policy evaluation favouring rather a constructivist view in which policy makers conduct dialogues about evaluation findings in order to reach their conclusions. Thus 'policy-making is conceived of as an ongoing dialogue, in which both governmental and societal actors contest their views on policy issues by exchanging arguments'. At heart this argument challenges the 'positivist idea that policy evaluators may provide the policy-maker with neutral or objective feedback information or recommendations'. Rather than enlighten the policy maker 'At best, the evaluator might contribute to the quality of policy discourse by entering the negotiations that compose the policy-making processes with informed arguments and a willingness to listen, argue, and persuade or be persuaded'. This shift from the rational to the argumentative is, according to Van der Knaap, a way to 'institutionalise policy orientated learning'. This is not to suggest that the evaluator is relieved of the responsibility to provide reliable information

and findings but needs also to supplement traditional analysis with material that will stimulate debate and allow different stakeholders to consider material presented from different perspectives. A similar logic informs a recent article by Ville Valovirta (2002). 'Rather than regarding evaluative information as indisputable knowledge, it is viewed as a collection of arguments, which can be debated, accepted and disputed'. According to the author utilisation of evaluation should be regarded as a process that runs through four stages:

1. Familiarisation with evaluation results and involvement in the evaluation process.
2. Interpretations based on expectations, assessments of the quality of research (truth test) and the feasibility of actions proposed or implied (utility test).
3. Argumentation in which 'individual interpretations are ... subject to collective deliberation, discussion, negotiation and decision making'.
4. Effects which may take the form of decisions and actions, new shared comprehensions and increased awareness; and increased or undermined legitimacy.

Within this perspective on evaluation use:

'Evaluations force people to present well-grounded arguments for refuting evaluation conclusions and recommendations. This opens up possibilities for new understandings to emerge (Page 77)'

An important distinction that Valovirta makes, is between the different contexts in which evaluation takes place. In particular he distinguishes between settings where there is a high level of consensus versus those with a high level of conflict; and settings where there is a low pressure for change versus those with a high pressure for change. He suggests that these contextual differences will determine the nature of the argumentation that takes place around evaluation findings.

Guidelines On Evaluation Practice And Standards

For many years there has been concern among evaluators about evaluation use. This has partly been fuelled by the academic debates reviewed above which have highlighted the problem of evaluations not being used. One approach to these concerns has been to produce guidelines usually for evaluators, but also for commissioners, that are intended to promote evaluation use. The earliest of these efforts have been produced by the American Evaluation Association and its precursor organisations, for example the Joint Committee on Standards for Educational Evaluation.

The Programme Evaluation Standards prepared as guidance to evaluators, particularly from a programme perspective, identifies standards under four main headings:

- Utility Standards
- Feasibility Standards
- Propriety Standards
- Accuracy Standards

For the purposes of this literature mapping exercise, utility standards which 'are intended to ensure that an evaluation will serve the information needs of intended users' are the most immediately relevant. These standards variously concern:

- being clear about stakeholders so that their needs can be addressed;
- ensuring the credibility of the evaluators so that their results are likely to be accepted;
- collecting relevant information from a broad range of sources as understood by clients and stakeholders;
- being clear about value judgement used to interpret findings;
- report clarity so that the information provided in reports is easily understood;
- reports that are well disseminated to intended users in a timely fashion; and,
- planning evaluations from the outset in a way that encourages follow through from stakeholders.

It is worth noting that the other standards also have implications for evaluation use. For example, Feasibility Standards include a sub-standard entitled *Political Viability*. This is concerned with obtaining the cooperation of different interest groups in order to limit bias or misapplication of results. Similarly the Propriety Standards include a sub-standard entitled *Service Orientation* which is designed to assist organisations to address and effectively serve the needs of a full range of targeted participants; and one entitled *Conflicts of Interest* which seeks to avoid compromising evaluations and their results. Accuracy Standards are also infused by notions of use, for example *Valid Information*, one of the sub-standards is justified in terms of assuring 'the interpretation arrived at is valid for the intended use'.

This set of standards has been widely imitated and adapted to different national contexts including, most recently, by the Deutsche Gesellschaft für Evaluation – DeGEval (German Evaluation Society). There are also discussions and plans in France and the UK to

develop standards. These discussions have however moved away from directly adopting the North American model.

Alongside the widespread adoption of the AEA Programme Evaluation Guidelines there has also been the emergence more recently of a set of guidelines for the Ethical Conduct of evaluations. These differ from programme guidelines, being more concerned with the ethical dilemmas that both commissioners and evaluators face in the course of the evaluation process. Such guidelines have been variously prepared by the Australasian Evaluation Society (AES), The Canadian Evaluation Society (CES) as well as the AEA. These Ethical Guidelines are relevant to a discussion about evaluation use because they too are concerned with the credibility of evaluations and also the feasibility of the evaluation task.

Both the CES and the AEA direct their attention at evaluators. Thus the CES guidelines under the three main headings of *Competence*, *Integrity* and *Accountability* are concerned that *evaluators* should be competent, act with integrity and be accountable. Similarly the AEA - under its various headings of Systematic Inquiry, Competence, Integrity/Honesty, Respect for People, Responsibilities for General and Public Welfare and Recommendation for Continued Work – also direct their attention to what evaluators ought to do.

The set of Ethical Guidelines which stand out from the others already referred to, are those produced by the Australasian Evaluation Society (AES). These are concerned with ethical behaviour and decision making among commissioners, users and teachers of evaluation as well as evaluators themselves. According to the AES the primary groups addressed by their guidelines are commissioners and evaluation teams or evaluators. In this regard also the AES diverges from some of the other guidelines referred to above. They acknowledge that evaluators often work in teams rather than mainly as individuals.

The AES *Guidelines for the Ethical Conduct of Evaluations* follows the evaluation cycle by grouping its guidance under three main headings:

- Commissioning anpreparing for an evaluation.
- Conducting an evaluation
- Reporting the evaluation results.

One view one can take about the AES guidelines is that they constitute a quality assurance framework for the entire evaluation process. These guidelines (which can be found at http://www.aes.asn.au/ethics_guidelines-1.pdf) are relevant to evaluation use because of the focus that they bring to:

- Questions of credibility;
- Shared expectations between evaluators and commissioners about what can be delivered through an evaluation;
- Strengthening the basis for evaluation judgements;
- Reducing conflicts during the course of the evaluation process;
- Ensuring balance and simplicity in the way reports are presented.

However it should be noted that the AES regards these guidelines as complementary to rather than as a substitute for other guidelines such as the Programme Evaluation Guidelines – indeed they encourage the use of the Guidelines for the Ethical Conduct for Evaluations jointly with the Programme Evaluation Standards.

There is a variety of other guidelines for evaluation that are relevant to this literature mapping exercise. For example, the Means Collection Volume 1 (Evaluation Design and Management) has a section on optimising the use of evaluation. This focuses on dissemination, distinguishing between both different communication channels (e.g. reports, synthesis, article, confidential note) and audiences (e.g. commissioners of the evaluation, steering groups, managers, European institutions, citizens and journalists). The Means Collection also briefly touches on different forms of use including for example:

- Raising levels of knowledge;
- Deciding on improvements; and
- Judging success and failure.

The Means Collection authors also acknowledge the ‘absence of a direct short-term link between recommendations and decisions’.

On the boundaries of the territory to be covered by this mapping exercise would be a wider literature on innovation and good practice. Thus in many studies that have been conducted at a European level there is a concern for dissemination of innovation, an interesting example of this is found in Simo Mannila et al, eds (2001). In this volume, Philip Sonderberg discusses ‘Mainstreaming, Innovation and Knowledge Networks’, the way in which innovations are spread is close to the problem of how to encourage the take-up of the results of evaluations. Sonderberg also focuses on knowledge networks (cf. Policy Communities and Communities of Practice) as a means of ensuring the dissemination and mainstreaming of new ideas. Arguably this broader literature on the diffusion of innovation could be regarded more generally as a source of insights into the use of evaluation.

Suffice it to say for the purposes of this review, much of this literature is consistent in its thinking with the literature on evaluation use outlined in this mapping exercise.

Examples Of Specific Evaluation Arrangements

In this section a number of specific high-profile examples of evaluation use in public administrations are identified. By way of introduction it should be noted that these examples are not evaluated here and as far as we are aware have not themselves been evaluated by others. Most of the publicly available source material is either uncritical (and sometimes verging on the propagandist), or is aspirational because the initiatives concerned are themselves at a very early stage. It is fair to say that none of these cases are without problems, and these problems are acknowledged below where information is in the public domain.

The World Bank has developed an elaborate evaluation process that seeks to link planning and resource allocation to results. The evaluation process relies heavily on a rating system, using criteria such as relevance, efficiency, effectiveness, sustainability, institutional development and the Bank and borrowers performance. This system involves a number of different methods which include the Bank's own corporate scorecard and ICR (Implementation Completion Report). This is filled in by Programme Managers and independently monitored by the Bank's Evaluation Department. This evaluation process is embedded in a results based management framework, and indeed a broader risk assessment and control environment.

In order to foster a learning culture, the Bank has initiated a knowledge management strategy based on networks intended to draw lessons across countries and regions and ensure that 'global best practices are used to shape country development objectives' (1998). Knowledge networks are made up of a number of different components including:

- Forming professional communities of practice;
- Developing an online management system ;
- Establishing help desks or advisory services;
- Making key statistics available online;
- Providing a directory of expertise;
- Tracking engagement or transactional information;
- Creating online dialogue spaces;
- Increasing external access to Bank knowledge.

The implementation of this system has been uneven, among the reasons for this have been the work pressure on staff and pre-existing 'information hoarding cultures' which have militated against the sharing of information across sectors and departments.

Similar difficulties have been found in linking planning and resource allocation to results. This has been somewhat improved by the gradual refinement of a results based management system within the bank. Thus in the latest Annual Report on Operations Evaluation (2000-2001), progress with knowledge management activities have been noted in a number of areas, in particular:

- The active role of the helpdesk which since 1999 has fielded 3000 enquiries about evaluation methodology, findings and recommendations. An interesting feature of this helpdesk is that it is accessed both by OED internal staff and external partners.
- Participation in evaluation conferences, sector weeks and knowledge fairs in order to diffuse evaluation findings and lessons as widely as possible.
- Collaboration across the Bank's different evaluation units in order to create 'an evaluation community of practice'. This provides a forum to exchange evaluation news, activity and products and there is now a shared web portal that collects together material developed by different evaluation units.
- Synthesising evaluation findings providing the basis for a 'Lessons Paper Series' examples of recent titles include, Recurrent Lessons in India, Utilization of Project Implementation Units, and the World Bank's Experience with Institutional Development.

One innovation in the Bank's evaluation practice which is of particular relevance in the context of this mapping project is the attempt to trace the influence of three of the past OED evaluations:

'the study reconstructs the circumstances under which experiences of use occur, what change they bring about, and what factors favor or impede the perception of usefulness associated with OED studies' (World Bank 2001).

These studies, which have not yet reported, were chosen to include cases where there was already some indication of utilisation as this was likely to generate greater understanding of the circumstances under which evaluation occurred.

The Netherlands has recently introduced a policy evaluation system that explicitly links budgets to operations and results. The new regulation 'Central government achievement

data and evaluation research' came into force in all Dutch government departments in January 2002. This regulation overseen by the Ministry of Finance is the result of an interdepartmental working group which had two aims:

'to ensure that the evaluation function within central government is sufficiently guaranteed, and;

to ensure that the policy information supplied within the framework of the departmental budget and the departmental annual report complies with the quality requirements applying to this information.

In order to achieve these targets, the regulation contains further conditions regarding:

- I. the integrated deployment of evaluation instruments;*
- II. the weighing up that needs to be made when making use of ex ante evaluation research;*
- III. the degree to which policy is covered by periodical evaluation research and its frequency (completeness and periodicity);*
- IV. the methodological-technical quality of evaluation instruments and the way in which policy information is obtained;*
- V. the way in which civil servants and top politicians are informed about the results of periodical evaluation research, and;*
- VI. the responsibilities within a department for proper implementation of the regulation.*

(Policy Evaluation in the Netherlands: Linking budgets to operations and results, Peter van der Knaap 2001))

The use of evaluation is a feature of much of the system that is now being introduced in the Netherlands. This includes:

- Gathering together information on what has been achieved and the effects of policies.
- Ex ante evaluations which compare two or more policy options systematically.
- Ex post evaluations regarding 'achieving goals, effectiveness and appropriateness'.

With particular regard to the use of evaluation results, the new regulation requires that civil servants and politicians are informed about results. However, departments are not obliged to follow a standard process, but rather to specify what their procedure is 'in harmony with the departmental cycles of budgeting, planning and control'. In the Netherlands, the debate around evaluation, budgeting and performance management are

also rooted in questions of legitimacy. As Van der Knaap argues this integration of evaluation and performance management: 'provides society with a clear answer to the question: what can we expect of the government. And this in turn helps to answer the question: how much are we willing to pay for this?'

The Canadian Government has a long history of results based management and the incorporation of evaluation within that governments management framework (see for example chapter by John Maine in Leeuw et al. 2000 discussed above). The latest statement of this policy was published in April 2001 and replaced an earlier version published in July 1994.

'The policy is based on three fundamental principles:

- *first, that achieving and accurately reporting on results is a primary responsibility of public service managers;*
- *second, that rigorous and objective evaluation is an important tool in helping managers to manage for results; and*
- *third, that departments with the support of the Treasury Board Secretariat are responsible to ensure that the rigour and discipline of evaluation are sufficiently deployed within their jurisdictions.'*

Treasury Board Manual (2001)

Within this policy framework, evaluation has two main purposes;'

- *to help managers design or improve the design of policies, programs and initiatives; and*
- *to provide, where appropriate, periodic assessments of policy or program effectiveness, of impacts both intended and unintended, and of alternative ways of achieving expected results.'*

The whole emphasis of this policy is centred around use, evaluation is seen as a support for managers and decision makers. The Canadian government sees itself as 'committed to becoming a learning organisation' and the Treasury Board Secretariat which oversees government activities is obliged to:

'support the practice of evaluation in departments by providing advice on best practices, setting standards, monitoring the evaluation capacity in departments and using the products of evaluation to inform decision-making at the centre'.

These policies are integrated into the job responsibilities of administrative managers within government departments thus deputy heads are expected to provide 'early warning of evaluation findings' that indicate major concerns. Departmental heads of evaluation communicate their findings that raise management concerns to senior management. Whilst departmental managers are obliged to 'incorporate approved evaluation findings and measures for improvement into priority setting, planning, reporting and decision-making processes'. Furthermore, this policy is reinforced by evaluation standards that have been adopted for the Government of Canada. These include standards of evaluation reports that are intended to ensure that senior managers and external readers can understand the issues, contains clear and actionable recommendations and an analysis of the exposure to risks associated with these recommendations.

The current government in the UK has also made considerable efforts to improve the policy making processes of central government, in particular the Cabinet Office has put forward a framework for 'Professional Policy Making for the Twenty First Century' (1999). This framework is concerned with 'delivering outcomes' and to that extent shares many assumptions with a results based management approach. A key part of this approach is the use of evidence which follows from a commitment to concentrate on what works, a slogan used by the Government has been 'what counts is what works'. The commitment to evidence based policy involves improving access to research based findings through many of the methods which have been referred to elsewhere in this review. For example, knowledge pools and 'hubs' are seen as a means of knowledge management which provides access not only to results of research but also to 'evidence held in the minds of both front-line staff in departments, agencies and local authorities and those to whom the policy is directed'. This approach continues to inform many of the initiative of the UK government including wide ranging reform of the Civil Service which is intended to 'deliver results of high quality and good value'.

Any system that seeks to encourage the use of evaluation results, process use of evaluations and related practices of information management, organisational learning and results based management will encounter problems. All the examples included in this section have such problems, but not because their conception is flawed. Rather it is because as the earlier parts of this mapping exercise have suggested, evaluation use is complex and embedded in organisational, policy, administrative and methodological realities. There is no quick fix, magic bullet or best practice solution that can be found and copied. This does not mean that there is nothing to learn from these cases, although the

similarity in many regards to existing Commission practice can be seen as an initial positive lesson.

Conclusions

Evaluation use has been a topic of concern within the evaluation and policy analysis communities for over 25 years. This overview of the available literature has therefore highlighted long standing as well as more recent debates. What we have uncovered also fits well with other parts of this study: many existing aspects of Commission practice – in its recognition of instrumental, process and more subtle ‘argumentation’ uses, in its attempts to link evaluation with decision making, in its approach to the organisation and institutionalisation of evaluation within and across DGs – accord with what we know are important determinants of successful evaluation use.

We can draw together some of the main relevant conclusions without being too sensitive to what are often hard fought debates, mainly because the protagonists in these debates often themselves recognise that their differences arise from differences in context and fields of interest. The main conclusions we would draw are as follows:

- Evaluation use understood as the direct ‘instrumental’ use or take-up of results, findings, conclusions and recommendations is unusual although it does occur. The likelihood of such direct take-up will be shaped by administrative and policy imperatives as well as the quality of evaluation reports.
- Evaluation use is more likely to be cumulative across evaluations or in the case of particular evaluations be transformed by argumentation: the process of challenge and counter challenge to evaluation evidence and interpretation. Evaluation is likely to be influential rather than the sole cause of subsequent action.
- Even where results are not used or only used indirectly, the ‘process’ of evaluation initiation and management with all their attendant opportunities for discussion and reflection can be useful by offering opportunities to exchange information, clarify thinking and develop frameworks. These indirect uses of evaluation require that evaluators provide inputs and arguments that evaluation users can feed into their debates and argumentations.
- Whatever the kinds of evaluation use are envisaged or occur, involving stakeholders alongside evaluators and administrators throughout the evaluation cycle is essential in order to build commitment to evaluation use.

- Sound methods are important. These should be suited to the evaluation object, context, purpose and questions. There are cycles in the advocacy of particular methods: one must conclude that in their place most are relevant. The contemporary interest in systematic reviews and the synthesis of evaluation findings is consistent with the consensus that most evaluation use occurs across several rather than any one single evaluation.
- To use evaluation findings and processes, administrations need to learn. In order to do so they need to build up their human resources and competencies in evaluation, create networks and 'communities of practice', institutionalise appropriately (sometimes in a centralised and sometimes in de-centralised ways - each have their strengths and weaknesses) and build up their knowledge management systems. A culture of evaluation, interpreted most commonly as a reflex of self critical openness, is widely seen as essential to organisational learning.
- The uses of evaluation are inextricably bound up with the different purposes of evaluation. Those who link evaluation closely to accountability, for example through results based management, take a different view of evaluation use than those who prioritise learning or knowledge production evaluation purposes. It could be concluded from some of the literature that these different purposes and their associated uses are difficult to combine in one system. This may be because of the contrary requirements of self criticism and openness required for learning and a more purposive and 'error-free' culture that is often associated with performance management.
- Standards appear to be one way in which quality is assured and use is made more likely. Many existing sets of standards in evaluation appear to be particular in their content and suited also to particular contexts. The most useful standards would address the needs of evaluators, administrations, and wider publics; would include technical and ethical components; and be integrated within the institutional logics where they were to be applied. Such standards would probably also include key principles that are known to encourage the diffusion of innovations.
- The brief review of actual cases of use strategies, suggests that many if not most administrations are facing some of the same problems of evaluation use as the European Commission and other European Institutions. However none appear to have solved the problem. It would appear that all the 'solutions' identified in this

section need to be managed, integrated and institutionalised and will need constant re-engineering.

References and Sources

- Alkin, Marvin C. (1990) *Debates on Evaluation*. SAGE.
- Barbier, Jean-Claude and Bernard Simonin (1997) 'European Social Programmes : Can Evaluation of Implementation Increase the Appropriateness of Findings ?' *Evaluation Vol 3, No. 4, pp391-407*. SAGE.
- Bemelmans-Videc, Marie-Louise, Ray C. Rist and Evert Vedung, eds (1998) *Carrots, Sticks & Sermons*. Transaction Publishers.
- Boyle, Richard and Donald Lemaire, eds (1999) *Building Effective Evaluation Capacity : Lessons from Practice*. Transaction Publishers.
- Canadian Treasury Board (2001) *Treasury Board Manual 'Review' Volume*. Canadian Treasury Board.
- Caracelli, Valerie J. and Hallie Preskill, eds (2000) *The Expanding Scope of Evaluation Use*. Jossey-Bass.
- Davies, H. and S. Nutley (2002) *Discussion Paper 2 Evidence-Based Policy and Practice: Moving from Rhetoric to Reality*. University of St. Andrews Research Unit for Research Utilisation.
- Davies, Huw T.O., Sandra M. Nutley and Peter C. Smith, eds (2000) *What Works?. The Policy Press*.
- Department of Transport, Local Government and the Regions (2000) *Evaluation Notes*. UK DTLR.
- Deutsche Gesellschaft für Evaluation e.V. (2002) *Standards für Evaluation*. Die Deutsche Bibliothek.
- Duffy, M. (2000) *HEBS Commissioned Research Reports: Style Guidelines*. HEBS.
- Efficiency & Effectiveness Group (1999) 'Learning Lessons: The Role of Policy and Programme Evaluation' *Modernising Government White Paper (UK)* . Efficiency & Effectiveness Group.

- Efficiency & Effectiveness Group (1999) *Modernising Government White Paper, Learning Lessons : The Role of Policy and Programme Evaluation*. EEG.
- European Commission – EuropeAid (2001) *Manual Project Cycle Management*. European Commission.
- European Commission Budgetary Overview and Evaluation Unit (1997) *Evaluating EU Expenditure Programmes*. European Commission.
- Forss, Kim, Claus C. Rebien and Jerker Carlsson (2002) 'Process Use of Evaluations: Types of Use that Precede Lessons Learned and Feedback' *Evaluation Vol 8, No. 1, pp29-45*. SAGE.
- Fulbright-Anderson, Karen, Anne C. Kubish and James P. Connell, eds (1998) *New Approaches to Evaluating Community Initiatives Volume 2, Theory, Measurement and Analysis*. The Aspen Institute.
- Gallo, P.S (1978) 'Meta-analysis – a Mixed Metaphor' *American Psychologist 33, pp515-517*. Gautam, Madhur (2002) *2000-2001 Annual Report on Operations Evaluation*. World Bank Operations Evaluation Department.
- Guy, K., J. Clark, K.Balazs, J. Stroyan and E. Arnold (1998) *Strategic Options for the Evaluation of the R&D Programmes of the European Union*. Technopolis Group.
- Hedges, A (1994) *Commissioning Social Research: A Good Practice Guide*. Social Research Association.
- Horton, D., R. Mackay, A. Andersen and L. Dupleich (2000) *Evaluating Capacity Development in Planning, Monitoring and Evaluation: A Case From Agricultural Research*. ISNAR.
- IST Monitoring Panel (2000) *Report of the 2000 IST Monitoring Panel*. DG Information Society.
- Lave, J. and Wenger, E. (1991) *Situated Learning : Legitimate Peripheral Participation*. Cambridge University Press.
- Leeuw, Frans L., Ray C. Rist and Richard C. Sonnichsen, eds (1994) *Can Governments Learn?*. Transaction Publishers.
- Macintyre, Sally, Iain Chalmers, Richard Horton and Richard Smith (2001) 'Using Evidence to Inform Health Policy : Case Study' *BMJ Vol 322, pp 222-322*. BMJ.

Mannila, S., M. Ala-Kauhaluoma and S. Valjakka, eds (2001) *Good Practice in Finding Good Practice: International Workshop Evaluation*. Rehabilitation Foundation.

Means Collection (1999) *Evaluating Socio-economic Programmes*. European Commission.

Monitoring and Evaluation Department of MSF (2000) *The Utilisation of Evaluation Recommendations in Medecins sans Frontieres – Holland: A Study of 10 Evaluation Reports 1997-1999, Evaluation Series No. 14*. MSF.

Norwegian Ministry of Foreign Affairs (1999) *Evaluation Activities in the Norwegian Ministry of Foreign Affairs, Annual Report 1998*. Norwegian Ministry of Foreign Affairs.

Norwegian Ministry of Foreign Affairs (2000) *Evaluation Activities of the Norwegian Ministry of Foreign Affairs, Annual Report 1999*. Norwegian Ministry of Foreign Affairs.

Norwegian Ministry of Foreign Affairs (2001) *Evaluation Activities of the Norwegian Ministry of Foreign Affairs, Annual Report 2000*. Norwegian Ministry of Foreign Affairs.

Operations Evaluation Department (1999) *1998 Annual Report on Operations Evaluation*. World Bank Operations Evaluation Department.

Patton, Michael Quinn (1997) *Utilization-Focused Evaluation: The New Century Text, Edition 3*. SAGE.

Pawson, R. (2002) 'Evidence Based Policy : In Search of a Method' *Evaluation Vol 8, No. 2, pp157-181*.

Pollitt, Prof. Christopher and Prof. Helen O'Neill (1999) *An Evaluation of the Process of Evaluation of EC External Aid Programmes*. COTA.

Preskill, Hallie and Rosalie T. Torres (2000) 'The Learning Dimension of Evaluation Use' *New Directions for Evaluation No.88*. Jossey-Bass.

Strategic Policy Making Team, Cabinet Office (1999) *Professional Policy Making for the Twenty First Century*. UK Cabinet Office.

The Research Council of Norway, Strategic Planning (1996) *Evaluation Strategy*. The Research Council of Norway.

UK Government (1999) *Modernising Government White Paper*. TSO.

Valovirta, Ville (2002) 'Evaluation Utilization as Argumentation' *Evaluation Vol 8, No. 1, pp60-80*. SAGE.

Van der Knapp, P. (2001) *Policy Evaluation in the Netherlands: Linking Budgets to Operations and Results*. Netherlands Ministry of Finance.

Van der Knapp, P. (2000) 'Performance Management and Policy Evaluation in the Netherlands : Towards an Integrated Approach ' *Evaluation Vol 6, No. 3 pp335-350*. SAGE.

Van der Knapp, P. (1995) 'Policy Evaluation and Learning : Feedback, Enlightenment or Argumentation' *Evaluation Vol 1, No. 2 pp189-216*. SAGE.

Weiss, Carol Hirschon (1999) 'The Interface between Evaluation and Public Policy' *Evaluation Vol 5, No. 4, pp468-486*. SAGE.

Elliot Stern

ekonomista, politolog, antropolog, badacz - wybitna osobowość w dziedzinie ewaluacji programów publicznych. Członek zarządu w Brytyjskiej Akademii Nauk Społecznych, prezes m.in. Brytyjskiego Towarzystwa Ewaluacyjnego (1994-1996), Europejskiego Towarzystwa Ewaluacyjnego (2002-2003), obecnie prezes Międzynarodowej Organizacji do Współpracy w Ewaluacji (IOCE). Wykładowca akademicki, redaktor czasopisma Ewaluacja wydawanego przez Tavistock Institute, którego przez wiele lat był dyrektorem. Prowadził ewaluacje m.in. dla brytyjskiego rządu i w Brytyjskim Towarzystwie Ochrony Zdrowia. Ponadto jest także niezależnym konsultantem ds. ewaluacji przy takich instytucjach jak Komisja Europejska, Bank Światowy, UNESCO, Międzynarodowa Organizacja do Spraw Badań nad Rolnictwem (CGIAR).

Rynek ewaluacji w Polsce

– diagnoza stanu i perspektywy rozwoju

Wprowadzenie

Rynek ewaluacyjny w Polsce kształtuje się od około dziesięciu lat. Dynamika jego rozwoju była od samego początku powiązana z integracją z Unią Europejską. Zawierają się w tym zarówno działania związane z realizacją projektów finansowanych z funduszy przedakcesyjnych (Phare, Sapard), jak i reform przeprowadzonych w różnych obszarach polityk publicznych. Wstąpienie Polski do Unii Europejskiej stało się bodźcem, który powinien silnie wzmacniać dynamikę rozwoju rynku ewaluacyjnego. W sposób oczywisty jest to związane ze znaczącym zwiększeniem skali funduszy napływających z budżetu Wspólnoty, zarówno w obecnym okresie programowym jak i w następnym (2007-2013), w puli, których znajdują się znaczne środki pomocy technicznej przeznaczone m.in. na finansowanie działań związanych z ewaluacją programów operacyjnych.

Integracja ze strukturami europejskimi wiązała się z potrzebą spełnienia przez polską administrację szeregu wymogów, określonych prawem krajowym i europejskim, a związanych zarówno z przygotowaniem sprawnego systemu instytucjonalnego zdolnego do zaabsorbowania przyznanej pomocy oraz do przeprowadzania różnych typów ewaluacji, które mają gwarantować efektywne – zgodne z celami polityki spójności społeczno-gospodarczej – wykorzystanie funduszy.

Zaznaczyć jednak należy, iż nie możemy mówić o bezpośredniej zależności pomiędzy wysokością napływających środków, a kondycją rodzimego rynku ewaluacyjnego. Popyt i podaż zależą, bowiem od wielu czynników. Wzmoczona aktywność administracji publicznej powodowana wdrażaniem unijnych programów wsparcia, przejawiająca się wzrastającą liczbą zlecanych ewaluacji może je korzystnie stymulować, jednak w ograniczonym tylko zakresie wpływa na poprawę jakości usług z zakresu ewaluacji.

Uwarunkowania organizacyjno - proceduralne

Ewaluacja jako instrument zarządzania programami o charakterze społeczno-gospodarczym finansowanymi z funduszy strukturalnych Unii Europejskiej posiada swoje umocowanie prawne. Jest ono zawarte w następujących aktach:

- Rozporządzeniu Rady Unii Europejskiej (WE) nr 1260/1999 z dnia 21 czerwca 1999 r., ustanawiających przepisy ogólne w sprawie funduszy strukturalnych;
- Podstawach Wsparcia Wspólnoty dla Polski na lata 2004-2006 przyjętych przez Komisję Europejską 22 czerwca 2004r;
- Ustawie z dnia 20 kwietnia 2004r o Narodowym Planie Rozwoju.

Wymienione akty prawne znajdują ponadto swoje uzupełnienie w rozporządzeniach wynikających z ustawy o Narodowym Planie Rozwoju i dotyczących programów operacyjnych.

Ustawa o Narodowym Planie Rozwoju uwzględniając wymogi europejskie wyróżnia i definiuje zakres 3 typów ewaluacji, której podlegają wszystkie dokumenty programowe współfinansowane ze środków wspólnotowych, tj. Narodowy Plan Rozwoju i programy operacyjne. Ustawa wyróżnia: ocenę szacunkową (ewaluację ex-ante) - przed rozpoczęciem realizacji programu, ocenę w połowie okresu realizacji (ewaluację mid-term) oraz oceną pełną (ewaluację ex-post) - po zakończeniu realizacji. Ponadto, ustawa wprowadza czwarty typ ewaluacji, tzw. ocenę uzupełniającą (ewaluację on-going), która polega na zbadaniu ad hoc, w trakcie realizacji programu stopnia efektywności wykorzystania środków publicznych oraz skuteczności funkcjonowania systemu ich wdrażania. Poprzez analizę problemów pojawiających się w trakcie wdrażania programu i proponowanie rozwiązań usprawniających funkcjonowanie systemu, ewaluacja ta przyczyni się do wsparcia procesu zarządzania NPR i programami operacyjnymi.

W celu realizacji zadań związanych z ewaluacją Narodowego Planu Rozwoju utworzona została Krajowa Jednostka Oceny ulokowana w Departamencie Koordynacji Polityki Strukturalnej Ministerstwa Gospodarki i Pracy. Za ewaluację programów operacyjnych odpowiadają 1-2 osobowe wyspecjalizowane zespoły ewaluacyjne ulokowane w strukturach instytucji zarządzających tymi programami. Na potrzeby obecnego okresu programowego: 2004-2006 powołano 6 takich zespołów ewaluacyjnych.

Zadaniem Krajowej Jednostki Oceny oraz poszczególnych zespołów jest zlecenie ewaluacji wymaganych regulacjami oraz identyfikacja – w porozumieniu z instytucjami

zarządzającymi, zarówno Podstawami Wsparcia Wspólnoty jak i poszczególnymi programami operacyjnymi - tematów/obszarów ewaluacji on-going oraz ich zamawianie. Ponadto Krajowa Jednostka Oceny i wszystkie zespoły ewaluacyjne prowadzą działania służące budowaniu potencjału ewaluacyjnego wśród jednostek administracji wszystkich szczebli zaangażowanych we wdrażanie funduszy strukturalnych, w tym prowadzą działania informacyjno-edukacyjne (organizacja szkoleń i spotkań, rozbudowa stron internetowych). Wszystkie działania przewidziane do realizacji w obszarze ewaluacji w obecnym okresie programowym, czyli do 2008 roku zaplanowano w ramach Planów ocen opracowanych dla NPR i poszczególnych programów operacyjnych⁴.

Dodatkowo powołano Grupę Sterującą Ewaluacją NPR/PWW, która pełni rolę forum konsultacyjno-informacyjne pomiędzy Krajową Jednostką Oceny, a zespołami prowadzącymi ewaluację w ramach poszczególnych instytucji zarządzających programami operacyjnymi⁵.

Otoczenie instytucjonalne

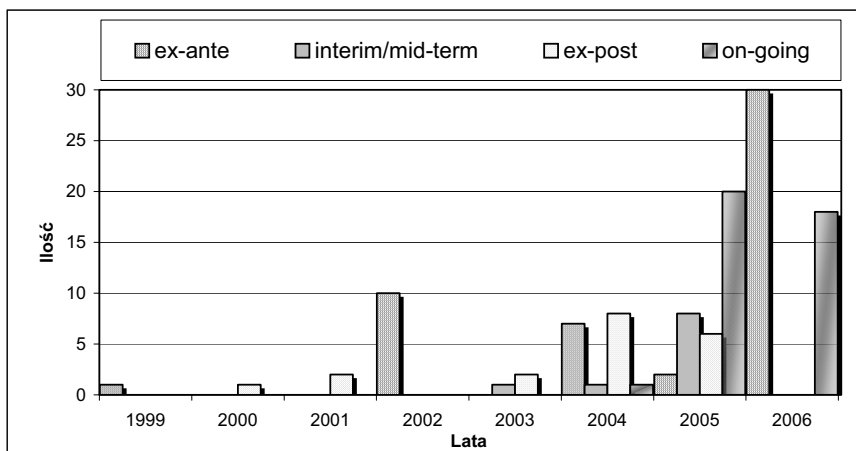
Zgodnie z opisanymi powyżej uwarunkowaniami proceduralnymi, na instytucjach zarządzających programami finansowanymi z funduszy strukturalnych, spoczywa obowiązek przeprowadzania ewaluacji tych programów. Doświadczenia polskie z wdrażaniem funduszy przedakcesyjnych w latach 1999-2005 dotyczą głównie ewaluacji interim/mid-term i ex-post (łącznie zlecono ok. 30 ewaluacji), por. wykres poniżej. Od 2002 roku rozpoczęto ewaluacje interwencji funduszy strukturalnych w latach 2004-2006, począwszy od zlecenia w tym roku przez Ministerstwo Gospodarki i Pracy ewaluacji ex-ante dokumentów programowych, zamówienia przez instytucje zarządzające programami operacyjnymi (Ministerstwo Gospodarki i Pracy, Ministerstwo Rolnictwa i Rozwoju Wsi i Ministerstwo Infrastruktury) ewaluacji ex-ante uzupełnień tych programów w 2004 r., po pierwszą ewaluację on-going dot. procesu wyboru wniosków do Inicjatywy EQUAL w Polsce.

W chwili obecnej możemy, więc mówić o pewnych doświadczeniach ewaluacyjnych poszczególnych instytucji. Sytuacja ta jednak coraz szybciej ulega zmianie, wraz z rosnącą ilością przeprowadzanych ewaluacji, oraz zacieśniającą się współpracą między instytucjami zarządzającymi i wdrażającymi.

⁴ Plany ocen udostępnione są na stronie internetowej MGİP: www.fundusze-strukturalne.gov.pl/Ewaluacja.

⁵ Szczegółowe informacje dotyczące proceduralnych uwarunkowań ewaluacji w Polsce odnaleźć można w poradniku przygotowanym przez MGİP: „Ewaluacja Narodowego Planu Rozwoju i programów operacyjnych w Polsce” (www.fundusze-strukturalne.gov.pl/Ewaluacja).

Wykres 1. Ewaluacje funduszy przedakcesyjnych (Phare, Sapard) i strukturalnych zlecone i planowane do zlecenia przez administrację publiczną w okresie 1999-2006.



Źródło: opracowanie na podstawie wyników badania rynku ewaluacyjnego w Polsce przeprowadzonego przez PARP oraz na podstawie „Planu Oceny NPR na lata 2004-2006”.

Rok 2005 przyniósł zdynamizowanie działań w obszarze ewaluacji programów finansowanych z funduszy strukturalnych. W pierwszej połowie roku zrealizowano ewaluację ex-ante wstępnego projektu Narodowego Planu Rozwoju na lata 2007-2013 oraz zamówiono ad hoc kilka ewaluacji on-going obejmujących analizę wybranych grup projektów oraz systemów realizacyjnych. Zgodnie z przygotowanymi dla NPR i programów operacyjnych *Planami oceny*⁶, do końca br. ma zostać zleconych łącznie ok. 20 takich ewaluacji, natomiast w przyszłym roku – kolejnych 18.

Wielkim wyzwaniem zarówno dla administracji, jak i ewalatorów zewnętrznych jest proces ewaluacji ex-ante interwencji funduszy strukturalnych planowanych na lata 2007-13. Zgodnie z projektem NPR na lata 2007-2013 ewaluacje ex-ante należy przeprowadzić w 2006 roku dla projektów 30 programów operacyjnych (por. wykres powyżej).

Do głównych instytucji zaangażowanych w zlecenie ewaluacji programów przedakcesyjnych (Phare, Sapard) należą Polska Agencja Rozwoju Przedsiębiorczości, Urząd Komitetu Integracji Europejskiej i Ministerstwo Rolnictwa i Rozwoju Wsi. Natomiast

⁶ Plany ocen, opracowane dla NPR i poszczególnych programów operacyjnych są dostępne na stronie MGiP: www.fundusze-strukturalne.gov.pl/Ewaluacja.

głównym zleceniodawcą ewaluacji funduszy strukturalnych jest Ministerstwo Gospodarki i Pracy, które obecnie skupia funkcje 4 instytucji zarządzających: PWW, SPO *Rozwój zasobów ludzkich*, *Wzrost konkurencyjności przedsiębiorstw*, ZPORR i PO *Pomoc Techniczna*.

Polska Agencja Rozwoju i Przedsiębiorczości (PARP)

Polska Agencja Rozwoju i Przedsiębiorczości od początku działalności wykonuje ewaluacje wdrażanych programów. W ramach PARP za prowadzenie ewaluacji odpowiedzialna jest Sekcja Ewaluacji (obecnie 4 osoby) działająca w ramach Zespołu Programowania i Ewaluacji.

Komórka ewaluacyjna PARP przeprowadza ewaluacje, zarówno zewnętrzne – realizowane przez zewnętrznych wykonawców, jak i ewaluacje wewnętrzne – realizowane we własnym zakresie. Większość projektów ewaluacyjnych ma charakter ex-post. Ograniczenie to uwarunkowane jest głównie środkami, jakie przeznacza się na przeprowadzanie badań. Zrealizowane dotychczas projekty dotyczyły programów krajowych i programów wsparcia z funduszy przedakcesyjnych Unii Europejskiej.⁷ W minionym roku zrealizowano 8 projektów (w tym dwa obejmujące cały komponent Phare 2000 RZL i MSP). Do końca roku 2005 r. PARP zrealizuje 5 projektów (w tym ewaluację całych komponentów RZL i MSP w ramach Phare 2001).

Urząd Komitetu Integracji Europejskiej (UKIE)

W ramach instytucji za ewaluację odpowiada Jednostka Ewaluacyjna Narodowego Koordynatora Pomocy. Komórka ta została utworzona w 2003 roku, przejmując obowiązki spoczywające uprzednio na Komisji Europejskiej. Obecnie zatrudnia 3 osoby na pełnym etacie oraz jedną osobę na pół etatu. W planach kadrowych przewiduje się poszerzenie tego stanu do 4 osób zatrudnionych na pełnym etacie. Wszystkie raporty ewaluacyjne UKIE są dostępne na stronach internetowych Urzędu. Przeprowadzane są one zgodnie z Rocznym Planem Ewaluacji UKIE. Obecnie jednostka prowadzi 8 projektów ewaluacyjnych. Wszystkie projekty wykonywane są przez wykonawcę zewnętrznego.

W ramach UKIE badaniami ewaluacyjnymi zajmuje się ponadto Departament Informacji Europejskiej. Prowadzi on ewaluacje wewnętrzne programu Sieci Regionalnych Centów Informacji Europejskiej. Są to projekty typu interim. Dotychczas przeprowadzono ewaluację 3 edycji programu.

⁷ Wszystkie raporty ewaluacyjne PARP są dostępne na stronie internetowej www.parp.gov.pl/ewalla.php

Ministerstwo Rolnictwa i Rozwoju Wsi

W Ministerstwie Rolnictwa i Rozwoju Wsi ewaluacją zajmuje się Wydział Monitorowania i Oceny PO SAPARD oraz Sektorowego Programu Operacyjnego. Za ewaluację bezpośrednio odpowiada dwóch pracowników Wydziału. Docelowo planuje się zatrudnienie dodatkowych 2-3 pracowników.

Wszystkie ewaluacje prowadzone przez MRiRW mają charakter zewnętrzny. Dotychczas zakończony został jeden projekt - ewaluacja mid-term PO SAPARD. W przyszłości planowane są ewaluacje SPO oraz znajdującego się obecnie na etapie projektowania Nowego Programu Rozwoju Obszarów Wiejskich.

Powyższy przekrój instytucji pokazuje, iż sieć komórek ewaluacyjnych w obrębie administracji publicznej znajduje się we wczesnej fazie kształtowania. Obecnie jednostki ewaluacyjne - w KJO oraz poszczególnych instytucjach zarządzających i wdrażających programy finansowane z funduszy Unii Europejskiej - zatrudniają łącznie 14 osób (por. tabela poniżej). Liczba ta jednak powinna stopniowo wzrastać by do końca 2005 roku osiągnąć poziom ok.. 30 osób.

Należy podkreślić, iż mimo dopiero zapoczątkowanych praktyk ewaluacyjnych możemy mówić już o wypracowanych wzorcach współpracy i wymiany doświadczeń. Pracownicy poszczególnych komórek i instytucji spotykają się regularnie zarówno w ramach posiedzeń grup roboczych, jaki w toku bieżących prac.

Tabela 2. Ulokowanie zespołów ewaluacyjnych w strukturach instytucji zarządzających NPR/PWW i programami operacyjnymi w latach 2004

Program	Ulokowanie funkcji ewaluacyjnych (jednostki ewaluacyjne)	STAN KADR ⁸	
		(na dzień 17.11.04)	(do końca 2005 r.)
Narodowy Plan Rozwoju/Podstawy Wsparcia Wspólnoty	Krajowa Jednostka Oceny Departament Koordynacji Polityki Strukturalnej, MGiP	3	5
Program Operacyjny Pomoc Techniczna	Wydział Pomocy Technicznej w Dep. Koordynacji Polityki Strukturalnej, MGiP	1	1
Zintegrowany Program Operacyjny Rozwoju Regionalnego	Wydział Raportowania i Monitorowania ZPORR, Dep. Wdrażania Programów Rozwoju Regionalnego, MGiP	1	1
SPO Rozwój Zasobów Ludzkich	Wydział Zarządzania, Departament Zarządzania Europejskim Funduszem Społecznym, MGiP	1	1
SPO Wzrost Konkurencyjności Przedsiębiorstw	Wydział Monitorowania, Raportowania i Oceny, Dep. Zarządzania Programem Wzrostu Konkurencyjności Gospodarki, MGiP	2	2
SPO Restrukturyzacja i modernizacja sektora żywnościowego oraz rozwój obszarów wiejskich	Wydział Monitorowania i Oceny Programów, Departament Rozwoju Obszarów Wiejskich, MRiRW	2	3
SPO Rybołówstwo i przetwórstwo ryb	Wydział Monitorowania i Raportowania Wykorzystania Funduszy Pomocowych, Departament Rybołówstwa, MRiRW	1	2
SPO Transport	Departament Programowania i Strategii, Ministerstwo Infrastruktury	1	1
Inicjatywa Wspólnoty Equal	Wydział Zarządzania, Departament Zarządzania Europejskim Funduszem Społecznym, MGiP	1 ⁹	1
Inicjatywa Wspólnoty Interreg	Wydział INTERREG, Departament Wdrażania Programów Rozwoju Regionalnego, MGiP	1	2
Fundusz Spójności	Departament Koordynacji Funduszu Spójności, MGiP	1	11 ¹⁰
	SUMA	14	30

Źródło: „Plan Oceny Narodowego Planu Rozwoju na lata 2004-2006”.

Perspektywa kolejnego okresu programowego 2007-2013 z liczbą programów operacyjnych ok. 30, doprowadzi do 4-krotnego wzrostu ilości instytucji odpowiedzialnych za zlecenie ewaluacji na poziomie programów operacyjnych, a w konsekwencji do

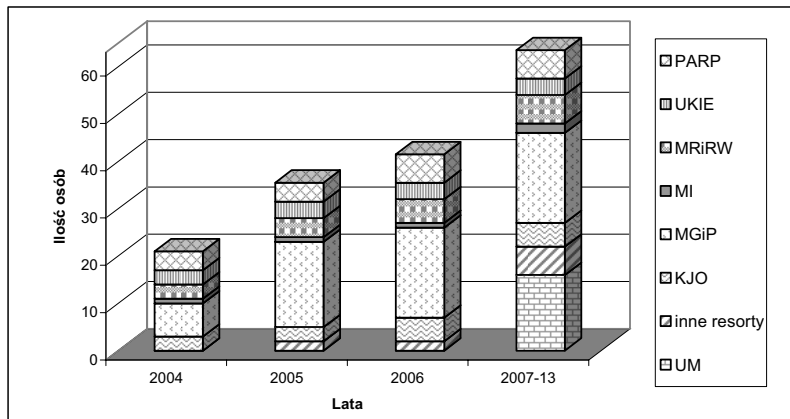
⁸ Osoby zaangażowane w prace ewaluacyjne.

⁹ Funkcjonuje jednoosobowe stanowisko ds. Oceny Programów SPO RZL i Inicjatywy Wspólnoty Equal.

¹⁰ Z 11 osób zatrudnionych w Wydziale Monitorowania 1 osoba pełni rolę koordynatora działań prowadzonych w ramach oceny Funduszu Spójności, natomiast 10 osób - będzie odpowiedzialnych jednocześnie za monitorowanie i ocenę poszczególnych projektów realizowanych przez Departament Koordynacji Funduszu Spójności.

zwiększenia liczby osób bezpośrednio zaangażowanych w te działania (patrz wykres poniżej).

Wykres 2. Kadry zaangażowane w ewaluację programów o charakterze społeczno-gospodarczym.



Źródło: Wyniki badania rynku ewaluacyjnego w Polsce przeprowadzonego przez PARP oraz na podstawie „Planu Oceny NPR na lata 2004-2006”. W prognozie na lata 2007-13 przyjęto, że do zadań ewaluacyjnych wyznaczona zostanie min. 1 osoba w każdej instytucji zarządzającej programem operacyjnym (instytucje te zostały nazwane jako „inne resorty”).

Informacje o rynku ewaluacji w Polsce

Najczęstszym zleceniodawcą projektów ewaluacyjnych jest bezsprzecznie administracja publiczna szczebla centralnego oraz administracja europejska. Wynika to w dużym stopniu z obowiązku przeprowadzenia ewaluacji, jako elementu rozliczenia wdrażanych programów. Instytucjami zamawiającymi są najczęściej Urząd Komitetu Integracji Europejskiej, Jednostka Finansująco-Kontraktująca, Komisja Europejska, Ministerstwo Gospodarki i Pracy oraz Polska Agencja Rozwoju Przedsiębiorczości.

Rynek ewaluacyjny w Polsce cechuje się dość dużym rozproszeniem. Utrudnia to ustalenie liczby realizowanych projektów, oraz wykonujących je podmiotów. Szacunkowo ilość podmiotów zajmujących się regularnie ewaluacją zamyka się w liczbie 15-20. Bezpośrednimi wykonawcami usług są zazwyczaj indywidualni ewaluatorzy realizujący poszczególne projekty na zamówienia firm badawczych, konsultingowych, lub we współpracy z innymi przedstawicielami branży ewaluacyjnej.

Podmiotami ubiegającymi się o zamówienia na przeprowadzenie ewaluacji są zazwyczaj:

- 1) **Firmy badawcze i konsultingowe.** W strukturach tych firm nie istnieją zasadniczo komórki lub zespoły wyspecjalizowane w prowadzeniu ewaluacji. Tworzone są one zadaniowo, na potrzeby konkretnego projektu. W ramach firm konsultingowych, częstokroć projekty ewaluacyjne prowadzone są przez zespoły zajmujące się sprawami europejskimi (funduszami strukturalnymi i innymi formami wsparcia).
- 2) **Firmy wyspecjalizowane w prowadzeniu ewaluacji.** Jest to wąskie grono firm wyspecjalizowanych w ewaluacji określonych obszarów polityk publicznych, a ewaluacja stanowi element ich *core business*.
- 3) **Ośrodki akademickie.** Częstokroć występują one w konsorcjach z partnerami zagranicznymi i krajowymi. Przykładu dostarcza tu Centrum Europejskich Studiów Regionalnych i Lokalnych UW. Przedstawiciele środowiska akademickiego są ponadto zatrudniani indywidualnie, jako eksperci, przez firmy badawcze i konsultingowe.
- 4) **Wykonawcy zagraniczni.** Są to podmioty legitymujące się doświadczeniem w obszarach, na których wykonawcy krajowi dotychczas nie działali. Podmioty te są zazwyczaj wykonawcami ewaluacji typu *ex-ante*. W praktyce chętnie korzystają oni z usług krajowych ewaluatorów i ekspertów celem obniżenia kosztów realizacji projektu.
- 5) **Krajowe stowarzyszenie branżowe** (Polskie Towarzystwo Ewaluacyjne). Od około roku PTE znacznie ograniczyło swoją aktywność, jako wykonawca projektów ewaluacyjnych. Obecnie nie uczestniczy ono w przetargach, a jedynie realizuje projekty szkoleniowe.
- 6) **Pozarządowe ośrodki i instytuty badawcze.**

Jakość produktów ewaluacyjnych z perspektywy administracji publicznej

Funkcjonujące wśród przedstawicieli administracji publicznej opinie, dotyczące jakości oferowanych na polskim rynku usług ewaluacyjnych są podzielone. Wskazuje się na następujące czynniki oddziaływujące negatywnie:

- 1) **Brak znajomości specyfiki programów wdrażanych w ramach określonej polityki publicznej.** Jest to związane z widocznymi brakami po stronie popytowej. W konsekwencji działający na polskim rynku ewaluatorzy pełnią niejako rolę

omnibusów, realizując projekty związane z dziedzinami wymagającymi zaangażowania wiedzy specjalistycznej.

- 2) **Brak znajomości specyfiki instytucji, zarówno zarządzających, jak i wdrażających.** Często wykonawcy zewnętrzni podejmując się oceny programu, na podstawie badań przedstawiają wnioski i zalecenia oderwane od rzeczywistości programu i specyfiki instytucji. Wielokrotnie ewaluatorzy przeprowadzając ocenę nie rozumieją cykliczności systemu monitoringu, logiki działania instytucji i wdrażania programu. Przejawia się to w nieuzasadnionej krytyce monitoringu i sposobu gromadzonych danych, braku możliwości uzyskania danych o programie, niezdolności instytucji do udzielenia informacji o wdrażanych programach i zgłaszaniu przez ewaluatorów trudności w zidentyfikowaniu właściwych osób lub departamentów zajmujących się ocenianym programem.
- 3) **Niedostateczna rola ekspertów posiadających doświadczenie w zakresie ewaluacji.** Wynika to ze wspomnianego powyżej zagadnienia. Silna konkurencja cenowa wpływa nieuchronnie na sposób budowania zespołów ewaluacyjnych przez potencjalnych wykonawców. Ponadto eksperci zostają dobierani w nie do końca właściwy sposób. Częstokroć reprezentują oni dyscypliny niepowiązane ze specyfiką danego projektu. Przekłada się to w prosty sposób na jakość rekomendacji, a więc rzutuje na użyteczność całego projektu.
- 4) **Zaniżenie cen usług ewaluacyjnych.** Niewielka liczba podmiotów zajmujących się ewaluacją na rynku i duża rola doświadczenia w tym zakresie przy wyborze wykonawców, powoduje, że firmy przeprowadzają badania ewaluacyjne „po kosztach”. Firmy badawcze chcąc zdobyć cenne doświadczenie i zaistnieć na rynku ewaluacyjnym zaniżają ceny, przez co w dużej mierze traci na tym jakość i zakres przeprowadzanych ewaluacji, a także trafność rekomendacji. Realizowane badania rodzą wiele pytań, jednak ograniczone zasoby wykonawców nie pozwalają na przeprowadzenie pogłębionych badań lub wynajęcie eksperta z badanej dziedziny.
- 5) **Ograniczone budżety badań ewaluacyjnych.** W przypadku tzw. ewaluacji wewnętrznych realizowanych przez instytucje wdrażające, ograniczone środki również obniżają jakość raportów z badań. Ewaluacje tego typu mają ograniczony zakres, oceniając najczęściej jeden wymiar rzeczywistości programu – badanie opinii beneficjentów. Zbadanie instytucji, wykonawców i innych podmiotów zaangażowanych w realizację programu ograniczają posiadane zasoby. Ograniczenia budżetowe determinują także przyjętą metodologię badań.

Praktyczne wykorzystanie ewaluacji

O skuteczności ewaluacji, czyli o tym czy spełniła ona swoją funkcję (tj. czy poprawiła funkcjonowanie programu, podniosła jakość interwencji publicznej, efektywność, etc.), czy też jest tylko spełnieniem wymogu administracyjnego decydują przede wszystkim: jakość raportu, proces rozpowszechnienia wyników oraz możliwość wykorzystania rekomendacji.

Wedle opinii zlecający ewaluację najczęściej występującym ograniczeniem w wykorzystaniu wyników ewaluacji stanowi niska jakość rekomendacji. Często ich adekwatność jest ograniczana przez niewłaściwe zrozumienie specyfiki ewaluowanego programu oraz instytucjonalnego kontekstu jego wdrażania. Ograniczenia stosowalności zaleceń mogą jednakże posiadać inne podłoże, niekoniecznie związane z jakością usług dostarczanych przez wykonawcę. Wskazać należy w tym miejscu na następujące zagadnienia:

- 1) **Niedostateczny rozwój kultury ewaluacyjnej.** Przejawia się to w braku rozumienia samego pojęcia ewaluacji wśród niektórych przedstawicieli administracji. Jest ona ponadto niezmiernie rzadko postrzegana jako narzędzie służące poprawie jakości funkcjonowania instytucji publicznych oraz świadczonych przez nie usług. W percepcji osób zaangażowanych we wdrażanie programów, ewaluacja stanowi najczęściej jeden z instrumentów kontrolnych, będąc utożsamiana z audytem i monitoringiem.
- 2) **Niestabilność instytucjonalna.** Zmiany struktury oraz priorytetów w obrębie administracji zakłócają proces wdrażania rekomendacji. Problem ten również szczególnie silnie związany jest z niestabilnością komórek ewaluacyjnych ulokowanych w instytucjach administracji publicznej
- 3) **Niewłaściwe ulokowanie ewaluacji w cyklu wdrażania projektu.** Rekomendacje formułowane są często zbyt późno, aby możliwe było ich wdrożenie. Jest to w dużym stopniu uwarunkowane złożonością procedur związanych z wyborem wykonawcy ewaluacji lub ograniczeniami zasobów.
- 4) **Zakłócenia decyzyjne związane z cyklem wyborczym.** Stosowalność rekomendacji wymagających określonych decyzji na politycznym szczeblu administracji jest zakłócana przez logikę cyklu wyborczego i jej wpływ na zapadające na tym szczeblu decyzje.

Podkreślenia wymaga jednak fakt, iż wszyscy przedstawiciele administracji publicznej, z którymi przeprowadzono rozmowy przytaczali przykłady udanego wykorzystania rekomendacji pochodzących z ewaluacji. Najczęściej wskazywano na udane modyfikacje

grup docelowych beneficjentów programów, oraz na zgodne z założeniami danej polityki zmiany stosowanych instrumentów pomocy publicznej.

Budowanie potencjału ewaluacyjnego

Krajowa Jednostka Oceny

Budowanie potencjału ewaluacyjnego polskiej administracji w zakresie ewaluacji programów o charakterze społeczno-gospodarczym stanowi jeden z priorytetowych celów wskazanych w *Planie Oceny Narodowego Planu Rozwoju na lata 2004-2006*.

Wdrożenie tego celu w okresie realizacji NPR 2004-2006, czyli do końca 2008 roku zostanie zapewnione poprzez wzmocnienie potencjału administracyjnego i koordynacyjnego Krajowej Jednostki Oceny do prowadzenia efektywnej koordynacji nad procesem budowy systemu ewaluacji w Polsce oraz przeprowadzania ewaluacji NPR/PWW, a także rozbudowę potencjału jednostek ewaluacyjnych ulokowanych w instytucjach zarządzających programami operacyjnymi.

Obok wzrostu liczebności kadr i wyodrębnienia tam gdzie okaże się to niezbędne specjalnych zespołów zajmujących się ewaluacją, powołanie Grupy Sterującej Ewaluacją NPR/PWW, Grupy konsultacyjnej ds. systemu oceny funduszy UE w Polsce (złożonej z ekspertów zewnętrznych) i przygotowanie Poradnika stanowiącego wykładnię obowiązujących procedur i zasad przy przeprowadzaniu procesów ewaluacyjnych, przyczyniają się do skoordynowania działań z zakresu ewaluacji na poziomie NPR/PWW i programów operacyjnych. Dodatkowo, wspomaganie rozwoju zdolności ewaluacyjnej zarówno wśród administracji publicznej, jak i ewaluatorów zewnętrznych odbywa się poprzez rozpowszechnianie wiedzy i informacji nt. ewaluacji (rozbudowa strony internetowej poświęconej ewaluacji, organizacja szkoleń dla administracji publicznej (szczególnie centralnego i regionalnego) i konferencji specjalistycznych z udziałem ekspertów).

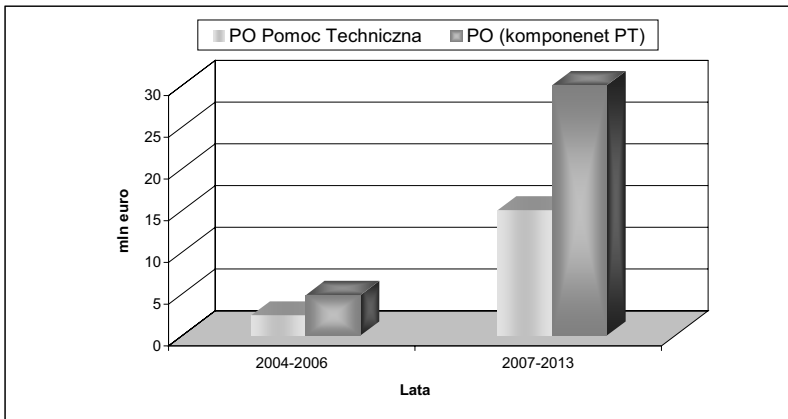
Głównym źródłem finansowania przedsięwzięć ewaluacyjnych w ramach NPR/PWW 2004-2006 jest Program Operacyjny Pomocy Technicznej na lata 2004-2006 (budżet ok. 2,4 mln Euro¹¹). Prace ewaluacyjne w instytucjach zarządzających poszczególnymi

¹¹ Budżet jest przeznaczony na wsparcie procesu oceny NPR/PWW: oceny (ewaluacji) efektywności wdrażania PWW i poszczególnych programów operacyjnych, oceny po zakończeniu okresu programowania (*ex-post*), oceny *ex-ante* przyszłych interwencji funduszy strukturalnych, oceny realizacji założonych celów horyzontalnych, jak również wypracowania i ciągłego doskonalenia procedur, metodologii i standardów oceny (ewaluacji).

programami operacyjnymi są finansowane w ramach działań z zakresu pomocy technicznej odpowiednich programów operacyjnych, których łączny budżet stanowi ponad 4 mln euro (por. wykres poniżej). Dodatkowo, w ramach umów bliźniaczych realizowanych w MGiP prowadzone są działania doradczo-szkoleniowe w zakresie ewaluacji.

Plany na kolejny okres programowy (2007-2013) zakładają alokacje w wysokości ok. 15 mln euro na ewaluację w ramach Programu Operacyjnego Pomoc Techniczna oraz ponad 7 mln euro w ramach komponentów pomocy technicznej programów operacyjnych.

Wykres 3. Środki na działania ewaluacyjne z Pomocy Technicznej



Źródło: Opracowanie własne na podstawie informacji przekazanych przez członków Grupy sterującej ewaluacją NPR/PWW oraz w oparciu o plan finansowy projektu Narodowego Planu Rozwoju 2007-13.

Polskie Towarzystwo Ewaluacyjne

Obok KJO znaczącym ośrodkiem kształtowania krajowego potencjału ewaluacyjnego jest Polskie Towarzystwo Ewaluacyjne. Jego realizacji służą działania ukierunkowane na stworzenie platformy wymiany doświadczeń, skupiającej możliwie szerokie grono rodzimych ewaluatorów. Jednym z tego typu działań jest organizowanie cyklicznych seminariów poświęconych zagadnieniu związanym ze specyfiką wykonywania zawodu ewaluatora na polskim rynku. PTE stawia sobie również za cel tworzenie katalogu dobrych praktyk i stymulowanie jakości świadczonych usług. Obecnie skupia ono około 50 członków. Ponadto PTE prowadzi lub współprowadzi szkolenia adresowane do różnego

rodzaju odbiorców (administracja, organizacje pozarządowe). W 2003r. PTE było także współorganizatorem konferencji poświęconej ewaluacji.

Inne czynniki kształtujące potencjał ewaluacyjny w Polsce

Czynnikami wpływającymi na budowanie potencjału ewaluacyjnego są wspólne inicjatywy i działania podejmowane przez różne instytucje administracji publicznej. Obok spotkań branżowych wymieniłem należy szkolenia organizowane na wszystkich poziomach, zarówno wewnętrznie - dla budowania kompetencji wewnątrz instytucji jak i dla instytucji zewnętrznych.

Na uwagę zasługuje również praktyka upowszechniania wyników zamawianych ewaluacji. Zarówno PARP¹², KJO, jak i UKIE umieszczają na swoich stronach internetowych wszystkie raporty wykonane na ich zlecenie oraz przez pracowników tych instytucji. Organizowane są ponadto prezentacje wyników ewaluacji oraz spotkania robocze, w trakcie, których przedstawiciele poszczególnych instytucji mają okazję do wymiany poglądów i wypracowania rekomendacji.

Oferta edukacyjna w zakresie ewaluacji

Rozwój wiedzy w obszarze ewaluacji ograniczany jest jednak przez słabą dostępność oferty edukacyjnej. Pierwsze kierunkowe propozycje wyższych uczelni w tym zakresie pojawiły się stosunkowo niedawno na przełomie lat 2000/2001 i początkowo związane były jedynie z ewaluacją w obszarze edukacji. Pierwsze studia poświęcone ewaluacji w innych obszarach niż edukacja rozpoczęto dopiero dwa lata temu (2003 rok/ Instytut Socjologii Uniwersytetu Warszawskiego). Obecnie tylko dwie uczelnie oferują studia podyplomowe związane z ewaluacją funduszy strukturalnych w Polsce, Uniwersytet Warszawski oraz po raz pierwszy w tym roku, Wyższa Szkoła Przedsiębiorczości i Zarządzania im. Leona Koźmińskiego. Wprawdzie na wielu uczelniach wyższych prowadzone są już zajęcia fakultatywne poświęcone ewaluacji programów o charakterze społeczno – gospodarczym, jednak to wciąż nie odpowiada stale rosnącemu zapotrzebowaniu rynku.

Duże ograniczenia w rozwoju dziedziny wynikają także z braku szkoleń specjalistycznych dostępnych na poziomie zaawansowanym oraz polskojęzycznej literatury fachowej.

¹² <http://www.parp.gov.pl/ewal1a.php>

Podsumowanie

Rynek ewaluacji w Polsce staje przed trudnościami właściwymi dla początkowego etapu rozwoju. Należy jednak zaznaczyć, iż w porównaniu z innymi rynkami nowych krajów członkowskich UE, nie mamy podstaw by mówić o zapóźnieniu. Ilustracją tego są wyniki badań potencjału ewaluacyjnego przeprowadzonych przez *Centre for Strategy and Evaluation Services*¹³. W ich świetle polska kultura ewaluacyjna cechuje się bardzo wysokim stopniem rozwoju, w porównaniu z innymi krajami Europy Środkowo-Wschodniej. Problemy, z którymi borykają się wykonawcy oraz instytucje zamawiające, znajdują swoje wyjaśnienie w niewielkim doświadczeniu z zakresu badań ewaluacyjnych.

Powyższe spostrzeżenia nie podważają jednak zarysowanych wcześniej ograniczeń występujących na polskim rynku ewaluacyjnym. Związane z nimi problemy nie mogą zostać jednak rozwiązane przez sam popyt na ewaluację, uwarunkowany napływającymi z funduszy UE środkami. Kwestia ta jest bowiem nierozłącznie powiązana z rozwojem potencjału ewaluacyjnego. Proces ten z kolei zwiększy swą dynamikę wyłącznie wówczas, gdy ewaluacja będzie postrzegana przez administrację publiczną jako narzędzie służące doskonaleniu jej codziennej praktyki, a przez wykonawców jako dochodowy i rozwojowy obszar działalności.

Opracowanie:

**Krajowa Jednostka Oceny, Ministerstwo Gospodarki i Pracy
Zespół Programowania i Ewaluacji, Polska Agencja Rozwoju Przedsiębiorczości**

¹³ Jack Malan „*Benchmarking evaluation capacity in the New Member States*” *Center for Strategy and Evaluation Services*”, www.cses.co.uk.

Krajowa Jednostka Oceny MGIP

Krajowa Jednostka Oceny (KJO) zgodnie z Ustawą o Narodowym Planie Rozwoju (art. 58 ust. 1) odpowiada za ewaluację Narodowego Planu Rozwoju. Krajowa Jednostka Oceny powstała w kwietniu 2004 r. jako Zespół ds. Oceny w ramach Departamentu Koordynacji Polityki Strukturalnej Ministerstwa Gospodarki i Pracy. Departament ten koordynuje zarówno działania z zakresu programowania interwencji funduszy Unii Europejskiej w Polsce, jak i sprawuje nadzór nad przebiegiem procesu realizacji Narodowego Planu Rozwoju, w tym pełni rolę Instytucji Zarządzającej Podstawami Wsparcia Wspólnoty.

Zespół Programowania i Ewaluacji PARP

Zespół Programowania i Ewaluacji jest komórką organizacyjną Polskiej Agencji Rozwoju Przedsiębiorczości odpowiedzialną za przeprowadzanie ewaluacji programów pomocowych wdrażanych w PARP, programowanie nowych instrumentów wsparcia oraz analizę sektora MSP.

Zespół PE jako samodzielna jednostka organizacyjna, istnieje od roku 2004. W ramach Zespołu działa Sekcja Ewaluacji bezpośrednio zaangażowana w proces oceny programów oraz Sekcja Programowania i Analiz prowadząca badania sektora MSP, dotyczące szeroko rozumianej przedsiębiorczości, barier jej rozwoju oraz otoczenia prawnego i instytucjonalnego. Wyniki badań znajdują zastosowanie nie tylko przy identyfikacji potrzeb, diagnozie sytuacji, czy przygotowaniu instrumentów wsparcia, ale także w badaniach porównawczych prowadzonych w ramach projektów ewaluacyjnych.

CZĘŚĆ I

Europejskie praktyki i rozwiązania w zakresie ewaluacji programów społeczno-gospodarczych sektora publicznego

David Hegarty

Ewaluacja funduszy strukturalnych w Irlandii¹⁴

Wprowadzenie

Niniejszy dokument opisuje zespół procedur i narzędzi wykorzystywanych w procesie ewaluacji funduszy strukturalnych w Irlandii od około dekady. Rozdział 7to prezentuje główne cechy trzech kolejnych programów Podstaw Wsparcia Wspólnoty (PWW) dla Irlandii na lata 1989-1993, 1994-1999 oraz 2000-2006. Następny rozdział opisuje zmiany zachodzące w strukturach ewaluacji, z uwzględnieniem zarówno strony organizacyjnej, jak i roli ewaluacji w procesie zarządzania programami finansowanymi ze środków funduszy strukturalnych. Wątek ten jest uzupełniony w rozdziale kolejnym, który poświęcony jest rozwojowi w podejściu do procesu ewaluacji oraz źródłom zainteresowań ewaluatorów. Warty podkreślenia aspektem w podejściu do procesu ewaluacji jest silny nacisk kładziony na kluczowe kwestie dotyczące finansowania różnego rodzaju inwestycji publicznych. Rozdział ostatni wyjaśnia jak ewaluacja wpłynęła na formułowanie strategii inwestycyjnych i alokację środków w ramach funduszy strukturalnych.

Tło

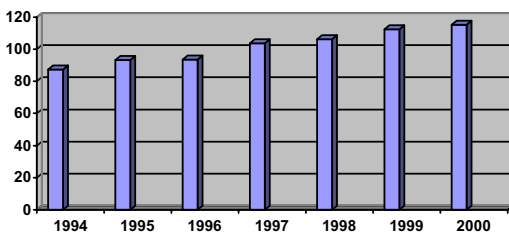
Począwszy od końca lat 80., Irlandia była odbiorcą znacznych środków w ramach funduszy strukturalnych. W trakcie okresów programowania 1989-1993 oraz 1994-1999 Irlandia była zaliczana do regionów objętych Celem 1. Wydatki wyniosły wówczas odpowiednio 4.2 mld € i 5.8 mld € i stanowiły 1,7 % średniego PKB z wszystkich tych lat. W obecnym okresie programowania 2000-2006, dwa regiony w Irlandii należą do regionów typu NUTS II. Są to: region Border, Midland and Western (BMW), który ma status regionu Celu 1 do 2006 r. oraz region Southern and Eastern (S&E), który kwalifikuje się do Celu 1 do 2005 r. jako region w trakcie transformacji. Włączając 4% rezerwę na

¹⁴ Niniejszy dokument stanowi materiał uzupełniający do prezentacji Davida Hegarty „Budowanie potencjału ewaluacyjnego – doświadczenie Irlandii”. Została ona opracowana przez autora, wówczas szefa Jednostki Oceny NPR/PWW w Irlandii, na potrzeby V Konferencji Europejskiej nt. ewaluacji funduszy strukturalnych zorganizowanej przez Komisję Europejską w Budapeszcie w czerwcu 2003 r.

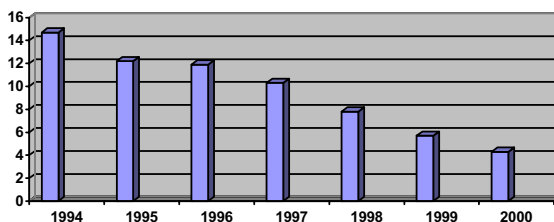
wykonanie, łączna alokacja z funduszy strukturalnych dla Irlandii na lata 2000-2006 wynosi 3,2 mld € (równowartość około 0,4% PKB w 2001 r.)

Gdy Irlandia należała do regionów objętych Celem 1, główny nacisk podczas obydwu okresów programowania był kładziony na długoterminowy wzrost potencjału gospodarki kraju i promocję gospodarczego zbliżenia Irlandii z Unią Europejską. Kolejnym celem było stawienie czoła brakowi równowagi na ówczesnym rynku pracy. Przejawiało się to w postaci ogromnych nakładów na rozwój zasobów ludzkich. (Stopa bezrobocia w Irlandii wynosiła średnio 14 % w ciągu I połowy lat 90., przy czym osoby trwale bezrobotne stanowiły 59 % wszystkich osób pozostających bez pracy). Pozytywny zwrot w irlandzkiej gospodarce, który nastąpił pod koniec lat 90., przyczynił się jednak do istotnej zmiany takiego stanu rzeczy. Podczas realizacji Podstaw Wsparcia Wspólnoty na lata 1994-1999 wzrost gospodarczy w Irlandii, mierzony PKB, wyniósł średnio 9 % rocznie (cel założony w PWW – 4 %). Stopa bezrobocia obniżyła się z prawie 15 % w 1994 r. do niespełna 6 % w roku 1999.

Wykres 1: PKB na osobę w PPS w Irlandii, EU-15=100 (Źródło: EUROSTAT)



Reagując na te zmiany, punkt ciężkości w obecnym NPR i PWW został przeniesiony na inne zagadnienia. W dużym stopniu nieoczekiwany wzrost gospodarczy z końca lat 90. przyczynił się do pojawienia się znaczących braków ilościowych i jakościowych w infrastrukturze gospodarczej i społecznej. Dotyczyło to m.in. transportu, mieszkalnictwa, ochrony środowiska oraz zdrowia. Pojawiły się także wątpliwości związane ze strukturą wzrostu gospodarczego w aspekcie terytorialnym, jak i stale pojawiającymi się problemami wykluczenia społecznego pewnych grup. W związku z tym, cele obecnego NPR/PWW są skierowane na deficyt w infrastrukturze gospodarczej, a także promocję zrównoważonego rozwoju regionalnego oraz zapobieganie wykluczeniom społecznym.

Wykres 2: Stopa bezrobocia w Irlandii (Źródło: CSO)

Struktury ewaluacji: ewolucja i rozwój

Irlandia stanowi ciekawe studium przypadku w dziedzinie rozwoju systemu ewaluacji funduszy strukturalnych. Po pierwsze, nie było wcześniej w irlandzkiej administracji publicznej tradycji związanych z formalną oceną wydatków publicznych. Po drugie, Irlandia może się pochwalić obecnie rozległym doświadczeniem w ewaluacji programów finansowanych ze środków funduszy strukturalnych, które nabyła w trakcie trzech kolejnych okresów programowania.

Prześledzona zostanie teraz skrótowo historia rozwoju systemów ewaluacji poczynszy od wczesnych lat 90.

PWW 1989-1993 i 1994-1999

Podczas pierwszego okresu programowania (1989-1993) wprowadzono wiele elementów związanych z ewaluacją. Ważnym wydarzeniem było utworzenie Jednostki Oceny EFS (1992 r.) oraz Jednostki Oceny Przemysłu (1993 r.). Jednostki te powstały na niezależnym gruncie pod egidą rządowych departamentów odpowiedzialnych odpowiednio za zarządzanie programami operacyjnymi dot. rozwoju zasobów ludzkich i rozwoju przemysłu. (W 1995 r. podobna jednostka została utworzona dla programu operacyjnego dot. Rolnictwa i rozwoju obszarów wiejskich). Od czasu ich powstania do 1999 r. ww. jednostki przeprowadziły wiele ewaluacji tematycznych oraz na poziomie działań.

W trakcie kolejnego okresu programowania (1994-1999) można było zaobserwować intensyfikację działań w zakresie ewaluacji. Na początku, PWW oraz programy operacyjne stały się przedmiotem oceny szacunkowej (ewaluacja ex-ante). Pomimo iż ocena ta została przeprowadzona na dosyć późnym etapie procesu negocjacji, a co za tym idzie, nie miała znaczącego wpływu na treść dokumentów programowych, stworzyła ona podstawę dla późniejszej oceny bieżącej (ewaluacja on-going) i oceny w połowie okresu

realizacji (ewaluacja mid-term). W czasie okresu programowania 1994-1999, w przypadku większości programów operacyjnych bieżąca ocena przeprowadzana była bądź przez wewnętrzne jednostki ds. ewaluacji bądź przez zewnętrznych ewaluatorów. W 1996 r. została utworzona Jednostka Oceny PWW (obecnie Jednostka Oceny NPR/PWW). Jej zadaniem była koordynacja prac jednostek ewaluacyjnych ulokowanych w strukturach instytucji zarządzających programami operacyjnymi oraz promocja dobrych praktyk z zakresu metodologii ewaluacji. W przedmiotowym okresie wykonano znaczącą ilość prac ewaluacyjnych. W 1998 r. w dokonanym przeglądzie przeprowadzonych w ramach PWW ewaluacji, Jednostka Oceny PWW wyodrębniła w latach 1995-1997 115 ocen bieżących zawartych w dziewięciu programach operacyjnych. Warto też podkreślić, że biorąc pod uwagę koszt procesu ewaluacji w połowie okresu realizacji, koszty ewaluacji do końca 1997 r. stanowiły 1,2 % wydatków w ramach PWW.

Należy zaznaczyć, iż mimo że nie było w tym czasie wymagań w zakresie przeprowadzania kompleksowej oceny PWW w połowie okresu realizacji, ocenę taką przeprowadzono w 1996 r. i 1997 r. i to zarówno na poziomie PWW, jak i programów operacyjnych. Przy czym jeśli chodzi o ocenę programów operacyjnych, działania zostały skoordynowane poprzez ustanowienie wspólnego zakresu usługi oraz uwzględnienie wyników ocen przy późniejszej ewaluacji na poziomie PWW¹⁵. Pozwoliło to skoncentrować się podczas oceny PWW na zagadnieniach systemowych. Ocena PWW w połowie okresu realizacji, przeprowadzona przez Instytut Badań Społeczno-Gospodarczych, jest powszechnie uznawana – zarówno w Irlandii, jak i w całej UE – jako przykład najlepszych praktyk w zakresie dużych, zróżnicowanych programów. Sposób podejścia do zadania i główne metody wykorzystywane podczas tej oceny były w późniejszym czasie wielokrotnie dyskutowane.

NPR/PWW 2000-2006

W obecnym okresie programowania nastąpiło w procesie ewaluacji wiele zmian. Obecnie stosowany zespół narzędzi i procedur jest wynikiem doświadczeń zdobytych w latach 1994-1999 kiedy to, jak wspomniano wyżej, miało miejsce wiele ocen przeprowadzanych zarówno przez wewnętrznych, jak i zewnętrznych ewaluatorów.

Jak już zostało powiedziane, w ramach PWW przeznaczone środki z funduszy strukturalnych przewidziane dla Irlandii na lata 2000-2006 to kwota rzędu 3,2 mld €, zaś

¹⁵ Zob. David Hegarty i Patrick Honohan, *Mid-term Evaluation of CSF In Ireland*, Europejska Konferencja nt. Praktyk w zakresie Ewaluacji Funduszy Strukturalnych, Sewilla, marzec 1998
52

łączna kwota przeznaczona na inwestycje publiczne w ramach PWW to prawie 5,5 mld €. PWW stanowią część Narodowego Planu Rozwoju (NPR), dokumentu o szerszym zakresie i całkowitym budżecie środków publicznych (włączając współfinansowanie UE) na poziomie 51 mld €. Jak z tego wynika, NPR składa się w głównej mierze z komponentów, które nie są współfinansowane ze środków funduszy strukturalnych lub też są, ale w niewielkim stopniu. Pomimo tego, władze irlandzkie postanowiły zastosować zasady dotyczące monitoringu ewaluacji funduszy strukturalnych do wszystkich inwestycji w ramach NPR, niezależnie od źródła finansowania. Z tego powodu, NPR i PWW są wdrażane jako programy zintegrowane. W związku z powyższym, w obecnym okresie programowania prace ewaluacyjne mają znacznie szerszy zasięg niż to miało miejsce wcześniej. Obecnie proces oceny NPR/PWW jest stosowany wobec działań angażujących ok. 1/5 rocznych wydatków publicznych w Irlandii.

W poniższych podrozdziałach zostanie przedstawiony proces oceny szacunkowej (ewaluacja ex-ante), oceny bieżącej (ewaluacja interim) oraz oceny w połowie okresu realizacji (ewaluacja mid-term).

Ocena szacunkowa (ewaluacja ex-ante)

Obecne, zintegrowane podejście do procesu ewaluacji ma swoje źródło w procesie oceny szacunkowej przeprowadzonej w dwóch etapach¹⁶. W etapie pierwszym Departament Finansów ogłaszał przetarg na ocenę krajowych priorytetów dot. inwestycji na lata 2000-2006 dokonaną na podstawie ustalonych wcześniej celów ogólnych. Rozstrzygnięcie przetargu miało miejsce we wczesnej fazie (marzec 1998 r.) przygotowań NPR i było zbieżne z procesem konsultacji treści ww. dokumentu z partnerami regionalnymi, społecznymi i przedstawicielami poszczególnych sektorów. Zabieg ten miał na celu dostarczenie rządowi Irlandii niezależnej analizy eksperckiej dotyczącej potrzeb inwestycyjnych gospodarki krajowej. Częścią pracy ewaluatorów (Instytut Badań Gospodarczo-Społecznych) była bowiem także recenzja uwag dostarczonych przez różne grupy zainteresowanych. Podejście metodologiczne ujęte w raporcie – Krajowe Priorytety Inwestycyjne na lata 2000-2006 (opublikowane przez IBGS w marcu 1999 r.) – zostało opisane w następnym rozdziale.

Drugi etap procesu ewaluacji miał miejsce, gdy Narodowy Plan Rozwoju był już prawie gotowy. Aby spełnić wymagania zapisane w Artykule 41 rozporządzeń dot. ewaluacji

¹⁶ Zob. David Hegarty i John Fitz Gerald (ESRI), *Ex Ante Evaluation Process for the 200-2006 Period In Ireland*, Czwarta Europejska Konferencja nt. Ewaluacji Funduszy Strukturalnych, Edynburg, wrzesień 2000

szacunkowej, Jednostka Oceny PWW (obecnie Jednostka Oceny NPR/PWW) miał za zadanie przygotowanie formalnej oceny szacunkowej NPR. Raport *Ocena Szacunkowa Narodowego Planu Rozwoju na lata 2000-2006* był, w pewnym stopniu, kontynuacją analizy opracowanej uprzednio przez IBGS.

Ocena bieżąca (ewaluacja on-going, interim)

Jak wspomniano wyżej, praktyka przeprowadzania ocen bieżących - tj. mających miejsce między oceną szacunkową i oceną w połowie okresu realizacji, a następnie między oceną w połowie okresu realizacji a oceną pełną, po zakończeniu realizacji – była istotną cechą procesu ewaluacji w latach 1994-1999. Praktyka ta została zachowana także w obecnym okresie programowania, jednak wprowadzono pewne zmiany w strukturze organizacyjnej procesu. Proces ewaluacji bieżącej odbywa się obecnie na poziomie NPR/PWW i jest prowadzony przez Jednostkę Oceny NPR/PWW, a nie, jak to miało miejsce w latach 1994-1999, na poziomie poszczególnych programów operacyjnych (tj. z udziałem jednostek oceny umiejscowionych w strukturach poszczególnych instytucji zarządzających i/lub zewnętrznych ewaluatorów).

Na początku obecnego okresu programowania (mniej więcej do końca 2000 r., czyli do momentu, kiedy uzupełnienia programów operacyjnych zostały dostarczone do Komisji Europejskiej) prace Jednostki Oceny koncentrowały się na opracowywaniu wskaźników. Opierając się na wytycznych przygotowanych uprzednio, Jednostka prowadziła bieżące doradztwo dla instytucji zarządzających nt. wyboru i kwantyfikacji wskaźników¹⁷.

Kolejnym etapem w procesie było opracowanie przez Jednostkę Oceny planu działań w zakresie ewaluacji bieżącej na lata 2001-2003 (tj. do momentu rozpoczęcia oceny w połowie okresu realizacji). Plan ten, zakładający zrealizowanie 16 projektów ewaluacyjnych dot. NPR/PWW, został przyjęty przez Komitet Monitorujący Pomocy Technicznej NPR/PWW w maju 2001 r. (Komitetowi przewodniczy instytucja zarządzająca NPR/PWW, a w jego skład wchodzi przedstawiciele instytucji zarządzających poszczególnymi programami operacyjnymi oraz, jako doradcy, przedstawiciele Komisji Europejskiej). Ww. plan działań stawiał sobie za cel osiągnięcie w trakcie ewaluacji takich rezultatów, które, stanowiąc jednocześnie wkład do oceny w połowie okresu realizacji, stanowiłyby same w sobie cenne źródło informacji. Określona liczba projektów do realizacji została wybrana w oparciu o wnioski z ewaluacji szacunkowej. W ten sposób

¹⁷ Jednostka Oceny PWW, *CSF Performance Indicators: Proposals for 2000-2006 Programming Period*, październik 1999 (dostępne na stronie <http://www.eustructuralfunds.ie/publications/evaluation/index.htm>)

plan działań stanowił pomost między oceną przed rozpoczęciem realizacji i w połowie okresu realizacji.

Ww. plan działań był przedmiotem wielu korekt, tak, ze względu na opóźnienia w rozpoczęciu realizacji programów (co przyczyniło się do rezygnacji z części projektów) jak i opóźnień w procesie naboru kadry do Jednostki Oceny NPR/PWW. Ostatecznie w latach 2001-2003 miało miejsce 7 ewaluacji bieżących; 5 z nich zostało przeprowadzonych z pomocą zewnętrznych ewaluatorów, a dwie pozostałe wewnętrznie, przez Jednostkę Oceny NPR/PWW. Główny nacisk podczas ww. ocen był położony na aspekty związane z zarządzaniem i wdrażaniem programu, włączając w to kwestie takie jak proces wyboru i oceny projektów, zarządzanie projektami, wskaźniki, formułowanie działań, a także problemy związane z duplikacją. Zgodnie z procedurami zawartymi w PWW, raporty z ewaluacji były formalnie rozpatrywane przez komitety monitorujące właściwych programów operacyjnych. Procedury te zakładają bowiem, że każda instytucja zarządzająca ma obowiązek formalnie ustosunkować się (po konsultacjach z instytucjami wdrażającymi) do każdego zalecenia ewaluatorów i odpowiedź dostarczyć komitetowi monitorującemu. Następnie, instytucja zarządzająca jest zobligowana do regularnego raportowania komitetowi monitorującemu o postępach we wdrażaniu przyjętych rekomendacji. Wyżej opisane procedury stanowią ważne novum w systemie ewaluacji w porównaniu z poprzednimi okresami programowania, kiedy to raporty ewaluatorów rzadko były formalnie brane pod uwagę przez komitety monitorujące.

Ocena w połowie okresu realizacji (ewaluacja mid-term)

Proces oceny w połowie okresu realizacji ma miejsce obecnie. Ewaluacje są przy tym prowadzone zarówno na poziomie NPR/PWW, jak i programów operacyjnych. Korzystając z doświadczeń z lat 1994-1999, cały proces został wysoce skoordynowany. Aby to osiągnąć, w połowie 2002 r. została podjęta decyzja o utworzeniu grupy planistycznej ds. oceny w połowie okresu realizacji na poziomie NPR/PWW. W skład grupy wchodziły przedstawiciele instytucji zarządzających programami operacyjnymi oraz właściwych dyrekcji generalnych KE, a przewodniczy jej instytucja zarządzająca NPR/PWW. Jednostka Oceny NPR/PWW pełni funkcję sekretariatu ww. grupy. Grupa funkcjonuje także jako komitet sterujący w procesie oceny w połowie okresu realizacji NPR/PWW. Aby wypełnić swoje funkcje koordynacyjne, komitet przyjął harmonogram prac ewaluacyjnych i kluczowe zakresy usług dla celów ewaluacji prowadzonych na poziomie programów operacyjnych. W praktyce, ww. zakresy usług wymagały jedynie niewielkiego dostosowania ich treści do konkretnego przypadku, czego rezultatem było stosowanie

jednolitego podejścia na poziomie trzech krajowych i dwóch regionalnych programów operacyjnych. Zakres usługi dla oceny NPR/PWW w połowie okresu ich realizacji był swoistym uzupełnieniem dla tych funkcjonujących na poziomie programów operacyjnych. Został on opracowany przez Jednostkę Oceny NPR/PWW i przyjęty przez grupę planistyczną ds. oceny (komitet sterujący).

Podejście do ewaluacji i główne obszary zainteresowań

Niniejszy rozdział stanowi uzupełnienie poprzedniego, traktującego o strukturach ewaluacyjnych. Opisuje on rozwój w podejściu do tematu ewaluacji na poziomie PWW, który można zaobserwować w Irlandii. Kluczową kwestią przedmiotowego zagadnienia jest rosnąca rola analiz dotyczących przyczyn porażek różnego rodzaju inwestycji realizowanych w ramach PWW. Kwestia ta została formalnie zaliczona do pięciu głównych pytań stojących przed ewaluatorami w procesie oceny prowadzonej w obecnym okresie programowania.

PWW 1989-1993 i 1994-1999

Ewaluacja ex-post PWW 1989-1993, ewaluacja ex-ante PWW na lata 1994-1999

Pierwsza znacząca ocena na poziomie PWW w Irlandii została przeprowadzona przez Instytut Badań Gospodarczo-Społecznych w 1993 r.¹⁸. Raport z ww. oceny obejmował zarówno ocenę po zakończeniu realizacji PWW 1989-1993, jak i wstępną analizę priorytetów przewidzianych w PWW na lata 1994-1999¹⁹. W trakcie badań skorzystano z metod stosowanych zarówno w analizach mikro jak i makroekonomicznych. Tak skonstruowana metodologia była wykorzystywana również w przeprowadzonych później ewaluacjach. Wykorzystując średniookresowy model gospodarki irlandzkiej, w raporcie oszacowano ekonomiczne efekty w skali makro dla irlandzkiej gospodarki w latach 1989-1993. Autorzy raportu przestrzegali jednak przed ograniczoną wiarygodnością prezentowanych wyników, twierdząc, że „Analiza ta została oparta na serii mocnych założeń dotyczących prawdopodobnego zwrotu z różnego typu inwestycji. Pomimo iż dokonując analizy zgromadziliśmy szereg dowodów na poparcie naszych założeń, dowody te pozostają słabe i wyrywkowe. Z tego powodu wyniki analizy powinny być

¹⁸ ESRI, *The Community Support Framework – Evaluation and Recommendations for the 1994-1997 Framework, Final Report to the Department of Finance*, kwiecień 1993

¹⁹ Następnie, KE zleciła przeprowadzenie osobnej oceny NPR na lata 1994-1999

traktowane z odpowiednią rezerwą.” Prace ewaluacyjne i badania prowadzone w późniejszym okresie umożliwiły dokonywanie bardziej wiarygodnych szacunków, wzmacniając tym samym wiarygodność modeli obliczeniowych.

Metodologia wykorzystana do określenia priorytetów na lata 1994-1999 miała naturę bardziej mikroekonomiczną. Silny nacisk został położony na potencjalną stopę zwrotu z różnego rodzaju inwestycji. W raporcie stwierdzono, że zwroty z wielu projektów mogą być środkami prywatnymi oraz, że „jeśli zwrot środków prywatnych w ramach projektu jest wysoki, środki unijne powinny być wykorzystywane tylko do finansowania takich inwestycji, gdzie istnieje prawdopodobieństwo niedoskonałości rynkowych, co uchroni beneficjentów od ponoszenia tego typu kosztów ze środków własnych.” W przyszłości, co zostanie opisane w dalszej części tekstu, ewaluatorzy opracują bardziej wiarygodną metodologię badania kwestii niedoskonałości rynkowych. Następne zagadnienie wskazywane i podkreślane w kolejnych ewaluacjach związane było z potrzebą uświadomienia faktu, iż koszty alternatywne (opportunity costs) środków UE były takie same jak w przypadku analogicznych kosztów środków skarbu państwa. Uwzględnienie tego faktu pozwoliło wypracować właściwy punkt wyjścia do obliczania stopy zwrotu.

Ewaluacja mid-term PWW na lata 1994-1999

Organizacja procesu oceny w połowie okresu realizacji PWW (1994-1999) została opisana powyżej. Ocena ta, przeprowadzona przez IBSG, swoje źródło miała w ewaluacjach prowadzonych na poziomie programów operacyjnych. Raport ewaluacyjny był w tym przypadku ponownie oparty zarówno na metodologii makro, jak i mikroekonomicznej i kontynuował opisane powyżej podejście zastosowane w ocenie ex-post PWW 1989-1993. W szczególności, ewaluacja wykorzystywała innowacyjną metodologię dla mikroekonomicznej analizy działań realizowanych w ramach poszczególnych programach operacyjnych.

Analiza mikroekonomiczna wychodziła z założenia, że polityka publiczna powinna kierować się w stronę korygowania zakłóceń rynkowych. Ewaluatorzy utrzymywali, że zamiast po prostu osiągać przypisane cele, każdy program powinien być dokładnie sprawdzony pod kątem tego, czy przyczynia się do redukcji zakłóceń w stopniu wystarczającym, aby uzasadnić zaangażowanie większej ilości środków (w postaci wyższego opodatkowania). Podejście to zakładało, że każde działanie będzie uzasadnione w zakresie kosztów alternatywnych do środków publicznych.

W oparciu o te zasady, analiza mikroekonomiczna zaproponowała nową funkcjonalną klasyfikację działań. Działania związane z wydatkowaniem środków publicznych w ramach dziewięciu programów operacyjnych zostały ujęte w 4 kategorie²⁰:

- Działania związane z wydatkami na usługi określane mianem „*dobra publicznego*”, które powodowałyby zahamowanie ich dostarczania do sektora prywatnego.
- Działania ukierunkowane głównie na modyfikację cen względnych dla prywatnych firm i konsumentów indywidualnych w celu skorygowania wpływu czynników zewnętrznych; innymi słowy, działania znane jako *subsydiowanie korygujące*.
- Działania *adresowane do ściśle określonej grupy odbiorców* mające na celu zmianę sytuacji, w której podmioty prywatne są uważane za niewłaściwie poinformowane, lub gdy istnieją specyficzne okoliczności zewnętrzne czy bariera komunikacyjna.
- Subsytia, których główny efekt ma *redystrybucyjny* charakter.

Każde działanie i poddziałanie zostało przypisane do jednej z ww. kategorii. W wyniku tego zabiegu, w ramach każdej kategorii łatwiej było dokonywać porównań, nawet działań ujętych w różnych programach operacyjnych. W każdej kategorii ewaluatorzy zdefiniowali tzw. „działanie kotwice”, czyli pewien dobrze rozumiany schemat, który stanowił punkt odniesienia dla pozostałych. Dla każdej z 4 kategorii stworzono także listę kryteriów, według których wstępnie oceniono ponad 100 działań w ramach 9 programów operacyjnych. To wstępne badanie pozwoliło zidentyfikować działania, które budziły wątpliwości i poddać je bardziej wnikliwej ocenie. Efektem tych działań było wyłonienie 18 działań, w odniesieniu do których ewaluatorzy zakwestionowali przewidziany poziom alokacji finansowej oraz 8, które ich zdaniem wymagały większych nakładów.

²⁰ Więcej szczegółów nt. metodologii i systemu klasyfikacji w dokumencie: *Mid-Term Evaluation of the CSF for Ireland*, David Hegarty, (Jednostka Oceny PWW) i Patrick Honohan (IBGS) opracowanym na potrzeby Europejskiej Konferencji nt. Praktyk Ewaluacyjnych w zakresie Polityki Strukturalnej, Sewilla, marzec 1998 lub w opublikowanym raporcie z ewaluacji: Honohan., P., (ed), *Structural Funds In Ireland, A Mid-Term Evaluation of the CSF, 1994-1999*, lipiec 1999

NPR/PWW na lata 2000-2006

Uzgodnione podejście do procesu ewaluacji

Znaczącym postępowaniem w bieżącym okresie jest formalne przyjęcie podejścia do ewaluacji, składającego się z zestawu 5 głównych pytań ewaluacyjnych. Zestaw ten został zawarty w PWW, w których czytamy: „Wszystkie przeprowadzone oceny będą miały za zadanie odpowiedzieć na 5 podstawowych pytań dot. *Uzasadnienia, Odpowiedniości, Skuteczności, Efektywności i Oddziaływania.*” To podejście do procesu ewaluacji zostało zarekomendowane w opracowanym przez Jednostkę Oceny PWW, raporcie dot. roli oceny bieżącej w okresie programowania 1994-1999²¹. W raporcie stwierdzono, że głównym problemem związanym z systemem ewaluacji w latach 1994-1999 był brak wspólnego, jednolitego zrozumienia celu procesu oceny, wyrażający się w różnym podejściu reprezentowanym przez poszczególne programy operacyjne. Podejście zawarte w 5 podstawowych pytaniach pojawiło się w wyniku przeglądu literatury fachowej, ale także w oparciu o wcześniejsze prace ewaluacyjne prowadzone na poziomie PWW w Irlandii. Wpływ wcześniejszej pracy IBGS – w której akcentowano kwestię uzasadnienia finansowania poprzez odnoszenie się do niedoskonałości rynkowych – jest oczywisty ze względu na obecność w zestawie pytania dot. racjonalnego uzasadnienia. Pozostałe pytania dotyczące odpowiedniości (tj., jaki jest wpływ okoliczności zewnętrznych na program?), skuteczności (tj., czy program spełnia postawione przed nim cele?), efektywności (tj., czy koszty programu są współmierne do korzyści?) oraz wpływu (tj., jakie są społeczno-gospodarcze efekty działań?) brzmią zapewne znajomo dla praktyków ewaluacji i wynikają z zapisów regulacji prawnych dotyczących funduszy strukturalnych. Jeśli chodzi o wprowadzenie ww. podejścia w życie, stwierdzono, że nacisk na poszczególne pytania będzie kładziony w zależności od momentu, w którym dana ocena jest przeprowadzana. W PWW czytamy: „Punkt ciężkości przy przeprowadzaniu poszczególnych ocen będzie zależał od specyficznych okoliczności, włączając zaproponowaną szerokość spektrum, etap realizacji programu, w którym ewaluacja ma miejsce, ograniczenia czasowe, a także potrzeba badania konkretnych zagadnień.” W praktyce, takie podejście do ewaluacji wkomponowano w cykl programu w sposób następujący:

²¹ Jednostka Oceny PWW, Review of Ongoing Evaluation Function in the CSF for Ireland, 1994-1999, październik 1998 (dostępne na: <http://www.eustructuralfunds.ie/htm/publications/evaluation/index.htm>)

- Podczas przeprowadzania oceny szacunkowej, nacisk był położony na pytanie dot. uzasadnienia;
- W centrum uwagi podczas oceny bieżącej (włączając ocenę w połowie okresu realizacji i bieżące prace ewaluacyjne bezpośrednio przed tą oceną) znajdują się pytania związane z odpowiedniością, skutecznością oraz w mniejszym stopniu efektywnością i wpływem programu (z zachowaniem tej kolejności);
- Po dokonaniu oceny w połowie okresu realizacji, prace ewaluacyjne koncentrują się na efektywności i wpływie programu;

Przyjęcie takiego ujęcia ewaluacji zapewniło cenny punkt odniesienia w procesie oceny NPR/PWW. Zaakceptowanie takiego schematu przyczyniło się do podniesienia świadomości i zrozumienia celu prac ewaluacyjnych wśród zarządzających programami. Podstawowe pytania ewaluacyjne stworzyły podstawy zarówno dla tworzenia planów prac ewaluacyjnych, jak i zakresów usług dla poszczególnych projektów. Na przykład, zakresy usług dla oceny NPR/PWW i programów operacyjnych w połowie okresu realizacji formułują w sposób jasny i czytelny cele ewaluacji w zakresie efektywności programu i jego związku z okolicznościami zewnętrznymi.

Ocena szacunkowa

Jak zostało podkreślone wcześniej, ocena szacunkowa obecnego NPR/PWW została przeprowadzona w dwóch etapach; wstępna ocena potrzeb poprzedzała formalną ocenę szacunkową NPR. Rezultaty obydwu ocen zostały zawarte w dwóch oddzielnych raportach. Zgodnie z ustalonymi wcześniej zasadami, silny nacisk podczas obydwu ocen został położony na racjonalne uzasadnienie lub uzasadnienie niedoskonałości rynkowych dla różnego typu inwestycji.

Jak wspomniano w rozdziale „Ocena szacunkowa”, w 1998 r. Departament Finansów zlecił wykonanie niezależnej analizy potrzeb inwestycyjnych. Opracowany przez IBGS raport obejmował następujące etapy analizy²²:

- Opracowanie, na podstawie przykładu zarówno Irlandii, jak i innych państw UE, nt. roli różnego rodzaju inwestycji publicznych w promocji wzrostu gospodarczego.

²² Więcej szczegółów: *Ex Ante Evaluation Process for the 2000-2006 Period in Ireland*, David Hegarty (Jednostka Oceny PWW) i John Fitz Gerald (IBGS), dokument zaprezentowany podczas Konferencji Komisji Europejskiej nt. Ewaluacji Funduszy Strukturalnych, Edynburg, 2000; lub pełny raport: Fitz Gerald J., *National Investment Priorities Report for the Period 2000-2006*, ESRI Policy research Series, No. 33, marzec 1999 (streszczenie raportu na stronie www.esri.ie)

- Raport badający potencjalny wzrost gospodarczy i wskazujący główne ograniczenia tego wzrostu.
- Raport ustanawiający priorytety inwestycji przeznaczonych do finansowania w ramach nowego NPR, a także inne, nieinwestycyjne działania wymagające wsparcia.
- Opierając się na niniejszej analizie i funkcjonalnej klasyfikacji niedoskonałości rynkowych inwestycji wykorzystanych w ewaluacji średniookresowej (zob. rozdział „Ewaluacja mid-term PWW na lata 1994-1999”), ewaluatorzy opracowali szczegółowe rekomendacje nt. priorytetów inwestycyjnych niezbędnych do osiągnięcia celów Planu.

Jak wspomniano wcześniej, formalna ocena szacunkowa Narodowego Planu Rozwoju została przeprowadzona przez Jednostkę Oceny PWW jesienią 1999 r. Podczas oceny wyraźnie odniesiono się do schematu pięciu podstawowych pytań ewaluacyjnych. W szczególności nacisk został położony na racjonalne uzasadnienie lub uzasadnienie niedoskonałości rynkowych dla różnego rodzaju inwestycji zaproponowanych w Planie. Biorąc to pod uwagę, ewaluacja ta ponownie korzystała z metodologii wykorzystanej w połowie okresu realizacji PWW 1994-1999 (zob. rozdział „Ewaluacja mid-term PWW na lata 1994-1999”). Ocena koncentrowała się także na zagadnieniu związku poszczególnych, zaplanowanych inwestycji ze znaczącymi zmianami, które pojawiły się w społeczno-gospodarczym otoczeniu w ostatnich latach. Podczas oceny korzystano z wyników prac ewaluacyjnych przeprowadzonych w poprzednich okresach programowania jako dowodu na skuteczność, efektywność i wpływ konkretnych działań. Zgodnie z zakresem usługi, w raporcie powstrzymywano się przed zaleceniami dotyczącymi alokacji finansowych, koncentrując się na zagadnieniach budzących wątpliwości, stawianiu pytań i wyciąganiu ogólnych wniosków. Część kwestii poruszonych podczas oceny poddano ponownej ocenie w kontekście procesu oceny bieżącej, o którym była mowa w rozdziale „Ocena bieżąca”.

Wnioski

Na koniec warto podjąć kwestię korzyści płynących z ewaluacji. Czy ewaluacja jest uważana przez zarządzających programami za istotny element? W jakim stopniu wyniki ewaluacji przyczyniły się do kształtu programów współfinansowanych w ramach funduszy strukturalnych i alokacji środków? Czy ewaluacja spowodowała jakieś zmiany?

Są dowody na to, że wyniki ewaluacji wpłynęły zarówno na kształt programów, jak i alokację środków finansowych. Ocena w połowie realizacji PWW 1994-1999, omówiona wcześniej z punktu widzenia organizacji procesu oraz specyfiki zastosowanego podejścia, jest na to doskonałym przykładem. Jak stwierdzono w dokumencie przygotowanym na konferencję w Sewilli (cytowanym wcześniej), wnioski płynące z oceny przeprowadzonej w połowie okresu realizacji stanowiły znaczący wkład w średniokresowy przegląd postępów i wpłynęły na decyzje podjęte przez Komisję we współpracy z władzami krajowymi²³ odnośnie realokacji środków funduszy strukturalnych. W lipcu 1997 r. Komitet Monitorujący PWW przyjął pakiet średniokresowych zmian finansowych dotyczących kwoty 163 mln €, stanowiącej ok. 7,5 % wydatków planowanych na rok 1998 i 1999. Pakiet ten obejmował redukcję budżetu 8 z 18 działań, w których poziom zaangażowania środków poddany został w wątpliwość przez ewaluatorów. W 5 z 8 obszarów, które zostały ocenione przez ewaluatorów jako wymagające dodatkowych środków, środki te zostały przyznane. Jak zostało stwierdzone w dokumencie opracowanym na konferencję w Sewilli, wyniki ewaluacji PWW wpłynęły także na decyzje nie związane bezpośrednio z kwestiami finansowymi.

Wyczerpująca ocena szacunkowa obecnego NPR 2000-2006, o której była mowa powyżej, miała wpływ zarówno na kształt NPR, jak i PWW. Jak stwierdzono w dokumencie opracowanym na konferencję w Edynburgu, okazało się, że wstępna analiza priorytetów wpłynęła na szeroko zakrojoną debatę społeczną nt. priorytetów Planu i była przedmiotem debaty dot. kształtu NPR²⁴. W ww. dokumencie stwierdzono także, że „Jest rzeczą oczywistą, że raport ten miał ogromny wpływ na Plan oraz, że istota Planu pokrywa się z rekomendacjami ewaluatorów.” Rezultaty oceny szacunkowej stanowiły także istotny element w negocjacjach między Komisją i władzami krajowymi nt. kształtu PWW. Na przykład, Komisja podzieliła wątpliwość, która wyniknęła podczas oceny, nt. podziału środków w sektorze produkcyjnym (program operacyjny dotyczący inwestycji w badania i rozwój, przemysł, turystykę, rolnictwo i rybołówstwo) w ramach NPR oraz ograniczonego wsparcia w ramach PWW dla tego sektora ukierunkowanego raczej na konkurencyjność, a nie na tworzenie miejsc pracy.

Przeprowadzenie oceny bieżącej było istotnym elementem ewaluacji w Irlandii. Bieżące prace ewaluacyjne, które koncentrowały się na kwestiach związanych z zarządzaniem i wdrażaniem programów, były prowadzone w ciągu kilku ostatnich lat, i miały pozytywny oddźwięk u zarządzających programami chętnie współpracującymi z ewaluatorami.

²³ David Hegarty i Patrick Honohan, *ibidem*

²⁴ David Hegarty i Patrick Honohan, *ibidem*

Okazało się, że proces ewaluacji jest z ich perspektywy pożyteczny na tym wczesnym etapie (przed oceną w połowie okresu realizacji). Ponadto, podczas ostatniej oceny bieżącej dotyczącej inwestycji w ramach NPR/PWW związanych z siecią dróg ewaluatorzy zwrócili szczególną uwagę na znaczący wzrost kosztów przewidzianych na realizację programu dróg krajowych²⁵. Kwestia ta była szeroko komentowana w mediach i doprowadziła do wznowienia dyskusji nt. zarządzania finansowego w programie. Jeśli chodzi o wpływ omawianych ocen na zmiany programowe, jest jeszcze zbyt wcześnie, aby wyciągać wnioski. Jednym z celów tych ocen jest opracowanie wkładu do oceny w połowie okresu realizacji, który nie jest jeszcze gotowy.

Jest faktem, że wymagania wynikające z regulacji UE przyczyniły się do promocji kultury i rozwoju potencjału ewaluacyjnego w Irlandii. Jak wspomniano na początku, przed napływem dużej ilości środków w ramach funduszy strukturalnych na początku lat 90 prowadzono w Irlandii bardzo niewielką ilość prac ewaluacyjnych. Jednak, podobnie jak w przypadku kwestii związanych z zarządzaniem i wdrażaniem programów, Irlandia odniosła się pozytywnie do wymagań UE dot. ewaluacji. Podczas trzech okresów programowania, Irlandia wypracowała struktury ewaluacyjne i podejście do przedmiotowej kwestii, które znajdują obecnie zastosowanie w stosunku do działań, na które przeznaczane jest ok. 1/5 rocznych wydatków publicznych.

Ogólne doświadczenie irlandzkie koncentruje się na fakcie, że dobrze zorganizowany i posiadający odpowiednie zasoby system ewaluacji, połączony z właściwymi strukturami i jasno określonym celem działań, jest istotnym instrumentem w procesie maksymalizacji korzyści płynących z dostępności środków w ramach funduszy strukturalnych. Oceny, przeprowadzone w odpowiednim czasie przez doświadczonych i obiektywnych ewaluatorów i oparte na właściwych pytaniach oraz wspierane przez decydentów, mają duże znaczenie.

Bibliografia

Hegarty, D. and Honohan, P., *Mid-Term Evaluation of CSF for Ireland*, Dokument przygotowany na Europejską Konferencję na temat Ewaluacji Funduszy Strukturalnych w Sewilli, marzec 1998

²⁵ Fitzpatrick Associates, Evaluation of Investment in the Road Network, Sierpień 2002 (dostępne na stronie <http://www.eustructuralfunds.ie.htm/publications/index.htm>)

Hegarty, D. and Fitz Gerald, J., *Ex Ante Evaluation Process for the 2000 to 2006 Period in Ireland*, Dokument przygotowany na Czwartą Europejską Konferencję na temat Ewaluacji Funduszy Strukturalnych, Edynburg, wrzesień 2000

CSF Evaluation Unit, *CSF Performance Indicators: Proposals for 2000 – 2006 Programming Period*, październik 1999 (dostępne na:

<http://www.eustructuralfunds.ie/html/publications/evaluation/index.htm>)

ESRI, *The Community Support Framework – Evaluation and Recommendations for the 1994-1997 Framework*, *Final Report to the Department of Finance*, kwiecień 1993

Government of Ireland, *European Union Community Support Framework for Ireland*, NDP/CSF Information Office, Department of Finance, grudzień 2001 (dostępne na

<http://www.eustructuralfunds.ie/html/publications/official/index.htm>)

Honohan, P., (ed), *EU Structural Funds in Ireland, A Mid-Term Evaluation of the CSF, 1994 –99*, ESRI Policy Research Series Paper No. 31, lipiec 1999

Fitz Gerland J. et al (eds.), *National Investment Priorities Report for the Period 2000-2006*, ESRI Policy research Series, No. 33, marzec 1999 (streszczenie dostępne na:

www.esri.ie)

CSF Evaluation Unit, *Review of Ongoing Evaluation Function in the CSF for Ireland, 1994 –1999*, październik 1998 (dostępne na:

<http://www.eustructuralfunds.ie/html/publications/evaluation/index.htm>)

Fitzpatrick Associates, *Evaluation of Investment in the Road Network*, sierpień 2002

(dostępne na: <http://www.eustructuralfunds.ie/html/publications/evaluation/index.htm>)

David Hegarty

Był szefem Jednostki ds. Ewaluacji PWW w Irlandii, od jej utworzenia w 1996 r. aż do niedawna. Jednostka działa niezależnie pod egidą irlandzkiego Ministerstwa Finansów (IZ PWW). W Jednostce był on odpowiedzialny za proces koordynacji ewaluacji PWW w Irlandii oraz za wypełnienie wymogów stawianych przez rozporządzenie dotyczące funduszy strukturalnych UE w zakresie ewaluacji. Posiada bogate doświadczenie w zlecaniu i prowadzeniu ewaluacji na różnych etapach cyklu programowania (tj. ewaluacja ex-ante, interim, mid-term i ex-post).

Sandra Speer

Ewaluacja programów społeczno-gospodarczych – doświadczenia niemieckie

Wprowadzenie

Ewaluacja stała się rutynowym działaniem prowadzonym przez wiele centralnych rządów, a jej przeprowadzanie wpisane jest w wymogi dotyczące projektów finansowanych ze środków UE. Ewaluacja jest z natury działaniem politycznym, jak również ma na nią wpływ wiele czynników wynikających z narodowych procesów demokratycznych i administracyjnych. W porównaniu z krajami takimi jak Wielka Brytania, USA, czy też Szwajcaria, Niemcy nie cechują się silną kulturą ewaluacyjną. Jednakże, liczba przeprowadzanych ewaluacji wciąż wzrasta – zarówno publikowanych, jak i nie publikowanych – przy czym występuje wiele, bardzo różnych koncepcji, jak również zróżnicowana jest ich użyteczność.

Ewaluacja w perspektywie historycznej

Zgodnie z opracowaniem Wollmanna (2003) można wyróżnić trzy „fale” ewaluacji:

Pierwsza „fala ewaluacji” miała miejsce w latach 1960-70. Charakteryzowała się myśleniem w kategoriach cykli polityki związanych z triadą tworzenia i planowania polityki, jej wdrażania oraz ewaluacji. W tej sytuacji ewaluacja była traktowana bardzo instrumentalnie, tzn. służyła wyłącznie zbieraniu i przekazywaniu informacji zwrotnej. Jeżeli chodzi o Europę, to Niemcy były jednym z krajów na czele peletonu podczas opisywanej pierwszej fali.

W połowie lat 70-tych, gdy decydenci polityczni przypisywali szczególną wagę do oszczędności budżetowych oraz kryterium efektywności kosztowej (druga fala), znaczenie ewaluacji w Niemczech uległo osłabieniu. Po poprzednim boomie ewaluacyjnym nastąpił kryzys tej dziedziny, ponieważ wyniki ewaluacji, którymi były często studia na wielką skalę, były rzadko wykorzystywane przez niemieckich polityków.

Trzecia fala ewaluacji rozpoczęła się pod koniec lat 80-tych i dominowała w latach 90-tych, gdy na szeroką skalę wdrożono „Nowe publiczne zarządzanie” oparte na myśleniu w kategoriach zarządzania cyklami ustanawiania celów, wdrażania i ewaluacji. Trzecia fala

koncentruje się bardziej na wewnętrznej działalności ewaluacyjnej. Obecnie w Niemczech przeprowadza się wiele ewaluacji. Ten wewnętrzny charakter trzeciej fali jest w niektórych dziedzinach polityki kompatybilny z długą tradycją procedur samooceny i ich rozwoju w krajach niemieckojęzycznych. Samoocena jest tutaj rozumiana w wąskim znaczeniu: pracownicy programu (a nie pracownicy wewnętrzni lub zewnętrzni) kontrolują proces ewaluacji. Samoocena jest czasami częścią szerszych systemów zarządzania jakością. W celu promowania samooceny Niemieckie Ministerstwo Federalne (BMFSFJ) opublikowało cykl materiałów i poradników służących zapewnieniu właściwej jakości usług kierowanych do dzieci i młodzieży, zwracając szczególną uwagę na samoocenę. Wsparcie federalne dla samooceny w tym obszarze polityki zostało wstrzymane przed około pięciu laty i obecnie nie są podejmowane żadne inicjatywy w celu jego zrekomensowania, ani na szczeblu rządu federalnego, ani na poziomie lokalnym. Koncepcja „samooceny” została najpierw wprowadzona do wielu dziedzin polityki społecznej i edukacji, po czym rozpoczęto realizację projektów ewaluacyjnych o charakterze zewnętrznym.

Dzisiejsza pozycja ewaluacji w Niemczech

Od lat 90-tych opublikowano w Niemczech wiele książek i artykułów. Opracowano również techniczny słownik dla potrzeb ewaluacji w Niemczech²⁶. W roku 1997 utworzone zostało Deutsche Gesellschaft für Evaluation (Niemieckie Towarzystwo Ewaluacyjne). Od roku 2002 wydawane jest dwa razy w roku niemieckojęzyczne czasopismo poświęcone ewaluacji „Zeitschrift für Evaluation”. W porównaniu do Wielkiej Brytanii, ewaluacja jako meta-dyscyplina jest nadal na wczesnym etapie rozwoju, a segmentacja dyscyplinarna jest bardzo wysoka. Dla ogółu społeczeństwa pojęcie „ewaluacja” często wiąże się z ewaluacją działalności edukacyjnej w szkołach i na uniwersytetach.

Instytucjonalizacja ewaluacji w Niemczech jest bardzo niska. Jedyne ministerstwem federalnym, które posiada centralną jednostkę ewaluacyjną jest Ministerstwo Współpracy Gospodarczej i Rozwoju (BMZ). Ogłoszono, że również Federalne Ministerstwo Gospodarki i Pracy (BMWA) stworzy centralną jednostkę ewaluacyjną, ale wybory, które odbędą się we wrześniu 2005 r. mogą zmienić politykę w tym zakresie. Potencjał ewaluacyjny jest szeroki, ale nie występuje intensywne wymiana doświadczeń. Niekiedy dla zamawiających ewaluacja jest nową (meta-) dyscypliną i nie stanowi ona części ich podstawowych praktyk i procedur.

²⁶ Patrz Struhkamp 2005; glosariusz patrz np. Das A-B-C der wirkungsorientierten Evaluation, Glossar Deutsch – Englisch. <http://www.univation.org/glossar/index.php>

Obecnie, oprócz wymogu przeprowadzania ewaluacji dla programów finansowanych przez UE, niektóre nowsze ustawy zawierają obowiązek przeprowadzania ewaluacji, np. reforma rynku pracy, tzw. reforma Hartza. Jeżeli miasta lub gminy chcą przeprowadzić eksperymenty (na małą lokalną skalę) – jeśli prawo krajowe dopuszcza taką możliwość – działanie takie zazwyczaj obejmuje obowiązek przeprowadzenia ewaluacji (tzw. klauzule eksperymentalne).

W celu pokazania ram ewaluacji funduszy strukturalnych, w niniejszym opracowaniu przedstawiony zostanie zarys koncepcji ewaluacyjnych stosowanych i omawianych w Niemczech, a także wymogi odnoszące się do ewaluacji oraz regulacje w zakresie dostępu do danych.

Ewaluacja programów społeczno-gospodarczych w Niemczech

Metodologia ewaluacji ma swoje korzenie w Ameryce Północnej, gdzie teoretycy opracowali wiele różnych podejść do ewaluacji. Każdego roku, antologie i podręczniki wprowadzają nowe warianty lub całkowicie nowe podejścia. Natomiast do chwili obecnej opracowano w Europie niewiele podejść odnośnie ewaluacji, za wyjątkiem „modelu nieukierunkowanego na cel” Vedunga (1998) i „realistycznej ewaluacji” Pawsona i Tilleya (1997). Poszczególne podejścia ewaluacyjne znacząco się różnią jeżeli chodzi o ich podstawy epistemologiczne, identyfikację i uwzględnienie wartości społecznych, koncentrację na określonych elementach programu poddawanego ewaluacji oraz inne kwestie. Zwłaszcza programy społeczno-gospodarcze charakteryzują się często napięciami w odniesieniu do wartości etycznych.

Różne podejścia do ewaluacji

W Niemczech, istnieje tradycja szczegółowych opisów – głównie poprzez innowacyjne podejścia naukowe – „społecznej dezintegracji”, „wykluczenia” czy też „ubóstwa”. W poprzednich dekadach pojawiły się np. „zmiany wskaźników społecznych”, „badania porównawcze dotyczące bezrobotnych”, co zostało uwzględnione w raportach dotyczących ubóstwa sporządzanych przez związki zawodowe i organizacje charytatywne, oraz w materiałach dotyczących dobrobytu i ubóstwa publikowanych przez rząd federalny. Doświadczenia w zakresie programów zorientowanych na wyniki i ich ewaluacje występują rzadko, przy czym ostatnio stosowane są innowacyjne podejścia do ewaluacji, takie jak podejścia wykorzystane w programie Federacji i krajów związkowych „Społeczne miasto” czy też w kontekście reformy systemu opieki społecznej.

W przypadku nasilonych problemów społecznych, celem powinno być wzmocnienie rozwiązań społecznych i korporacyjnych na rzecz poprawy i stabilizacji sytuacji. Ewaluacje zorientowane na wyniki mogą przyczynić się do rozwoju tego aspektu. Zwłaszcza w kontekście społeczno-politycznym, ewaluacje powinny przykładać się do rozwiązywania konfliktów, powstających często na tle celów polityki zwalczania ubóstwa oraz podejmowanych w jej ramach działań. Zatem, w ramach zakresu ewaluacji należy określić, która instancja będzie odpowiedzialna za określenie wartości oraz stopnia, w jakim stanowiłyby one integralną część samego procesu ewaluacji.

Niemieckie Federalne Ministerstwo Zdrowia i Opieki Społecznej (BMGS) zleciło perspektywiczne studium dotyczące „ewaluacji zorientowanej na wyniki dotyczącej zasięgu dobrobytu i ubóstwa” (Beywl/Speer/Kehr 2004).²⁷ To perspektywiczne studium koncentrowało się na ewaluacjach, w ramach których oceniane są programy, zadania i polityki, zwłaszcza w obszarze zapobiegania i zmniejszania ubóstwa, takie jak np. polityka rynku pracy, polityka społeczna, usługi dla dzieci i młodzieży. Głównym celem tego studium było stworzenie podstawy dla włączenia ewaluacji w niemiecki system sprawozdawczości dotyczący narodowego dobrobytu i ubóstwa. Również w jedenastym raporcie narodowym dotyczącym sytuacji młodych ludzi znaczenie ewaluacji podkreślone zostało w kilku paragrafach i dodatkowo poświęcony został temu tematowi osobny rozdział.

W tym miejscu należy podkreślić, że ewaluacja wyników powinna stanowić część systemu sprawozdawczości społecznej dotyczącej sytuacji młodych ludzi. Nie przeprowadzono żadnej syntezy ani analizy ewaluacji. Włączenie wyników poszczególnych ewaluacji krajowych – finansowanych ze środków unijnych i ze środków krajowych – do sprawozdawczości społecznej mogłoby prowadzić do lepszego zrozumienia rezultatów programu, jak również ich znaczenia w państwie opiekuńczym. Do chwili obecnej ewaluacje programów funduszy strukturalnych w Niemczech zazwyczaj znajdują swoje odzwierciedlenie w ramach „środowiska funduszy strukturalnych”, a sporadycznie np. w innych departamentach ministerstw.

Końcowy raport perspektywicznego studium (Beywl/Speer/Kehr 2004) dotyczącego ewaluacji w kontekście sprawozdawczości na temat dobrobytu i ubóstwa liczy 251 stron i składa się z wyszczególnionych poniżej pięciu rozdziałów (raport jest dostępny jedynie w języku niemieckim):

²⁷ Zwane następnie „studium perspektywicznym”.

1. Stworzenie teorii ewaluacji
2. Modele ewaluacyjne
3. Wymogi odnoszące się do ewaluacji w obszarach polityki likwidacji ubóstwa
4. Udostępnianie danych na potrzeby ewaluacji
5. Perspektywy i rekomendacje

Zgodnie z postępowaniem w zakresie rozważania wartości społecznych w planowaniu i prowadzeniu ewaluacji, można wyróżnić cztery główne typy podejść do ewaluacji. W perspektywnym studium dwanaście omówionych modeli ewaluacyjnych zostało przypisanych do powyżej wskazanych czterech typów.

- Ewaluacja obiektywna w odniesieniu do wartości (Value-distant evaluation) wyklucza osąd wartościujący i traktuje wartości jako dane ustalone z góry w drodze procedur demokratycznych, które nie wchodzą w zakres ewaluacji. Specjaliści ds. ewaluacji powinni dostarczyć obiektywnych i wiarygodnych danych, rezultatów i wniosków odnoszących się do faktów społecznych, zgodnie ze ścisłymi zasadami. Ewaluacje takie są oparte na: (1) celu programu (2) projekcie eksperymentalnym (3) projekcie quasi-eksperymentalnym (4) analizie kosztów i korzyści dla programu (5) kontekście i mechanizmie lub (6) teorii programu.
- Ewaluacja relatywizująca wartości (Value relativistic evaluation) koncentruje się na wartościach społecznych, planowaniu, wdrażaniu i stosowaniu. Bez zajmowania stanowiska, taki szczególny typ ewaluacji wzmaga istniejące konflikty. Specjaliści ds. ewaluacji wraz z pracownikami zajmującymi się programem omawiają wartości tak, aby wywołać motywację i energię do aktywnego zaangażowania i wykorzystania rezultatów ewaluacji.
Na ewaluację relatywizującą wartości wpływ mają: (7) kwestie wrażliwe (sensitive issues) / delikatne tematy (tender subjects) (8) dialogi.
- Ewaluacja hierarchizująca wartości (Value prioritising evaluation) dąży do uszeregowania poszczególnych wartości. Głównym celem jest przedstawienie szeroko akceptowanych wartości jako podstawy dla ewaluacji i jej rezultatów.
Omówione i akceptowane wartości są warunkiem koniecznym dla odpowiednich pytań ewaluacyjnych, danych, analizy i wniosków. Równocześnie można dopuścić, aby udział poszczególnych grup i ich wpływ na tworzenie priorytetów był różny, w zależności od faktycznego rozłożenia władzy.

Ewaluacje takie koncentrują się na: (9) decyzjach (10) wykorzystaniu (11) i interesach zaangażowanych stron.

- Ewaluacja nakierowana na wartości (Value positioned evaluation) zakłada, że systemy społeczne charakteryzują się nierównowagą i dysproporcją władzy. Poprzez zajęcie określonego stanowiska, ewaluacja taka zakłada działanie jako przeciwwagę do istniejącej hegemonii wartości na płaszczyźnie politycznej, społecznej i kulturowej w celu wzmocnienia pozycji grup nieuprzywilejowanych. Ewaluacja osadzona na wartościach jest oparta na (12) samoorganizacji (empowerment evaluation).

Tabela: Modele ewaluacji i sposób w jaki uwzględniają one wartości

Odniesienie do wartości	Model ewaluacji skoncentrowany na:	Synonimy
1. Obiektywne	Celach	Studia efektywności
	Projekcie eksperymentalnym	Gold standard
	Projekcie quasi-eksperymentalnym	Ewaluacje ekonometryczne
	Analizie kosztów i korzyści	Efektywność kosztowa
	Mechanizmie kontekstu	Ewaluacja realistyczna
	Teorii programu	Ewaluacja oparta na teorii
2. Relatywistyczne w stosunku do wartości	Kwestiach (<i>Issue focused</i>)	Reagujący
	Dialogu	Konstruktywny / czwartej generacji
3. Hierarchizujące wartości	Decyzjach	Zorientowany na odpowiedzialność
	Wykorzystaniu	Pragmatyczny
	Interesach zainteresowanych stron	Przemyślany, demokratyczny
4. Zorientowane na wartości	Samoorganizacji	Ewaluacja upodmiotawiająca (<i>Empowerment evaluation</i>)

Na podstawie Beywl/Speer/Kehr 2004.

Każdy model ewaluacji był wykorzystywany w Niemczech. Dla większości z powyżej przedstawionych dwunastu modeli ewaluacji, można zidentyfikować i pokrótce przedstawić kilka przykładów zastosowania obejmujących kwestie dobrobytu i ubóstwa w Niemczech. Wynik analizy aktualnych niemieckich raportów ewaluacyjnych jest niezadowolający. Standardy metodologiczne, zwłaszcza ewaluacji przeprowadzanych na małą skalę znacznie się różnią i czasami budzą wątpliwości. W większości przypadków

nie przedstawiono podstawowych zasad teorii ewaluacji. W niewielkim stopniu omówiono w raportach ewaluacyjnych odniesienie do wartości i ich planowane wykorzystanie.

W ramach ewaluacji złożonych programów często łączy się różne podejścia do ewaluacji. Ma to szczególne znaczenie dla programów UE w ramach funduszy strukturalnych, gdyż nie tylko same programy są złożone, ale również cele i pytania ewaluacyjne wymagają zróżnicowanego podejścia. W odniesieniu do ewaluacji europejskich funduszy strukturalnych można zaobserwować tendencję do ewaluacji obiektywnej w odniesieniu do wartości i ewaluacji podsumowującej, ale nie należy przy tym pomijać funkcji dialogu. Czasami elementy partycypacyjne są wprowadzane do schematu ewaluacji, ale zazwyczaj mają one niewielkie znaczenie dla procesu ewaluacji (Toepel/Schwab 2005). Podejście do ewaluacji oparte na teorii nie jest zbyt często wykorzystywane, mimo że może być bardzo użyteczne w ewaluacji programów wielowarstwowych realizowanych na poziomie krajowym, jak również lokalnym, czy też na poziomie partnerstw. Odbiór takiego podejścia do ewaluacji jest jednak nadal niezadowolający.

Obecnie w Niemczech szeroko omawiane są możliwości i ograniczenia dla ewaluacji jeżeli chodzi o mierzenie rezultatów osiąganych w programach i ich odpowiednie zorientowanie na wartości. Powyżej omówione ramy podejść mogą zostać wykorzystane do analizowania ewaluacji. Kolejna konferencja Niemieckiego Towarzystwa Ewaluacyjnego zostanie, w szczególności, poświęcona „mierzeniu rezultatów”.²⁸

Obok dyskusji dotyczących istniejących modeli ewaluacji oraz ich zróżnicowanych umiejętności radzenia sobie z pomiarem rezultatów jak również zróżnicowanego zorientowania na wartości, nowoczesne programy polityczne wymagają zastosowania nowych metod ewaluacji. Programy UE są często przygotowywane w drodze współpracy i partnerstwa – często podawanym przykładem jest program EQUAL. Również niemiecki rząd promuje tworzenie sieci wymiany doświadczeń – np. w ramach sektora weterynaryjnego i MŚP. Charakter takich sieci i ich dynamika znacznie się różnią od programów nakierowanych na poszczególne osoby. Agregacja rezultatów ewaluacji sieciowych jest szczególnie dużym wyzwaniem. Na początku bieżącego roku zorganizowana została konferencja dotycząca „ewaluacji sieciowych”, niedługo zostaną wydane publikacje pokonferencyjne.²⁹

²⁸ DeGEval-Jahrestagung: „Wirkungen auf der Spur”, 12.-14.10.05 w Essen, Niemcy.

²⁹ Tagung „Netzwerk-Evaluation”, 27./28.01.05 w Kolonii.

<http://www.univation.org/index.php?class=artikel&id=174&PHPSESSID=ee678abec7443b9bda27a44b0706f8dd>

Wymagania stawiane ewaluacji wysokiej jakości

Wymagania stawiane ewaluacji wysokiej jakości wyrażane są na kilka sposobów: przez podmioty zlecające jak i ze strony środowiska ewaluatorów.

Cechy charakterystyczne dobrej ewaluacji, zidentyfikowane w perspektywnym studium i w modelach ewaluacji, zostały zestawione z wymaganiami krajowych ekspertów zajmujących się ewaluacją dotyczącą dobrobytu i ubóstwa. Zostało to zrealizowane w oparciu o wywiady z doświadczonymi naukowcami i grupami fokusowymi obejmującymi urzędników z różnych ministerstw federalnych i z ministerstw krajów związkowych. Do tego czasu komunikacja w zakresie ewaluacji pomiędzy poszczególnymi ministerstwami była bardzo słaba lub też wcale jej nie było. Podobne obserwacje poczynili również ewaluatorzy EFS.

- Wielu ekspertów popiera postępowanie partycypacyjne i pluralistyczne w obszarze ewaluacji, włączając w to dobór odpowiednich ewaluacyjnych modeli i postępowań w zależności od celu i pytań zawartych w umowie o przeprowadzenie ewaluacji. Ze względu na postępowania zorientowane tematycznie, oczekują one zasadniczo realistycznych rezultatów i inicjatyw na rzecz wzmocnienia programów i polityk.
- Zgodnie z ich punktem widzenia, systematyczna weryfikacja oddziaływania programów społeczno-politycznych jest pierwszoplanowym zadaniem ewaluacji. Równocześnie eksperci wskazują na problemy natury metodologicznej i koncepcyjnej (problem wyjaśniania zależności przyczynowo-skutkowej pomiędzy interwencją a rezultatem), które są trudno rozwiązywalne i powiązane są z działaniami operacyjnymi, bazą danych i projektem badania. Proponowane rekomendacje odpowiadają znaczeniu poszczególnych modeli ewaluacyjnych.
- Wielu ekspertów przewiduje szansę wykorzystania ewaluacji jako narzędzia planowania i zarządzania przy wstępnym omawianiu, przygotowywaniu i wdrażaniu programów, co zapewni skuteczność i adekwatność odpowiednich środków, które pozwalają na zwalczanie ubóstwa. Rekomendują oni stosowanie modelowania, identyfikacji i weryfikacji rezultatów przed rozpoczęciem oraz w trakcie pierwszego etapu realizacji, celem wpisania ewaluacyjnego know-how w cały cykl i politykę programu.

- Kwestia, w jakim stopniu ewaluacja może przyczynić się do sprecyzowania celów programu i jego perspektyw, wynikających z interpretacji i wniosków jest przedmiotem kontrowersji.
- Występuje ogromne zainteresowanie wykorzystaniem obecnie prowadzonych rozmów na temat teorii ewaluacji, jak również praktycznych rezultatów ewaluacji, w celu uzyskania wiarygodnych, przynoszących korzyści i precyzyjnych metodologicznie ewaluacji. Ogólnie rzecz biorąc, partnerzy dialogu liczą na korzystny wkład do wdrażania strategii nastawionej na zmniejszanie, czy też likwidację ubóstwa. W tym miejscu, należy zaangażować politykę, naukę i społeczeństwo.
- Jednym z głównych celów jest opublikowanie studiów ewaluacyjnych wkrótce po ich zakończeniu, w celu zwiększenia ich wkładu w doprecyzowanie celów i strategii oraz dla dalszego rozwoju ewaluacji jako koncepcji metodologicznej. Zintegrowane analizy badań ewaluacyjnych są uważane za istotne, zwłaszcza w przypadku rozbieżnych rezultatów indywidualnych badań.

Niemieckie Towarzystwo Ewaluacyjne opublikowało „Standardy przeprowadzania ewaluacji” (2002) oraz „Rekomendacje odnośnie stosowania Standardów w kontekście samooceny” (2004a). W najbliższym czasie standardy te zostaną poddane przeglądowi i ewentualnej weryfikacji. Obecnie są one niekiedy umieszczane w specyfikacji istotnych warunków zamówienia.

Fragment aktualnej specyfikacji istotnych warunków zamówienia stosowanej przez niemieckie Federalne Ministerstwo Spraw Wewnętrznych (tłumaczenie):

...

metoda ewaluacji:

ewaluacja powinna być oparta na „Standardach wspólnego komitetu”

- standardy dot. wykorzystania,
- standardy dot. wykonalności,
- standardy dot. etyki badawczej,
- standardy dot. dokładności.

Wybór wnioskodawców dokonywany będzie w oparciu o wspomniane powyżej kryteria oraz dodatkowo z uwzględnieniem kryteriów kwalifikacyjnych uwzględniających wytyczne Amerykańskiego Towarzystwa Ewaluacyjnego („Guiding Principles for Evaluators”³⁰). W wytycznych tych zawarte są następujące oczekiwania wobec ewaluatorów:

- systematyczny wywiad (Systematic Inquiry),
- kompetencje: wiedza i umiejętności przeprowadzania ewaluacji (muszą zostać potwierdzone) (Competence);
- zapewnienie uczciwości i integralność całego procesu ewaluacji (Integrity/Honesty);
- szacunek dla ludzi – poszanowanie godności i prywatności uczestników ewaluowanych programów (Respect for People),
- odpowiedzialność za cele rozpatrywanego programu z uwzględnieniem interesów wszystkich grup odbiorców programu. (Responsibilities for General and Public Welfare);

Po drugie, Niemieckie Towarzystwo Ewaluacyjne opublikowało „Rekomendacje dotyczące szkoleń i edukacji w dziedzinie ewaluacji – wymogi dla ewaluatorów” (2004b), są one szczególnie ważne, ponieważ rynek szkoleń w zakresie ewaluacji w Niemczech jest nadal bardzo młody, a kształcenie w zakresie ewaluacji na uniwersytetach jest bardzo heterogeniczne.

Federalne Ministerstwo Gospodarki i Pracy (BMWA 2004) utworzyło zespół ekspertów w celu sformułowania wytycznych dla ewaluacji. Obecnie następujące wytyczne są stosowane w ramach BMWA:

- Wytyczne dla ewaluacji szacunkowej (ex ante) programów w ramach BMWA
- Wytyczne dla procesu i ewaluacji ex post programów badawczych i rozwoju
- Wytyczne dla ewaluacji systemowych instytucji i władz
- Wytyczne dla ewaluacji w kontekście programów MŚP
- Wytyczne dla ewaluacji w kontekście programów rynku pracy

³⁰ <http://www.eval.org/Guiding%20Principles.htm>

Można stwierdzić, że wytyczne dla ewaluacji skupiają się wokół kwestii celów, skuteczności, efektywności programów i ich sytuacji względem faktów. Kwestia wartości społecznych nie odgrywa jednak znaczącej roli.

Podsumowując, oprócz wyczerpującego materiału dotyczącego ewaluacji opublikowanego przez Komisję UE, również na szczeblu krajowym wypracowanych szereg standardów i wytycznych.³¹ Jednakże kwestią otwartą nadal pozostaje ich systematyczne wykorzystywanie w praktyce.

Wymagania w stosunku do danych wykorzystywanych w procesie ewaluacji

W ramach zakresu programów i polityk nakierowanych na walkę z ubóstwem pojawia się pytanie, w jaki sposób udoskonalona ostatnio baza danych może zostać efektywnie i zgodnie z kanonami metod naukowych wykorzystana w ewaluacji zorientowanej na wyniki. Analiza literatury oraz wywiady z ekspertami zostały wykorzystane do przygotowania odpowiedzi w ramach perspektywicznego studium.

Ogólnie rzecz biorąc, przychylnie traktowana jest dostępność istniejących źródeł danych, jednak pożądana jest poprawa baz danych oraz integracja różnych typów danych.

- Koszt uzyskania danych administracyjnych jest niski. Mogą one być często wykorzystywane w kontekście regionalnym w celu opisanego wyjściowej sytuacji grup docelowych, ogólnych warunków konkretnego programu, jak również w celu umożliwienia porównań międzyregionalnych. Dla uzyskania istotnych wniosków, konieczna jest jednak precyzyjna wiedza na temat powstawania procesu i ograniczeń w wykorzystywaniu danych administracyjnych. Zastosowanie standardowych metod elektronicznego przetwarzania danych (EDP), które mogłyby generować dane na poziomie lokalnym, jest stanem pożądanym. Jest to szczególnie istotne w przypadku Niemiec, gdyż system federalny stwarza odmienne praktyki w poszczególnych krajach związkowych, jak również na szczeblu lokalnym. Dla lepszego wykorzystania danych administracyjnych zaleca się połączenie wybranych, istniejących źródeł danych, w celu zwiększenia bazy danych i dostarczenia danych pionowych.
- Centralnie skonsolidowane dane statystyczne dają podstawę porównań programów krajowych i federalnych, i powinny być intensywnie wykorzystywane. Pożądanym jest istnienie sekcji pionowych dla danych administracyjnych, jak

³¹ Odnośnie porównania różnych krajowych i europejskich wytycznych patrz Beywl/Speer (2004).

również dla danych statystycznych. Ponieważ określone podgrupy osób dotkniętych lub zagrożonych ubóstwem nie mogą zostać wystarczająco precyzyjnie zidentyfikowane, dane statystyczne powinny zostać poszerzone o dane administracyjne oraz dane panelowe.

- Dane panelowe są przywoływane jako istotne odniesienie dla badań opartych na procesach. Jednakże, ze względu na relatywnie małą liczbę badanych grup, możliwości wnioskowania w oparciu o wyniki tych badań są ograniczone. W przypadku większych funduszy można uzyskać proporcjonalnie większą liczbę prób oraz zredukować liczbę istniejących ograniczeń dotyczących odpowiedzi na określone pytania ewaluacyjne.

W 2001 r. utworzona została komisja ds. poprawy infrastruktury informacyjnej pomiędzy nauką i statystyką (Kommission zur Verbesserung der informationellen Infrastrukturen zwischen Wissenschaft und Statistik) przez niemieckie Federalne Ministerstwo Edukacji i Badań (BMBF). Opublikowało ono zalecenia dotyczące następujących tematów:

- konieczność lepszej współpracy nauki i statystyki dla użytkowników danych i producentów danych
- zalecenia dotyczące zaangażowania nauki w programowanie
- zalecenia dotyczące kontynuowania jak również poszerzenia istniejących statystyk
- zalecenia dotyczące badań związanych z wytwarzaniem danych, ich prezentacją i zapisem
- zalecenia dotyczące kształcenia i szkolenia
- zalecenia dotyczące dostępu do danych
- zalecenia dotyczące dopasowania danych
- zalecenia dotyczące ram prawnych

Poprawa dostępu do danych jest często warunkiem koniecznym dla niektórych modeli ewaluacyjnych. Zwłaszcza ewaluacje quasi-eksperymentalne wymagają istnienia rozbudowanych baz danych, które nie mogą zostać stworzone przez ewaluatorów. Mimo, że komisja nie zajmuje się w szczególności sposobem ewaluacją, to jej raport i zawarte w nim zalecenia mają istotne znaczenie dla ram ewaluacji.³² Ośrodki statystyczne krajów związkowych, jak również federalny urząd statystyczny utworzyły niedawno ośrodki

³² Obecnie komisja ds. infrastruktury danych została przekształcona w stałą radę ds. danych społeczno-gospodarczych, <http://www.ratswd.de/home.htm>.

badania danych oraz ośrodki usługowe, których prace mogą być również wykorzystane na potrzeby ewaluacji.

Poprawiony dostęp do danych dotyczących rynku pracy w Niemczech pozwala na przeprowadzanie ewaluacji ekonometrycznych, tak jak to zostało zrealizowane w ramach ewaluacji reformy Hartza. Dane dotyczące rynku pracy mogą się w chwili obecnej charakteryzować najlepszym dostępem i najlepszą możliwością powiązań pomiędzy różnymi zestawami danych, ponieważ centralna agencja IAB administruje większością zestawów danych. Ocena reformy Hartza jest najbardziej rozległą ewaluacją w ramach jednego obszaru polityki w Niemczech. Na potrzeby przeprowadzenia tej ewaluacji po raz pierwszy wybrano dwuetapową procedurę zamówień: najpierw wybrano równoległe dwa konkurencyjne projekty ewaluacyjne, a w kolejnym roku wybrano pojedyncze zestawy. Ewaluacja tej reformy ma charakter ewaluacji uzupełniającej i z pewnością będzie wywierać wpływ na niemiecką kulturę ewaluacyjną i jej rozwój. Przyszłość pokaże w jakim stopniu ewaluacja reformy Hartza zostanie wykorzystana i w jakim stopniu przyczyni się do stworzenia kształcenia instytucjonalnego.

Poprawiony dostęp do danych statystycznych pomaga w przeprowadzeniu ewaluacji EFS, na obszarze analiz kontekstu programu. Ułatwiony dostęp do danych przekłada się także bezpośrednio na poprawę jakości monitoringu EFS (Jaedicke 2005), jednocześnie zagrożeniem dla tego stanu rzeczy jest uproszczenie procedur biurokratycznych, co może skutkować powrotem do sytuacji z roku 2000.

Podsumowanie

Spółeczność ewaluacyjna w rozszerzającej się Europie wymaga wyraźnej tożsamości i samoświadomości. Nie jest to łatwe zadanie dla młodej, nadal kształtującej się europejskiej społeczności ewaluacyjnej, która jest sama w sobie międzynarodowa i międzykulturowa. Występuje również rywalizacja pomiędzy poglądami środowiska a poglądami poszczególnych praktyków związanych z różnymi dyscyplinami.

Kultura ewaluacyjna w Niemczech powoli dojrzewa. Ramy dla ewaluacji europejskich funduszy strukturalnych oraz ich głównych czynników oddziaływania zostały streszczone. W Niemczech dyskusja dotycząca teorii ewaluacji została ostatnio pogłębiona. Wymogi płynące z ministerstw stały się dużo bardziej konkretne. Zapoczątkowana została poprawa dostępu do danych. Segmentacja dyscyplinarna jest nadal wysoka i może również stanowić problem przy realizacji ewaluacji funduszy strukturalnych.

Wykorzystanie ewaluacji mogłoby zostać wzmocnione poprzez połączenie ich z federalnym systemem sprawozdawczości społecznej. Warunkiem koniecznym byłoby przeprowadzenie syntezy ewaluacji. Jej wykorzystanie mogłoby zostać wówczas poprawione na dwóch poziomach: po pierwsze, poprzez podniesienie użyteczności bezpośrednich wyników ewaluacji w odniesieniu, zarówno do badanego programu, jak i innych projektów wchodzących w ten sam obszar polityki publicznej (niez względu na ich źródło finansowania), po drugie poprzez przyspieszenie procesu uczenia na poziomie ogółu działań ewaluacyjnych.

Bibliografia

Beywl, Wolfgang/Speer, Sandra/Kehr, Jochen (2004), Wirkungsorientierte Evaluation im Rahmen der Armut- und Reichtungsberichterstattung, Bundesministerium für Gesundheit und Soziale Sicherung (BMGS), Bonn. www.bmgs.bund.de/download/broschueren/A323.pdf

Beywl, Wolfgang/Speer, Sandra (2004), Data- and Literature-Based Reflections on Western European Evaluation Standards and Practices, w: New Directions for Evaluation, Vol. 104, str. 43-54.

Bundesministerium für Wirtschaft und Arbeit (BMWA) (ed.) (2005), Evaluation staatlicher Interventionen aus Auftraggebersicht - Schlussbericht der Projektgruppe "Evaluation", Berlin. <http://www.bmwa.bund.de/Navigation/Service/bestellservice,did=45546.html>

Bundesministerium für (BMFSFJ) (ed.) (2002) Elfter Kinder- und Jugendbericht, Berlin. <http://www.bmfsfj.de/Kategorien/Publikationen/Publikationen,did=4994.html>

Deutsche Gesellschaft für Evaluation (DeGEval) (ed.) (2002), Standards für Evaluation, Köln/Alfter. [Standardy ewaluacji] Streszczenie w jęz. angielskim: <http://www.degeval.de/calimero/tools/proxy.php?id=72> Streszczenie w jęz. rosyjskim: <http://www.degeval.de/calimero/tools/proxy.php?id=73>

Deutsche Gesellschaft für Evaluation (DeGEval) (ed.) (2004), Empfehlungen für die Aus- und Weiterbildung in der Evaluation, Alfter. [Rekomendacje dotyczące szkoleń i edukacji w dziedzinie ewaluacji – wymogi dla ewaluatorów] <http://www.degeval.de/calimero/tools/proxy.php?id=51>

Deutsche Gesellschaft für Evaluation (DeGEval) (ed.) (2004), Empfehlungen zur Selbst-Evaluation [Rekomendacje dotyczące stosowania Standardów w kontekście samooceny] <http://www.degeval.de/index.php?class=artikel&id=265>

„Guiding Principles for Evaluators” <http://www.eval.org/Guiding%20Principles.htm>

Jaedicke, Wolfgang (2005), Die Halbzeitbewertung zum Europäischen Sozialfonds in Deutschland – Ist der Sprung nach vorn gelungen?, w: ZfEV, Nr 1, str. 93-107.

Kommission zur Verbesserung der informationellen Infrastruktur zwischen Wissenschaft und Statistik (Hrsg.) (2001), Wege zu einer besseren informationellen Infrastruktur, Baden-Baden.

Pawson, Ray /Tilley, Nick (1997), Realistic Evaluation, Thousand Oaks, Ca.

Speer, Sandra/Kotischke, Henry (2005), Evaluationsstandards in der Berufsbildung, w: Berufsbildung, Vol. 91/92, str. 36-37.

Struhkamp, Gerlinde (2005), Evaluation in Germany – An Overview, w: Journal of Multi-Disciplinary Evaluation, <http://evaluation.wmich.edu/jmde/> [publikacja wkrótce].

Toepel, Kathleen/ Schwab, Oliver (2005), Evaluation in der europäischen Strukturpolitik, w: ZfEV, Nr 1, str. 63-74.

Univation (ed.) (2004), Das A-B-C der wirkungsorientierten Evaluation – Glossar Deutsch-Englisch, Köln. <http://www.univation.org/glossar/index.php>

Vedung, Evert (1998), Evaluation im öffentlichen Sektor, Vienna.

Wollmann, Hellmut (2003), Evaluation in public-sector reform: Towards a „third wave“ of evaluation? W: Wollmann, Hellmut (ed.), Evaluation in Public-Sector Reform, Cheltenham, str. 1-11.

Sandra Speer

Jest partnerem w Univation – Instytucie Ewaluacji i Szkoleń – prywatnej firmie konsultingowej, mieszczącej się w Kolonii. Jej praca skupia się na krajowych i międzynarodowych badaniach ewaluacyjnych dla urzędów publicznych takich jak Federalne Ministerstwo Zdrowia i Ubezpieczenia Społecznego, a także na ewaluacjach w zakresie szkoleń i kształcenia ustawicznego, rynku pracy i polityki społecznej. Jest zainteresowana wzmocnieniem kontaktów między różnymi europejskimi kulturami i sieciami ewaluacji. Jedną z jej specjalizacji jest metodologia ewaluacji, relacje między różnymi istniejącymi i nowymi modelami ewaluacji. Pracowała jako badacz w Seminarium Nauk Społecznych Uniwersytetu w Kolonii. Obecnie jest nieetatową wykładowcą na Uniwersytecie Nauk Stosowanych w Gummersbach and Dortmundzie.

CZĘŚĆ II

Zarządzanie oparte na wiedzy – wykorzystanie wyników ewaluacji w procesie zarządzania

Jarosław Górniak

Ewaluacja jako czynnik doskonalenia zarządzania strategicznego w administracji publicznej: projekt SAPER

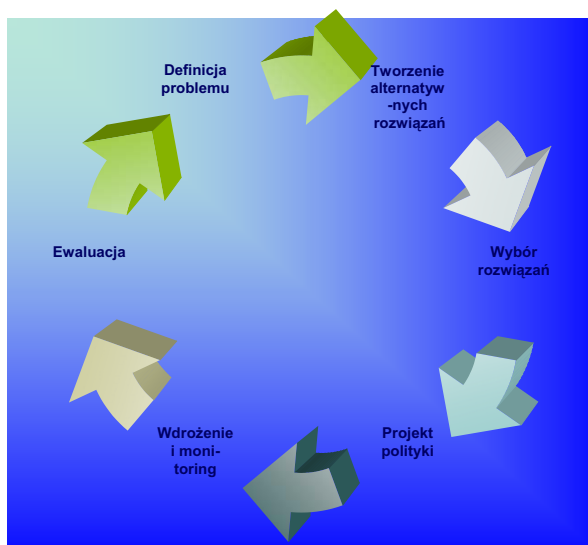
Wstęp

Agendy państwa realizują i będą realizować wiele programów i interwencji. Wynika to między innymi z naszego funkcjonowania w ramach Unii Europejskiej. Programy publiczne i interwencje powinny być poddane ścisłym reżimom programowania, monitoringu i ewaluacji. Choć w tym wypadku obowiązują wytyczne Unii Europejskiej, to często są one realizowane w sposób formalistyczny i powierzchowny, co wynika po części z nieprzygotowania administracji publicznej do wdrożenia tych zasad, a po części z interesu realizujących projekty, a nawet jednostek zarządzających funduszami, które zainteresowane są w bezkonfliktowości i płynności całego procesu. Niezbędne jest zbudowanie dobrego systemu ewaluacji programów publicznych i wdrożenie poprawnych metodologicznie zasad ewaluacji. Można w ten system włączyć świat akademicki, ale nadzór i koordynacja powinna pozostać w ręku wyspecjalizowanych służb publicznych. Skutkiem rozwoju ewaluacji powinno być także udoskonalenie systemu projektowania interwencji i regulacji.

Wyniki ewaluacji muszą być w oczywisty sposób publicznie dostępne. Podobnie, przejrzyste musi być podejmowanie decyzji o wyborze projektów do finansowania oparte o ewaluację ex-ante składanych wniosków. Wydatkowanie publicznych środków stanie się dzięki temu bardziej przejrzyste, co stopniowo powinno zwiększyć zaufanie społeczne do organów państwa oraz obliczalność tych ostatnich. W zasadzie jawność ta powinna być czymś oczywistym, w świetle ustawy o dostępie do informacji publicznej, jednak nie zawsze się tak dzieje, czego przykładem są niejasności przy kwalifikacji projektów realizowanych z funduszy europejskich, czy trudności z dostępem do dokumentacji projektowej, zdarzające się nawet w praktyce ewaluacyjnej.

Nie będę się tutaj wdawał się w rozważania nad definicjami ewaluacji, przyjmując pragmatyczną koncepcję ewaluacji jako przedsięwzięcia zmierzającego do określenia w oparciu o właściwie zgromadzone i przetworzone informacje, w jakim stopniu interwencja publiczna: polityka, program lub projekt, spełnia ustalone kryteria, w tym, w szczególności, w jakim stopniu osiągnęła cele dla osiągnięcia których została podjęta i jakie są relacje pomiędzy nakładami i działaniami a wynikami tej interwencji.

Rys. 1. Cykl polityk publicznych



W klasycznym cyklu polityk publicznych (Rys. 1) ewaluacja jest nie tyle zakończeniem łańcucha działań, lecz spoiwem pozwalającym na stałe doskonalenie polityk, programów i projektów zorientowanych na rozwiązywanie problemów społecznych. Ewaluacja jest naturalnym elementem każdego racjonalnego działania opartego na planowaniu. Ewaluacja pełni trzy podstawowe funkcje³³:

- funkcję rozliczenia i egzekwowania odpowiedzialności (accountability)
- funkcję poznawczą (knowledge)
- funkcję stymulowania usprawnień i rozwoju organizacyjnego (development).

³³ Por. Chelimsky E., *The Coming Transformations in Evaluation*, w: Chelimsky E. and Shadish W. R. (eds) *Evaluation for the 21st Century*. London: Sage Publications, 1997.

W dalszym ciągu zaprezentowane zostaną szerzej praktyczne warunki pełnienia przez ewaluację tych funkcji, wskazana zostanie potrzeba zbudowania systemu meta-ewaluacji i opisane jej funkcje oraz przedstawiony zostanie pragmatyczny projekt adresujący te wyzwania na szczeblu regionalnym. Przedstawione zostaną argumenty na rzecz tezy o kapitalnym znaczeniu ewaluacji dla procesów doskonalenia zarządzania strategicznego w administracji publicznej, a w szczególności: umiejętności precyzyjnego formułowania celów i pomiaru efektów działania, właściwego rozpoznawania związków przyczynowo-skutkowych pomiędzy interwencjami publicznymi a obserwowanymi skutkami społecznymi oraz doskonalenia innych aspektów procesu programowania i implementacji polityki publicznych, z większym społecznym pożytkiem.

Jednym z kluczowych warunków osiągnięcia takich efektów jest ściśle wkomponowanie ewaluacji w proces tworzenia i modyfikacji polityk publicznych, doprowadzenie do tego, by decydenci traktowali badania ewaluacyjne jako wartościowe i potrzebne źródło informacji o efektach działań przeszłych i prawdopodobnych konsekwencjach działań przyszłych. Wymaga to legitymizacji badań ewaluacyjnych przez ich wysoki standard metodologiczny i profesjonalny oraz obiektywizm. To jednak nie wystarczy: potrzebne jest także wyposażenie decydentów w instytucjonalną możliwość weryfikacji poziomu owych standardów, a także agregacji, analizy i syntezy wyników ewaluacji różnych projektów i programów oraz przekształcania ich w wiedzę użyteczną w procesie projektowania działań i podejmowania decyzji. Zadania te powinny wypełniać stosowne ośrodki zajmujące się na rzecz decydentów meta-ewaluacją.

Funkcja rozliczenia i egzekwowania odpowiedzialności

Jednostki organizacyjne, które zlecają ewaluację przede wszystkim ze względu na formalne wymagania związane z odbiorem projektów i programów, koncentrują swoją uwagę głównie na wymiarze jej funkcji kontrolno-rozliczeniowym. Ewaluacja bywa często traktowana jako jeszcze jedna forma kontroli. W Polsce jest o to szczególnie łatwo, gdyż kontrola – czyli sprawdzanie zgodności prowadzonych działań oraz przeznaczenia środków z założeniami i procedurami – jest instytucją dobrze znaną i wkomponowaną w biurokratyczny paradygmat organizacyjny, natomiast ewaluacja jest pojęciem nowym.

Ewaluacja nie ogranicza się, jak wiadomo, do analizy sposobu rozdysponowania zasobów, choć takie podejście jest dobrze widoczne w wielu udających ewaluację raportach z realizacji programów, które spotykamy w Polsce, na różnych szczeblach

administracji państwowej i samorządowej. Królują w nich analizy dotyczące właśnie alokacji środków. Na przykład, osiągnięcia programów zwiększania bezpieczeństwa publicznego prezentowane są poprzez informacje o wzroście nakładów na policję czy wzroście liczby patroli, a nie z perspektywy poziomu przestępczości czy poczucia bezpieczeństwa obywateli; efekty programu reintegracji mierzone są liczbą i wielkością wypłaconych zasiłków, a nie zmianą wartości mierników wykluczenia społecznego itp.

Po to, by skutecznie realizować funkcję rozliczeniową ewaluacji (i nie tylko), konieczny jest dobry pomiar wyników (performance measurement). Pomiar wyników jest integralnym składnikiem zarządzania przez wyniki (performance management), przenieszonego coraz chętniej do administracji publicznej po to, by zwiększyć efektywność i innowacyjność działań. Zwolennicy tzw. nowego zarządzania publicznego nie mają wątpliwości: po to, by mogła być egzekwowana odpowiedzialność za wykorzystanie publicznych środków zgodnie z interesem publicznym, konieczne jest wprowadzenie pomiaru efektów działań publicznych w oparciu o klarownie zdefiniowane systemy wskaźników. Jak podkreśla Picciotto: „Skuteczne zarządzanie politykami publicznymi wymaga pełnego zrozumienia związków pomiędzy nakładami, produktami, wynikami i wpływami”.³⁴ (*Effective management of public policies requires a full understanding of the linkages between inputs, outputs, outcomes and impacts*).

Każdy z wspomnianych przez Picciotto elementów musi zostać właściwie zmierzony, a zatem musi zostać ustalony odpowiedni zestaw wskaźników nakładów, produktów, rezultatów i wpływów. Zarządzanie przez wyniki (performance management) oparte jest na określaniu wyników w kategoriach skutków działania (produktów, rezultatów, wpływu), ustalaniu planowanego poziomu wyników (targets), pomiarze wyników przy pomocy zestawu wskaźników osiągnięć (performance indicators), rozliczaniu wyników w relacji do użytych zasobów i w konsekwencji dokonywaniu alokacji zasobów w oparciu o informację o wynikach³⁵. Pomiar wyników jest bazą zarówno monitoringu jak i ewaluacji, choć częściej to właśnie monitoring traktuje się jako właściwe ucieleśnienie pomiaru wyników. Monitoring prowadzony jest w trybie ciągłym i ma za zadanie sprawdzanie przebiegu realizacji projektu i dostarczanie podstaw do uruchamiania działań korygujących. Ewaluacja jest dokonywana w określonych, dyskretnych punktach programu czy projektu i nie tylko ustala stopień realizacji zaplanowanych wyników, lecz

³⁴ Picciotto, R., *Towards an Economics of Evaluation*, Evaluation Vol 5(1)/1999 s.18.

³⁵ Davis I.C., *Evaluation and Performance Management in Government*, Evaluation Vol. 5(2)/1999 s. 151.

ma dawać wiedzę o trafności programu, jego wpływie na rozwiązanie problemów społecznych czy gospodarczych, które miał rozwiązać, a zwłaszcza ustaleniu związku przyczynowego pomiędzy podejmowanymi działaniami a wynikami. Dobry monitoring znakomicie ułatwia dokonywanie ewaluacji dostarczając jej potrzebnych danych. Bez dobrego monitoringu ewaluacja staje się na ogół przedsięwzięciem kosztownym i kłopotliwym, o ograniczonych praktycznych efektach.

Jestem głęboko przekonany, że bez dobrego pomiaru wyników polityk, programów i projektów nie da się doprowadzić do tego, by publiczne środki były wydawane skutecznie i odpowiedzialnie. Dobry pomiar wyników w oparciu o porządne wskaźniki jest też bramą do dwóch pozostałych wymienionych wyżej funkcji ewaluacji: funkcji poznawczej i rozwojowej, do których jeszcze wrócimy. Dobry pomiar wyników nie jest jednak sprawą prostą.

Wśród ewaluatorów przetoczyła się debata na temat relacji między pomiarem wyników a realizacją funkcji rozliczeniowej ewaluacji. Wskazuje się zwłaszcza na to, że pomiar wyników przy pomocy wskaźników nie pozwala na dobre uchwycenie jakości programów, a to ze względu na złożoność owej jakości, reprezentującej ludzkie doświadczenia. Jakości nie da się – zdaniem krytyków pomiaru wyników – sprowadzić do zestawów wartości uzyskanych w mniej czy bardziej wyrafinowanych pomiarach, które zawsze prowadzą do uproszczeń i w konsekwencji naruszenia pluralizmu będącego immanentną cechą demokracji³⁶. Ostra krytyka ograniczania funkcji rozliczeniowej (accountability) do pomiaru wyników dokonywana jest z zwłaszcza z pozycji postmodernistycznych, antypozytywistycznych, nastawionych na traktowanie ewaluacji raczej jako metody interwencji społecznej i zmiany organizacyjnej niż zobiektywizowanego badania programów: (używa się wręcz terminu: aktywna ewaluacja³⁷).

Krytyka akcentowania pomiaru wyników jako kluczowego elementu nowoczesnego podejścia do zarządzania publicznego daje się podeprzeć szeregiem przykładów naiwnego, bezrefleksyjnego używania wskaźników i biurokratycznie prowadzonego monitoringu i pomiaru ewaluacyjnego, który nie prowadził do żadnych istotnych wniosków dla polityk publicznych. Wskaźniki są często traktowane formalistycznie, bywa

³⁶ Por. Greene, J.C., *The Inequality of Performance Measurements*, Evaluation Vol. 5(2)/1999 s. 160-172.

³⁷ Kuipers, H., Richardson, R., *Active Qualitative Evaluation. Core Elements And Procedures*, Evaluation Vol. 5(1)/ 1999 s. 61-79.

że są po prostu kopiowane z rozmaitych wzorców i przewodników. Ponieważ przedmiotem naszej troski jest Polska, warto zauważyć, że w większości polskich programów, także w kluczowych programach rządowych, wskaźniki zdają się pełnić funkcję dekoracyjną. Dodawane są zwykle jako odrębna część programu, bez jasnego związku logicznego z celami programu. Nie wiadomo za bardzo, jaka ma być ich rola: czy mają służyć celom ilustracyjnym, czy ogólnemu pomiarowi bardziej kontekstu programu niż jego wyników, czy może są po prostu – co jest dość prawdopodobną hipotezą – tylko uczynieniem zadość formalnym wymogom przewodników metodycznych Komisji Europejskiej, wymagających zaopatrzenia programów we wskaźniki. Wskaźniki są źle używane w praktyce programowania, monitoringu i ewaluacji projektów i programów publicznych w Polsce. Można powiedzieć, że bardzo często nie jest właściwie rozumiana ich istota i miejsce w zarządzaniu przedsięwzięciami publicznymi, a w konsekwencji nie jest rozumiana rola monitoringu i ewaluacji.

Trzeba zatem wrócić do podstaw metodologicznych. Wskaźnikiem jakiegoś zjawiska lub cechy (obserwowalnych lub nieobserwowalnych) nazywamy inne (względnie łatwo) obserwowalne zjawisko lub cechę, którego zaobserwowanie pozwala określić z wystarczająco wysokim prawdopodobieństwem czy precyzją o wystąpieniu badanego zjawiska, posiadaniu określonej cechy, a także – o ile to możliwe – ustaleniu jej poziomu. Wskaźniki zatem nie są jakimś samoistnym elementem programów lecz służą sprawnemu pomiarowi elementów, które go konstytuują: nakładów, wyników, skutków działań i ich wpływu na adresowane problemy społeczne.

Warto też pamiętać, że wskaźniki mogą pozostawać w różnych relacjach do cechy rzeczywistości, która jest właściwym przedmiotem zainteresowania: wskaźnik może definiować analizowane zjawisko (np. bezrobocie może zostać zdefiniowane w kategoriach wskaźnika użytego do jego pomiaru), może też być z nim skorelowane (np. indeks posiadania rozmaitych dóbr w gospodarstwie domowym jest niezłym, choć niedoskonałym wskaźnikiem bieżącego dochodu gospodarstwa domowego, a słabszym jego wskaźnikiem może być także wykształcenie). Badane zjawisko może być wreszcie konstruktem, który nie jest bezpośrednio obserwowalny (np. satysfakcja obywateli) i musi być zmierzony pośrednio przy pomocy obserwowalnych wskaźników (np. odpowiednio skonstruowanego indeksu opartego o zestaw pytań zadanych w badaniu sondażowym). Dysponujemy zatem dużymi możliwościami doboru wskaźników, których można użyć do pomiaru interesujących nas aspektów programu.

Wskaźniki są więc jedynie sposobem – bezpośrednim lub pośrednim – ustalenia stopnia realizacji poszczególnych wymiarów realizowanego przedsięwzięcia. Problem z ewaluacją – a nawet szerzej: z całym procesem projektowania i zarządzania programami publicznymi – nie wynika aż tak bardzo z ubóstwa informacyjnego zarzucanego wskaźnikom, ani z tego, że nie są one w stanie oddać dobrze jakości programów, ale z tego, że są często oderwane od istoty programu. W mniejszym stopniu dotyczy to pomiaru nakładów, gdyż tu rządzą zasady rachunkowości oraz dobrze opanowana sztuka sprawozdawczości z podejmowanych działań. W większym stopniu problem dotyczy pomiaru wyników i na tym się skupimy.

Punktem wyjścia do dobrego systemu wskaźników jest jasne zdefiniowanie celów do których się zmierza. Wystarczy przejrzeć rozmaite programy (a także projekty) publiczne, by stwierdzić, że jasne formułowanie celów nie jest ich najmocniejszą stroną. Nie będę tu przytaczał przykładów, ale nie jest o nie trudno. Trudniej jest o przykłady pozytywne. Cel jest pewnym pożądanym stanem rzeczy, który – o ile nie jest mglistym marzeniem – powinien dać się opisać w kategoriach jasno określonych cech – zmiennych. Dość natrętą manierą jest natomiast opisywanie celów programu w kategoriach działań lub nawet intencji. A więc znajdziemy w programach czy projektach sformułowania w stylu: tworzenie warunków, działanie na rzecz, sprzyjanie itp., natomiast brak jest jasnego określenia owego pożądanego stanu rzeczy, do którego osiągnięcia zmierza projekt. Jeśli zaś cel jest sformułowany w kategoriach stanu przyszłego, to często zawodzą kolejne elementy: precyzyjna definicja pojęć użytych w określeniu celu i wreszcie – tu wracamy do naszego głównego problemu – brak przełożenia celu na jego mierzalne charakterystyki, czyli na wskaźniki.

Jestem gotów zgodzić się z wieloma zastrzeżeniami formułowanymi przez część specjalistów od ewaluacji pod adresem pomiaru wyników, podnoszonych zwłaszcza przez reprezentantów szkoły konstruktywistycznej, stawiającej na ewaluację jako interwencję społeczną i praktyczną zmianę organizacyjną. Rzeczywiście, często trudno jest uchwycić dobrze wszystkie istotne aspekty ewaluowanego programu przy pomocy ilościowych wskaźników. Dlatego jestem otwarty na multiplizm – wykorzystywanie wielu metod badawczych w celu uzyskania wszechstronnego i pogłębionego oglądu ewaluowanych programów³⁸. Niemniej krytyka wskaźnikowego pomiaru wyników jako istotnego składnika ewaluacji spotykana w literaturze i wśród praktyków nie powinna

³⁸ Cook, T. D. *Postpositivist Critical Multiplism*, w: R.L. Shotland and M. M. Mark (eds) *Social Science and Social Policy*, pp. 21–62. Thousand Oaks, CA: Sage. 1985.

odwracać uwagi od istoty problemu – niskiej jakości projektowania polityk i programów publicznych. Jestem głęboko przekonany, że w Polsce konieczne jest położenie nacisku na pomiar wyników, a tak naprawdę na formułowanie celów strategii, programów i projektów w sposób jasny, możliwy do zoperacjonalizowania – czyli zbudowania zestawu wskaźników, które pozwalają ustalić czy lub w jakim stopniu cel programu został osiągnięty. To dyscyplinuje twórczych projekty i decydentów, pozwala ocenić racjonalność alokacji zasobów do programu i porównać efekty z innymi projektami.

Dobrze zdefiniowane i zoperacjonalizowane cele strategii, programów i projektów są w praktyce warunkiem przeprowadzenia sensownej ewaluacji, gdyż przecież ma za jedną ze swych podstawowych zadań ocenę tego, czy i jak zrealizowane zostały cele programu. Oczywiście, powinna sięgać dalej: ustalać związki pomiędzy działaniami a wynikami, sprawdzać, czy leżące u podstaw programu założenia teoretyczne (czy quasi-teoretyczne) okazały się adekwatne, dostarczać wiedzy po przyczynach niepowodzeń, które dzięki temu mogą okazać się nauką wartą wydanych pieniędzy. Ewaluacja powinna stymulować zmiany w politykach publicznych, zarządzaniu i alokacji środków. To, co jednak porządna ewaluacja, w tym *ex-ante*, może spowodować względnie szybko, to poprawa standardów projektowania przedsięwzięć publicznych, w tym zwłaszcza precyzji formułowania celów. Wiemy na ten temat bardzo dużo, a i tak powstają i są wdrażane poważne programy bez poważnej matrycy logicznej projektu, która jest techniką porządkującą myślenie o celach, wskaźnikach i monitoringu, a w konsekwencji – ewaluacji. Ten stan rzeczy trzeba zmienić.

Zwykle kojarzy się programy i projekty publiczne, a w konsekwencji pomiar wyników i ewaluację, z działaniami alokacyjnymi. Natomiast nie traktuje się w sposób analogiczny działań regulacyjnych – tworzenia ustaw, rozporządzeń i innych aktów prawnych – choć ich skutki są na ogół wymierne. Powinno stać się zasadą, że wprowadzanie regulacji prawnych jest traktowane jako projekt, który także powinien mieć jasno zdefiniowane cele i system pomiaru efektów.

Warto też powiedzieć kilka słów na temat samych wskaźników. Operacjonalizacja celów polityk publicznych nie jest łatwa. Najprościej jest ustalić oczekiwane produkty programu czy projektu, gdyż te są bezpośrednim, namacalnym lub przynajmniej opisywalnym wynikiem podejmowanych działań. Trudniej jest mierzyć wpływ interwencji na rozwiązanie problemów społecznych, którym jest dedykowane. „Dobrą praktyką”, godną naśladowania, jest podjęta w Komisji Europejskiej praca nad wskaźnikami wykluczenia społecznego w ramach Otwartej Metody Koordynacji (OMK) polityk społecznych. Nie ma

tu miejsca na dyskusowanie wskaźników z Laeken przyjętych w ramach tego przedsięwzięcia. Warto natomiast zwrócić uwagę, że w raporcie, który opracował zespół kierowany przez prof. Atkinsona dla potrzeb EU Social Protection Committee³⁹ skatalogowano zestaw cech, które powinny mieć dobre wskaźniki i portfele wskaźników dla określonej domeny polityk publicznych. Tutaj przytaczam trzy pierwsze, które w procesie ewaluacji wyników programów publicznych odgrywają kapitalną rolę (co nie znaczy, że inne można pominąć):

1. Wskaźnik powinien identyfikować istotę problemu oraz mieć jasną i normatywnie akceptowalną interpretację, to znaczy zmiany jego wartości powinny być jasno interpretowalne jako pozytywne lub negatywne;
2. Wskaźnik powinien być reprezentatywny i statystycznie wiarygodny, co oznacza konieczność krytycznej analizy źródeł danych, sposobu ich gromadzenia i samego sposobu konstrukcji miar;
3. Wskaźnik powinien reagować na polityczne interwencje lecz nie podlegać manipulacji.

Trwająca debata nad wskaźnikami z Laeken pokazuje, że stosowanie tych zasad nie jest łatwe, a nade wszystko, że dobór wskaźników jest funkcją definicji celu polityk publicznych, a więc jest wyborem sensu stricto politycznym, choć podlegającym metodologicznym rygorom. Dlatego dobór wskaźników, zwłaszcza tych określających ostateczne wpływy programu, nie może być wyizolowany z politycznej debaty nad podstawowymi, strategicznymi celami rozwojowymi. Mówiąc krótko i dosadnie – **to wyboczenie wskaźników z polityki jest źródłem impotencji pomiaru wyników i ewaluacji.**

Na rozmaitych szczeblach administracji państwowej i samorządowej powstają diagnozy stanu dla potrzeb strategii rozwoju. Ich autorzy zasypują decydentów danymi statystycznym, czyli wskaźnikami, które nijak się nie mają do ostatecznie przyjmowanych, *explicite* bądź *implicite*, celów działania. Dlatego często diagnozy te nie mają większego znaczenia i traktowane są jako swoista dekoracja. Trzeba podjąć trud opracowania i przyjęcia zestawu wskaźników dla poszczególnych obszarów polityk publicznych (częściowo będą przyjęte rozwiązania wypracowane w Unii Europejskiej, ale nie zaspokoją one wszystkich potrzeb). Trzeba także zwiększyć zaangażowanie

³⁹ Atkinson, T., Cantillon, B., Marlier, E. and Nolan, B. *Social Indicators: The EU and Social Inclusion*, Oxford University Press, Oxford 2002; por. też EU Social Protection Committee, *Report on Indicators in the field of poverty and social exclusion*, October 2001.

Głównego Urzędu Statystycznego w gromadzenie potrzebne dane, a zwłaszcza przetwarzanie ich we wskaźniki przyjęte na potrzeby polityki publicznych, aktualne na tyle, by miały praktyczne znaczenie dla podejmowania decyzji. Alternatywą, ale nie zawsze dostępną, jest zamówienie odpowiednich badań w renomowanych ośrodkach naukowych.

Funkcja poznawcza ewaluacji – gromadzenie wiedzy

W swoim klasycznym artykule poświęconym warunkom skutecznej implementacji polityki publicznych, opublikowanym w 1981 r., Sabatier i Mazmanian na pierwszym miejscu wymieniają jasność i precyzję celów, już na drugim zaś znajduje się na liście: trafna definicja związków przyczynowo-skutkowych pomiędzy podejmowanymi działaniami a problemami, które mają rozwiązać⁴⁰. Jak widać, owe pierwszorzędne warunki skutecznej implementacji znajdują swoje pokrycie w podstawowych funkcjach ewaluacji. Pierwszy – jasność i precyzja celu – wspierany jest przez pomiar wyników, którego właściwe egzekwowanie wymusza jasne i operacyjne określenie celów przedsięwzięć publicznych. Drugi z podanych wyżej warunków skutecznej implementacji powinien być wspierany przez właściwą realizację funkcji poznawczej ewaluacji. Jednym z kluczowych zadań ewaluacji, odróżniającym ją od bieżącego monitoringu, czy ogólniej – od ograniczenia się do oceny w oparciu o pomiar wyników – jest ustalenie związków przyczynowo-skutkowych pomiędzy podejmowanymi interwencjami, a efektami społecznymi.

Ustalenie związków przyczynowo-skutkowych w projektach publicznych jest trudne ze względu na złożoność procesów społecznych i problemy z kontrolowaniem wpływu czynników kontekstowych na uzyskiwane wyniki. Klasyczny, pozytywistyczny paradygmat ewaluacyjny, silnie promował stosowanie quasi-eksperymentu jako kluczowego planu badawczego pozwalającego zmniejszyć ryzyko nieprawidłowego ustalenia związków przyczynowych pomiędzy działaniami i efektami. Kłopoty z jego stosowaniem w realiach implementacji programów publicznych oraz łatwość popadania w formalizm badawczy (mam tu na myśli sprowadzanie wyjaśnienia do raportu statystycznego dotyczącego istotności związków pomiędzy bodźcami i reakcjami, bez uzyskiwania wglądu w mechanizm transmisyjny) doprowadziły do krytyki tego podejścia. Wskazuje się na konieczność oparcia ewaluacji na teorii dotyczącej mechanizmów wiążących działania z wynikami i na potrzebę wglądu jakościowego w procesy związane

⁴⁰ Sabatier P.A., Mazmanian D.A., *The Implementation of Public Policy: A Framework of Analysis*, w: D.A. Mazmanian, P.A. Sabatier (Eds.), *Effective Policy Implementation*, Heath, Lexington, Toronto 1981.

z realizacją programów czy projektów, wpływające zarówno na ich powodzenie, jak i na porażki.

Rozsądne stanowisko w tej kwestii nie odrzuca dorobku paradygmatu pozytywistycznego. Uznaje go jedynie za podejście niewystarczające dla właściwego spełnienia zadań ewaluacji. Wykorzystywanie wielu metod badawczych, zarówno ilościowych jak i jakościowych (czy to w ujęciu określanym mianem triangulacji metodologicznej, czy to w postaci wspomnianego już „multiplizmu” metodologicznego, w tym równoległego stosowania różnych podejść badawczych i porównywania konkluzji), oparcie ewaluacji na teorii lub konkurujących teoriach nie podważa w niczym zalet planów badawczych zmierzających do ustalenia związków przyczynowo-skutkowych w oparciu o trafny i rzetelny pomiar. Tam, gdzie quasi-eksperyment jest technicznie i etycznie możliwy i ekonomicznie uzasadniony, należy go stosować. Tam, gdzie nie jest on możliwy, należy wykorzystywać nowoczesne techniki statystyczne pozwalające na modelowanie wpływu podejmowanych działań na wskaźniki celu interwencji przy statystycznej kontroli wpływu czynników kontekstowych. W oparciu o teorię średniego zasięgu dotyczącą wprowadzanej interwencji należy budować modele będące przedmiotem weryfikacji w oparciu o rzetelny i trafny pomiar. A teorię można budować i modyfikować posługując się w procesie generowania hipotez materiałem z badań jakościowych.

Porażki i niedomogi związane w przeszłości ze stosowaniem na świecie paradygmatu pozytywistycznego w ewaluacji nie powinny stanowić wygodnego pretekstu do zarzucenia rygoryzmu metodologicznego na rzecz rozmaitych, słabo kontrolowanych procedur, często określanych mianem „jakościowych”, a faktycznie nie uwzględniających zasad metodologii badań ewaluacyjnych. W Polsce stoimy ciągle przed koniecznością wdrożenia dobrych standardów badawczych, także w ich pozytywistycznym rozumieniu. Nie będzie to łatwe, gdyż wymaga, między innymi, podniesienia jakości kształcenia metodologicznego na kierunkach kształcących potencjalnych ewaluatorów, a więc przede wszystkim na socjologii, psychologii i ekonomii. Jak dotąd, zagadnienia badań ewaluacyjnych są tam traktowane marginalnie. Zarówno w kręgach praktyki ewaluacyjnej, jak i wśród naukowców zajmujących się tymi zagadnieniami, problematyka metodologii badań ewaluacyjnych jest w moim odczuciu marginalizowana, a uwaga koncentrowana jest na samym rozumieniu ewaluacji, jej funkcjach, aspektach instytucjonalnych i regulacjach jej dotyczących (zwłaszcza zaleceniach instytucji UE). Jeśli ewaluacja ma spełniać dobrze swoje zadania, musi być oparta na wysokich

standardach metodologicznych, wyższych od typowo akceptowanych w sektorze biznesowym.

Po to, by ewaluacja mogła dobrze realizować swoją funkcję poznawczą – gromadzenia wiedzy przydatnej w doskonaleniu projektowania polityk publicznych i zarządzaniu ich wdrożeniem - konieczne jest zbudowanie zaplecza meta-ewaluacyjnego działającego na rzecz administracji publicznej. Przez meta-ewaluację rozumiem tutaj zarówno a) gromadzenie, analizowanie i syntezę wiedzy zawartej w wynikach większej liczby pokrewnych ewaluacji, jak i b) „ewaluację ewaluacji”, czyli weryfikację jakości przeprowadzonych ewaluacji z punktu widzenia przyjętych standardów profesjonalnych i wymogów metodologicznych. Oba te elementy otwierają drogę do wykorzystania ewaluacji w procesie uczenia się i usprawniania procesów programowania i wdrażania interwencji publicznych. Jak słusznie zauważają Uusikylä i Virtanen, żeby wyniki ewaluacji mogły być wykorzystane w tym celu, muszą być spełnione trzy warunki: a) wyniki ewaluacji muszą być ogólnie zrozumiałe, b) ich ustalenia muszą być aktualne, trafne i rzetelne i wreszcie c) muszą być uzasadnione merytorycznie i poprzez użytą metodologię, a także oparte na solidnych etycznych regułach ich prowadzenia⁴¹. Oczywiście, warunkiem wykorzystania wyników meta-ewaluacji opartej na obu wskazanych filarach, jest chęć wykorzystania ich przez decydentów, a także wyjście z nimi poza technokratyczne opłatki i uczynienie przedmiotem debaty publicznej. Trudno jednak zachęcać polityków i decydentów w zakresie polityk publicznych do korzystania z meta-ewaluacji, skoro jest ona w praktyce polskiej nieobecna.

Funkcja rozwojowa ewaluacji

Wskazując na potrzebę wykorzystania wyników ewaluacji, a zwłaszcza meta-ewaluacji, przez decydentów dotykamy trzeciej z wspomnianych funkcji ewaluacji: stymulowania zmiany i doskonalenia zarządzania publicznego. Właściwe rozumienie tej funkcji każe inaczej podchodzić nie tylko do ewaluacji, lecz także do samych interwencji publicznych w postaci programów czy projektów. Jeśli instytucje publiczne mają się uczyć i rozwijać, to trzeba założyć prawo do innowacji, co oznacza prowadzenie działań dotąd nie sprawdzanych w konkretnych warunkach społecznych, choć racjonalnie uzasadnionych. Musi to zakładać dopuszczalność błędów i niepowodzeń. Jednak każde niepowodzenie powinno być przedmiotem starannej analizy jego przyczyn i prowadzić do konkluzji dla

⁴¹ Uusikylä, P., Virtanen, P., Meta-Evaluation as a Tool for Learning. A Case Study of the European Structural Fund Evaluations in Finland, *Evaluation* Vol. 6(1)/2000, s. 52.

przyszłych interwencji. W ten sposób niepowodzenia nie muszą oznaczać marnotrawstwa środków publicznych lecz koszt uczenia się, które prowadzi do skuteczniejszego rozwiązywania problemów społecznych w przyszłości. Niestety, ewaluacja jest powszechnie utożsamiana z kontrolą, traktowana z niechęcią, a realizatorzy projektów i programów wykazują dużą rezerwę wobec ujawniania problemów i niepowodzeń. W gruncie rzeczy cierpi na tym nie tylko zdolność identyfikowania niepowodzeń i ustalania ich przyczyn, lecz także możliwość obiektywnego ustalenia czynników sukcesu, choć w tym wypadku, z oczywistych względów informacje są łatwiej dostępne.

Jednostki organizacyjne odpowiedzialne za meta-ewaluacje powinny być także zaangażowane w upowszechnianie standardów metodologicznych ewaluacji, ich egzekwowania oraz wsparcie doradcze agend prowadzących ewaluację, w szczególności w trudnym, początkowym okresie rozwoju ewaluacji. Sprostanie wymogom stawianym przez nowoczesną, metodycznie poprawną ewaluację – wstępną, w trakcie projektu i po jego zakończeniu – wymusza bowiem wprowadzenie właściwych metodyk projektowania opartych na jasno sformułowanych celach, miernikach poziomu ich realizacji, oddzielonych od nich działań/zadań wraz z właściwymi im wskaźnikami wykonania i kalkulacji nakładów. Dobra metodycznie ewaluacja prowadzi z czasem – za pośrednictwem meta-ewaluacji – do zgromadzenia bazy wiedzy o zależnościach przyczynowo-skutkowych i dobrych praktykach, która ułatwia programowanie i wdrażanie polityk publicznych.

Podkreślam tu funkcje poznawcze ewaluacji i ich związki z procesem racjonalizacji polityk publicznych, gdyż są one często spychane w cień przez funkcję rozliczeniową (*accountability*) ewaluację. Jest też oczywiste, że wszystkie trzy funkcje: rozliczeniowa, poznawcza i rozwojowa wspomagają się wzajemnie i dają możliwość doskonalenia polityk publicznych. Po to, by ewaluacja mogła spełniać swoje funkcje rozwojowe konieczna jest jednak zmiana instytucjonalna – zbudowanie systemu meta-ewaluacji mającego wpływ na ośrodki podejmujące decyzje w zakresie polityk publicznych, a jednocześnie niezależne od nacisku jednostek realizujących projekty czy odpowiadających za ich sukces.

Projekt SAPER – System Analizy Polityk i Ewaluacji Regionalnej

Przekonanie o konieczności doskonalenia polityk publicznych i roli, jaką odgrywa metodologicznie poprawna i traktowana poważnie przez decydentów ewaluacja doprowadziło do zainicjowania przez Małopolską Szkołę Administracji Publicznej Akademii Ekonomicznej w Krakowie i Zakład Socjologii Gospodarki i Edukacji Instytutu Socjologii Uniwersytetu Jagiellońskiego prac nad przedsięwzięciem, któremu nadana została nazwa **System Analizy Polityk i Ewaluacji Regionalnej (SAPER)**⁴². Postanowiliśmy opracować projekt zbudowania potencjału meta-ewaluacyjnego na poziomie województwa w celu rozwinięcia na poziomie regionalnym:

- zdolności do oceny samych ewaluacji, ich poprawności metodologicznej i zgodności z przyjętymi zasadami;
- zdolności do analizy i syntezy wyników ewaluacji i przekształcania ich w przesłanki dla programowania w następnych okresach;
- zdolności do wspierania metodycznego innych jednostek samorządowych w regionie.

Dlaczego postanowiliśmy zaproponować rozwiązanie na poziomie regionalnym? Z kilku powodów.

Pierwszym z nich i być może kluczowym było przekonanie, że władze samorządowe województwa małopolskiego będą zainteresowane doskonaleniem procesu programowania i (meta)ewaluacji regionalnych polityk publicznych i przygotowanie w tym zakresie rozwiązań instytucjonalnych zintegrowanych z przygotowywaną Strategią Rozwoju Województwa Małopolskiego na lata 2007-2013. Zainteresowanie decydentów jest bowiem podstawowym wymogiem zrealizowania wszystkich funkcji ewaluacji. Rzeczywiście, Zarząd Województwa wyraził już formalnie zainteresowanie projektem i podjął decyzję o uruchomieniu prac nad jego wdrożeniem. Będzie to dotyczyło w pierwszym etapie stworzenia systemu pomiaru wyników Strategii Rozwoju. Liczymy także, że zrealizowane zostaną kolejne etapy projektu, w tym zwłaszcza te, które zakładają rozwiązania instytucjonalne.

Drugim powodem było przekonanie, że wszystkie województwa stają przed analogicznym wyzwaniem. Liczba projektów publicznych realizowanych w województwach i zarządzanych na szczeblu wojewódzkim szybko rośnie. Wszystkie

⁴² W skład zespołu, który pracował nad założeniami projektu SAPER wchodzi: Jarosław Bober, Jarosław Górniak, Stanisław Mazur, Wiesław Wańkiewicz i Marcin Zawidzki.

będą musiały przejść proces ewaluacji. Będą to zarówno ewaluacje wewnętrzne jak i zewnętrzne, dokonywane przez kontraktowane w tym celu podmioty. Same te ewaluacje trzeba będzie profesjonalnie zweryfikować oraz dokonać syntezy ich wyników. Konieczne będzie zatem korzystanie z dobrodziejstw meta-ewaluacji w obu jej kluczowych aspektach. Ponadto wszędzie opracowywane są wojewódzkie strategie rozwoju i konieczne będzie zbudowanie systemu ich monitoringu i ewaluacji. Uważamy, że docelowo konieczne będzie budowanie potencjału eksperckiego i meta-ewaluacyjnego na szczeblu urzędów marszałkowskich, we współpracy z środowiskiem akademickim i eksperckim województwa. Traktujemy projekt SAPER jako rozwiązanie pilotażowe, które – mamy nadzieję – może stanowić w przyszłości „dobrą praktykę”, ułatwiająca i stymulująca powstanie podobnych rozwiązań w innych województwach.

Po trzecie, sądzimy, że w razie powodzenia tego przedsięwzięcia łatwiej będzie przekonać także rząd i parlament o konieczności stworzenia potencjału dla profesjonalnej analizy polityk publicznych i meta-ewaluacji opartej na najwyższych standardach metodologicznych.

Po czwarte, integralnym elementem projektu SAPER jest zbudowanie systemu wskaźników służących pomiarowi efektów polityk publicznych na szczeblu regionalnym. Dane statystyczne publikowane przez służby statystyczne państwa nie są jeszcze dobrze zintegrowane z koniecznością dokonywania pomiaru efektów polityk publicznych (choć zaznacza się pod tym względem postęp, także pod wpływem harmonizacji ze standardami europejskimi). Dobrą ilustracją tego problemu jest fakt, że w Polsce tylko jednorazowo, ad hoc, obliczone zostały w GUS wartości wskaźników z Laeken (niekompletne, z powodu braku wszystkich niezbędnych danych), już mało aktualne w momencie wykorzystania ich we „Wspólnym Memorandum” Komisji Europejskiej i Rządu RP w sprawie reintegracji społecznej. Później GUS nie kontynuował publikowania tych wskaźników, a są one przecież kluczowym elementem Otwartej Metody Koordynacji polityk społecznych w UE. Postanowiliśmy opracować system wskaźników tak, jak powinno być to robione, to znaczy zaczynając od eksplikacji i operacjonalizacji (w sensie metodologicznym) celów Strategii Rozwoju. Następnie zamierzamy zweryfikować dostępność danych statystycznych i opracować – w wykonalnym zakresie – standardy pomiarów uzupełniających. W przypadku wskaźników dotyczących polityk społecznych, zwłaszcza w zakresie walki z bezrobociem i wykluczeniem społecznym, uważamy za bardzo ważne dopracowanie się nie tylko klarownej bazy statystyk, ale także ich społecznej interpretacji, tzn. uzyskania wiedzy jakościowej o problemach społecznych,

które są identyfikowane przez wybrane wskaźniki. Podjęcie takiego pionierskiego przedsięwzięcia na poziomie regionalnym wydało się nam bardziej wykonalne ze względów logistycznych i ekonomicznych. Towarzyszy nam przy tym głębokie przekonanie, że wiele wniosków będzie można wprost wykorzystać w innych województwach i na poziomie całego kraju.

Na marginesie tego wątku poświęconego wskaźnikom warto wspomnieć, że w oparciu o wizję zarysowaną przez SAPER został już opracowany odrębny projekt badań nad systemem wskaźników społecznych w zakresie rynku pracy i wykluczenia społecznego w województwie małopolskim, który ma pozwolić na zbudowanie listy tych wskaźników, ich weryfikację metodologiczną i zbudowanie jakościowej „interpretacji społecznej”. Ponadto zakładamy pomiar ich aktualnego poziomu i opartą o to diagnozę sytuacji w województwie. Projekt ten przeszedł pomyślnie procedurę konkursową i uzyskał dofinansowanie ze środków Europejskiego Funduszu Społecznego⁴³; rozpoczęcie realizacji planowane jest na wrzesień 2005. Wspominam o tym projekcie także z tego względu, że będzie on prowadzony wspólnie przez jednostki eksperckie (Małopolska Szkoła Administracji Publicznej AE w Krakowie i Zakład Socjologii Gospodarki i Edukacji Instytutu Socjologii UJ) oraz jednostki publiczne (Wojewódzki Urząd Pracy w Krakowie i Wojewódzki Urząd Statystyczny). W ten sposób w praktyce chcemy zbliżyć się do idei współpracy środowisk naukowych i eksperckich z sektorem publicznym w procesie budowania potencjału ewaluacji, meta-ewaluacji i analizy polityk publicznych. Liczymy na to, że wnioski z tych badań pozwolą nam nie tylko zaproponować rozwiązania na szczeblu województwa i kraju, lecz także aktywnie włączyć się w debatę nad wskaźnikami wykluczenia społecznego, do której nowy impuls dała konferencja “Taking forward the EU Social Inclusion Process”, zorganizowana przez Prezydentkę Luksemburską UE przy wsparciu Komisji Europejskiej 13-14 czerwca 2005 w Luxemburg City⁴⁴. Zakładamy też, że wskaźniki opracowane w tym i innych projektach realizowanych z inspiracji bądź ramach SAPER-a będą także odgrywały rolę wskaźników wpływu (impact indicators), które będą wpisywane w projekty realizowane w województwie, co

⁴³ Projekt kierowany jest przez dra Stanisława Mazura (MSAP AE) – koordynacja ogólna – i prof. UJ dra hab. Jarosława Górniaka (Instytut Socjologii UJ) – koordynacja merytoryczna; w przygotowaniu projektu bardzo aktywnie uczestniczyli: dr Jolanta Perek-Białas (UJ), dr Barbara Worek (UJ) i mgr Michał Kudłacz (MSAP AE).

⁴⁴ Na konferencji został poddany pod dyskusję nowy raport poświęcony wskaźnikom wykluczenia społecznego: Atkinson, A. B., Cantillon, B., Markier, E., and Nolan, B., Taking Forward the EU Social Inclusion Process. An Independent Report Commissioned by the Luxembourg Presidency of the Council of the European Union, 2005. Raport jest dostępny na stronach internetowych: (www.fm.etat.lu) oraz (www.ceps.lu/eu2005_lu/inclusion).

zapewni ich lepsze zintegrowanie ze strategią województwa. Wskaźniki, zgodnie z założeniem przyjętym w projekcie SAPER, powinny spełniać przytoczone wcześniej zasady dotyczące wskaźników sformułowane przy okazji prac nad wskaźnikami z Laeken, gdyż są to dobrze sformułowane reguły, ukierunkowujące pracę nad wskaźnikami społecznymi.

Docelowo oczekujemy powołania w województwie małopolskim jednostki administracyjno-eksperymentalnej dedykowanej meta-ewaluacji w obu zaszyfrowanych wyżej wymiarach, prowadzącej monitoring Strategii Rozwoju województwa, przygotowującej analizy polityk publicznych, wspierającej proces projektowania strategii i programów regionalnych i upowszechniającej standardy ewaluacyjne w województwie. Pierwszym etapem realizacji SAPER-a będzie jednak opracowanie i wdrożenie systemu ewaluacji i monitoringu nowej Strategii Rozwoju Województwa.

Przyjmujemy, że przygotowanie systemu zajmie ok. 2 lat, a po jego wdrożeniu konieczne będą zapewne modyfikacje – sam SAPER wymaga ewaluacji i powinien korzystać z jej owoców. Chodzi przede wszystkim o rozpoczęcie wielkiego procesu uczenia się zdobywania wiedzy poprzez ewaluacje projektów i wykorzystywanie ich wyników w procesach decyzyjnych.

Zakończenie

Po to, by ewaluacja mogła się przyczynić do doskonalenia procesów zarządzania publicznego, w tym projektowania i implementacji interwencji publicznych w formie programów czy projektów, musi ona zostać wyprowadzona poza swoje funkcje rozliczeniowe. Ewaluacja powinna dostarczać nie tylko informacji, czy i w jakim stopniu zostały osiągnięte cele interwencji, lecz także – w jaki sposób zostały one osiągnięte, a w sytuacji porażek – dlaczego cele nie zostały osiągnięte⁴⁵. Wymaga to dobrego pomiaru wyników i nakładów osadzonego w jasno zdefiniowanych celach działania i prawidłowo określonych i zaplanowanych zasobach potrzebnych do ich osiągnięcia. Trudno się nie zgodzić z popularnym wśród teoretyków ewaluacji stanowiskiem, że wymaga to sięgnięcia po bardziej złożoną metodologię, wykraczającą poza paradygmat pozytywistyczny skupiony na pomiarze wyników i eksperymencie lub quasi-eksperymencie jako technice ustalania związków przyczynowo-skutkowych między działaniami a wynikami. Zapewne trzeba uwzględnić metody jakościowe i techniki

⁴⁵ Sanderson, I., *Evaluation in Complex Policy Systems*, *Evaluation* Vol. 6(4)/2000, s. 440.

ewaluacji partycypacyjnej. Trzeba silniej uwzględniać w ewaluacji kontekst społeczny i instytucjonalny działań oraz wartości i perspektywę beneficjentów interwencji. Ewaluacja musi być nastawiona na eksplorację i wyjaśnienie a nie tylko na pomiar wyników.

Jednak w Polsce nie jesteśmy na etapie kryzysu ewaluacji – my dobrze funkcjonującej ewaluacji jeszcze nie mamy. O ile w krajach, w których ewaluacja interwencji publicznych upowszechniła się dwadzieścia czy trzydzieści lat temu narasta frustracja na skutek ograniczonego wpływu wyników ewaluacji na realne procesy podejmowania decyzji, o tyle w przypadku Polski trudno jeszcze mówić o szerokiej możliwości wykorzystania wartościowych badań ewaluacyjnych, bo takich nie ma zbyt wiele. Ewaluacja ma zbyt często charakter formalny i służy spełnieniu formalnych wymogów.

Po to, by ewaluacja mogła być wykorzystywana w procesie strategicznego zarządzania publicznego i służyć jego doskonaleniu, musi być przede wszystkim realizowana naprawdę profesjonalnie. Wymaga to transferu wiedzy i doświadczeń, zarówno z praktyki innych krajów, jak i z praktyki innych dziedzin społecznych badań stosowanych w Polsce, jak choćby badań i analiz realizowanych dla sektora biznesowego. Wymaga też korzystania z rozwoju metodologii badań społecznych. Po to, by poziom badań ewaluacyjnych się podniósł

niezbędne jest egzekwowanie wysokiego standardu tych badań przez jednostki zarządzające, odbierające wyniki ewaluacji. Jednostki odpowiedzialne za odbiór i weryfikację ewaluacji powinny mieć kompetencje na poziomie co najmniej równym realizatorom ewaluacji. Ponadto, by umożliwić przekształcanie wyników ewaluacji w wiedzę użyteczną w projektowaniu polityk publicznych i podejmowaniu decyzji wdrożeniowych, trzeba zbudować w administracji publicznej potencjał meta-ewaluacyjny. Jest to możliwe do realizacji we współpracy pomiędzy administracją publiczną a środowiskami naukowymi. Wymaga jednak także ukierunkowania wysiłków nauki na ten obszar poprzez konkursową alokację zasobów i budowanie zinstytucjonalizowanych form współpracy uniwersytetów i akademii z urzędami.

Zaprezentowany tutaj projekt SAPER jest w moim przekonaniu dobrym krokiem w kierunku stworzenia monitoringu i ewaluacji strategii rozwoju oraz systemu meta-ewaluacji sprzyjającego wykorzystaniu ewaluacji w zarządzaniu publicznym opartym o usystematyzowaną i zweryfikowaną wiedzę.

Jarosław Górniak

(1962), dr hab., profesor Uniwersytetu Jagiellońskiego, absolwent socjologii (UJ) i ekonomii (AE w Krakowie), kierownik Zakładu Socjologii Gospodarki i Edukacji Instytutu Socjologii UJ, kierownik specjalności Socjologia gospodarki i Badania Rynku, prodziekan Wydziału Filozoficznego UJ, ponadto profesor w Wyższej Szkole Europejskiej im. Ks. J. Tischnera oraz współpracownik Małopolskiej Szkoły Administracji Publicznej Akademii Ekonomicznej w Krakowie; w latach 2002-2005 wykładał w Akademii Ekonomicznej w Krakowie analizę polityk publicznych oraz komunikację i partycypację społeczną. Redaktor naczelny "EMERGO Journal of Transforming Economies and Societies", członek redakcji "Studiów Socjologicznych" i czasopisma metodologicznego PAN: "ASK ASK. Społeczeństwo, Badania, Metody"; członek Rady Naukowej Centrum Badania Opinii Społecznej. Specjalista w zakresie metodologii badań i analizy danych oraz socjologii gospodarki i organizacji, badań rynku i komunikacji społecznej. Od wielu lat jego aktywność badawcza wiąże się z informacyjnym wspomaganie procesów decyzyjnych, najpierw w zakresie analizy zachowań konsumenckich i budowania strategii rynkowych, zwłaszcza w sektorze finansowym, a następnie także budowania rozwiązań analitycznych dla biznesu; jest głównym konsultantem firmy SPSS Polska - czołowego dostawcy rozwiązań w zakresie analizy danych. Od ok. dziesięciu lat coraz intensywniej zajmuje się analizą polityk publicznych, a od kilku lat jest to obszar wiodący w jego aktywności badawczej. Był m.in. międzynarodowym ewaluatorem Poverty Strategies Initiative - programu realizowanego przez UNDP w ponad 100 krajach, jako ekspert UNDP i ILO zajmował się zagadnieniami ubóstwa na Litwie, był koordynatorem obszaru Komunikacja i Partycypacja Społeczna i jednym z trzech koordynatorów regionalnych w Programie Rozwoju Instytucjonalnego doskonalącym administrację samorządową w Polsce (program MSWiA i Banku Światowego), oraz uczestnikiem szeregu innych przedsięwzięć programowych, analitycznych i implementacyjnych w obszarze polityk publicznych. Aktualnie kieruje zespołem prowadzącym pilotażową ewaluację śródkresową Działania 1.2 i 1.3 Sektorowego Programu Operacyjnego Rozwoju Zasobów Ludzkich.

CZĘŚĆ III

Budowanie potencjału ewaluacyjnego w Polsce

Olivier Rouland

Budowanie potencjału ewaluacyjnego w Unii Europejskiej: wskazówki dla Polski⁴⁶

Wprowadzenie

Pragnę skorzystać z okazji, aby powiedzieć kilka słów na temat stymulowanego przez fundusze strukturalne budowania potencjału ewaluacyjnego państw członkowskich oraz związanych z tym wyzwaniach i szansach, jakie stoją przed Polską.

Wraz z rozpoczęciem wdrażania funduszy strukturalnych w 2004r., opracowanie skutecznych systemów monitorowania i ewaluacji w nowych państwach członkowskich nabrało priorytetowego znaczenia.

W okresie poprzedzającym akcesję działania ewaluacyjne były w tych krajach ograniczone przeważnie do ewaluacji ex-ante programów przeznaczonych do wdrożenia w okresie wydatkowania 2004-2006 oraz ewaluacji interim (okresowej) programu Phare i powiązanych z nim projektów. W następstwie rozszerzenia, napływ funduszy ze środków Wspólnoty radykalnie się zwiększył, co zrodziło potrzebę przywiązywania szczególnej uwagi do problematyki monitoringu i ewaluacji. Inną ważną kwestią jest w tym kontekście także właściwe przygotowanie potencjału administracji publicznych nowych krajów członkowskich do kolejnego okresu programowania funduszy strukturalnych, także poprzez rozwój ich potencjału ewaluacyjnego.

W obszarze ewaluacji nowe kraje członkowskie nie zaczynają oczywiście od zera. Wszystkie spośród nich posiadają bowiem bogate tradycje w dziedzinie badań społeczno-ekonomicznych. Realizowano w nich ponadto różne inicjatywy z zakresu budowania potencjału ewaluacyjnego, finansowane z funduszy wspólnotowych, a także przekazano im niezbędną wiedzę praktyczną dzięki współpracy z ewaluatorami pracującymi na co dzień w państwach „Piętnastki”.

⁴⁶ Prezentacja ta przedstawia wyłącznie poglądy jej autora i w żadnych okolicznościach nie może być uważana za oficjalne stanowisko Komisji.

Niemniej jednak, potencjał 10 nowych państw członkowskich UE, w zakresie ewaluacji programów finansowanych ze środków publicznych, powinien być nadal pogłębiany w celu wypracowania użytecznych i wiarygodnych narzędzi zarządzania na potrzeby administracji publicznej.

Ewaluacja funduszy strukturalnych w Unii Europejskiej

- ramy administracyjne i instytucjonalne

W Unii Europejskiej ewaluacja programów finansowanych z funduszy strukturalnych jest zazwyczaj prowadzona na dwóch poziomach:

1) Na poziomie krajowym prowadzona jest koordynacja poszczególnych działań ewaluacyjnych. W tym zakresie istnieją dwa ogólne rozwiązania organizacyjne:

- niezależne zespoły ewaluacyjne w ramach instytucji krajowych (np. Krajowa Jednostka Oceny we włoskim Ministerstwie Gospodarki, Krajowa Jednostka Oceny w polskim Ministerstwie Gospodarki i Pracy) lub
- zespoły ewaluacyjne ulokowane poza strukturami ministerstw, lecz pozostające pod ich nadzorem (np. NUTEK w Szwecji, ÖROK w Austrii).

2) Ewaluacja na poziomie programu prowadzona jest przez zespoły odpowiedzialne za ocenę poszczególnych projektów. Działania ewaluacyjne na tym poziomie mogą być prowadzone:

- w ramach administracji regionalnej (np. Instytucji Zarządzającej programami regionalnymi, jak to ma miejsce w Niemczech), co jest powszechną praktyką lub
- w ramach organów krajowych zajmujących się programami horyzontalnymi (np. Instytucji Zarządzającej horyzontalnymi programami operacyjnymi w Polsce).

Zespoły odpowiedzialne za ewaluację funduszy strukturalnych w administracji publicznej liczą zazwyczaj od 2-3 do 30 osób (np. 7 na Łotwie, 30 we Włoszech, w Krajowej Jednostce Ewaluacyjnej). Są one finansowane z funduszy przeznaczonych na pomoc techniczną.

Dość powszechną praktykę stanowi organizowanie współpracy wewnętrznej między jednostkami albo zespołami zajmującymi się ewaluacją. Współpraca ta może być zorganizowana na szczeblu międzyministerialnym lub międzydepartamentalnym. Organizuje się także współpracę pomiędzy zespołami zajmującymi się różnymi programami (np. Podkomitet Ewaluacyjny na Węgrzech, Sieć Jednostek Oceny we

Włoszech). W ten sposób można zdecydowanie wzmacnić koordynację działań ewaluacyjnych i zapewnić lepszą analizę na potrzeby decydentów politycznych.

Równie dużą popularnością na płaszczyźnie ewaluacji funduszy strukturalnych cieszy się współpraca z podmiotami zewnętrznymi. Zasada partnerstwa jest szeroko stosowana, a rola towarzystw ewaluacyjnych i jednostek badawczych w koordynowaniu działań ewaluacyjnych wciąż wzrasta (np. Platforma Ewaluacyjno-Koordynująca KAP-EVA w Austrii; Grupa Konsultacyjna ds. ewaluacji Narodowego Planu Rozwoju/Podstaw Wsparcia Wspólnoty w Polsce).

Ponadto większość nowych państw członkowskich opracowała plany ewaluacyjne, które obejmują szczegółowe harmonogramy działań ewaluacyjnych w latach 2004-2006 i wskazują obszary interwencji finansowanych z funduszy strukturalnych, będące w danym kraju przedmiotem szczególnego zainteresowania.

Ocena programów finansowanych z funduszy strukturalnych

- porównanie potencjału ewaluacyjnego

Starych i Nowych Krajów Członkowskich.

„Stare” kraje członkowskie od dawna przeprowadzają ewaluację programów finansowanych przez Wspólnotę, a ich potencjał administracyjny w tej dziedzinie został rozbudowany wraz z rozwojem systemów funduszy strukturalnych. Uwzględnienie tej informacji tłumaczy obecność, poza zwykłymi utrudnieniami i wyzwaniem, uderzających różnic pomiędzy systemami i kulturą ewaluacji w starych i nowych państwach członkowskich.

W porównaniu z państwami „Piętnastki”, nowe państwa członkowskie napotykają przede wszystkim na brak doświadczonych ewaluatorów krajowych. Jednocześnie bardziej widoczna jest zależność od usług zagranicznych ekspertów w zakresie ewaluacji (pochodzących z krajów „Piętnastki”). Z drugiej strony, w krajach Piętnastki proces ewaluacji jest najprawdopodobniej przeprowadzany na zbyt wczesnym etapie programowania, podczas gdy w pozostałych 10 państwach UE nie stanowi to jeszcze problemu, gdyż dotychczas nacisk był położony na ewaluację ex-ante.

Z punktu widzenia metodologii, głównym utrudnieniem dla ewaluatorów zarówno w nowych, jak i w starych państwach członkowskich, są trudności z właściwą oceną wpływu programów. Jakość ewaluacji jest ponadto obniżana przez brak danych z monitoringu,

problemy z kwantyfikacją wyników oraz przez komplikacje w zakresie analizy związków pomiędzy nakładami finansowymi a osiągniętymi efektami. Dla nowych państw członkowskich zagadnieniami metodologicznymi stanowiącymi największe wyzwanie wydają się być: brak danych wyjściowych oraz nieprecyzyjnie zdefiniowane cele i wskaźniki wydajności, co prawdopodobnie odzwierciedla różnice w charakterze dotychczas przeprowadzonych badań ewaluacyjnych.

Brak wysokiej jakości danych z monitoringu stanowi podstawowy problem dla wszystkich ewaluatorów w całej UE. Tłumaczy to ich trudności z przeprowadzeniem badań.

Zarówno w starych, jak i w nowych państwach członkowskich, programy dotyczące zagadnień „horyzontalnych” (np. równych szans) oraz interwencje, których rezultatem ma być zwiększenie kapitału społecznego, postrzega się jako największe wyzwanie dla ewaluatorów.

Co więcej, w nowych państwach członkowskich urzędnicy zlecający ewaluację twierdzą, że jakość badań ewaluacyjnych w ich krajach pozostawia wiele do życzenia. Generalnie, analiza wpływ społeczno-ekonomicznego, analiza wyników programu oraz przedstawiane wnioski i zalecenia są postrzegane jako najsłabszy element badań ewaluacyjnych prowadzonych na potrzeby UE.

Jednymi z najbardziej uderzających problemów w całej Unii Europejskiej są: brak trwałego wkomponowania wyników ewaluacji funduszy strukturalnych w cykl tworzenia polityk publicznych, oraz fakt, iż decyzje dotyczące priorytetów budżetowych rzadko kiedy podejmowane są w oparciu o rzeczowe i spójne badania ewaluacyjne. Odnotowanie wymaga fakt, iż sceptyczne opinie na ten temat są szczególnie silne w nowych krajach członkowskich.

Z drugiej strony należy uwzględnić, że w wielu państwach Piętnastki wymogi określone w przepisach dotyczących ewaluacji funduszy strukturalnych stanowiły katalizator rozwoju kultury ewaluacyjnej, która rozprzestrzeniła się i objęła swym wpływem, nie tylko projekty związane z oceną wydatkowania środków UE, ale także ewaluację programów krajowych. Potencjał ewaluacyjny państw członkowskich był ponadto wzmacniany pod wpływem wytycznych Komisji Europejskiej oraz pomocy technicznej.

Mocne i słabe strony systemu ewaluacji funduszy strukturalnych w Polsce

Biorąc pod uwagę możliwości ewaluacyjne państw członkowskich UE warto przeprowadzić wstępną ocenę słabych i mocnych stron polskiego potencjału ewaluacyjnego w zakresie oceny funduszy strukturalnych.

Należy podkreślić, że w polskiej administracji publicznej istnieje obecnie rozwinięta sieć zespołów ewaluacyjnych z dobrze określonym podziałem obowiązków w jej obrębie. Podobnie jak w przypadku innych państw członkowskich, obowiązki te są podzielone pomiędzy Krajową Jednostkę Oceny oraz zespoły pracujące na poziomie programów.

Na poziomie tym, już na początku cyklu programowania, zostały utworzone Komitety Sterujące, które zarządzają procesami ewaluacyjnymi w sposób bardzo płynny i skuteczny.

Dodatkową silną stroną polskiego potencjału ewaluacyjnego jest długa tradycja badań socjoekonomicznych, która może zostać wykorzystana dla celów ewaluacyjnych. Dobrym przykładem są tutaj studia podyplomowe w zakresie ewaluacji prowadzone na Uniwersytecie Warszawskim. Ponadto, ewaluacja instrumentów przedakcesyjnych oraz ewaluacja ex-ante funduszy strukturalnych pomogły wzmocnić potencjał rodzimych ewaluatorów i przygotować ich do prowadzenia oceny programów finansowanych z funduszy strukturalnych w okresie 2004-2006 r.

Współpraca pomiędzy instytucjami administracji publicznej a środowiskiem ewaluatorów jest dobrze rozwinięta, co przejawia się udziałem przedstawicieli tego środowiska w pracach Grup Sterujących oraz Grupy Konsultacyjnej ds. ewaluacji Narodowego Planu Rozwoju / Podstaw Wsparcia Wspólnoty. Warto również wspomnieć, że Polska jest jedynym spośród nowych państw członkowskich, w którym istnieje ewaluacyjne stowarzyszenie branżowe (PTE), działające w bliskiej współpracy z administracją publiczną.

Należy wreszcie pamiętać, iż polski rynek ewaluacyjny, ze względu na samą wielkość kwot z funduszy strukturalnych mających zasilić gospodarkę narodową w okresie programowania 2007-2013 r., będzie bardzo atrakcyjny zarówno dla zagranicznych jak i krajowych instytutów badawczych i firm. Powinno to między innymi zaowocować

zwiększoną konkurencją pomiędzy wykonawcami oraz poprawą jakości badań ewaluacyjnych.

Z drugiej jednak strony, można zaobserwować słabe strony polskiego systemu ewaluacyjnego. Po pierwsze, w administracji publicznej wciąż brakuje wystarczającej ilości personelu niezbędnej do zarządzania rosnącą liczbą badań ewaluacyjnych. Jest to poważny problem, którym władze powinny zająć się w pierwszej kolejności.

Ponadto, mimo obecności dobrze wykwalifikowanej kadry akademickiej, zamawiający ewaluację w dużej mierze korzystają z usług ewaluatorów zagranicznych, a rynek wykonawców krajowych jest w dalszym stopniu niedostatecznie rozwinięty. W celu zmiany tej sytuacji instytucje publiczne powinny wykazać bardziej aktywną postawę (np. powinny zlecać więcej badań ewaluacyjnych i zwiększać świadomość społeczności akademickiej poprzez kampanie informacyjne).

Dodatkowo, polski system ewaluacyjny jest stymulowany głównie przez ocenę programów realizowanych w ramach funduszy strukturalnych i innych interwencji finansowanych ze środków UE. Znaczenie ewaluacji programów finansowanych ze środków krajowych jest natomiast ograniczone. Mając to uwaźe, bardzo przydatne może okazać się stworzenie ogólnego systemu ewaluacyjno-monitorującego, obejmującego zarówno programy finansowane ze środków UE, jak i z funduszy krajowych czy regionalnych. Na przykład takiego rozwiązania napotykamy w regionie Emilia Romagna (Włochy). System monitorujący obejmuje tam wszystkie interwencje wdrożone w regionie Emilia Romagna – zarówno te mające zasięg krajowy, regionalny czy lokalny, jak i interwencje finansowane w ramach EFS – bazując na spójnym zestawie wskaźników.

Wnioski

Polski system ewaluacyjny posiada dobrze zorganizowaną strukturę z jedną z najbardziej zaawansowanych kultur ewaluacyjnych wśród nowych państw członkowskich. Niemniej jednak, powinien być on nadal rozwijany, w pierwszej kolejności poprzez poprawę skuteczności współpracy z instytucjami prowadzącymi badania społeczno-ekonomiczne oraz poprzez zwiększenie liczby personelu zajmującego się ewaluacją po stronie administracji publicznej.

Biorąc pod uwagę powyższe obserwacje, mam wielką nadzieję, iż rozwój potencjału ewaluacyjnego w Polsce będzie przebiegał bez przeszkód, podobnie jak to miało miejsce w ciągu poprzednich dwóch lat, oraz że wypracowane w Polsce „najlepsze praktyki” w tej

dziedzinie będą stanowiły wartościowy wkład w proces wymiany doświadczeń pomiędzy nowymi państwami członkowskimi. Ponadto, współpraca z tymi państwami, oparta na partnerstwie w zakresie ewaluacji EFS i EFRR, może znacznie przyspieszyć ten proces, dostarczając polskim władzom nowych rozwiązań zaczerpniętych z działań ewaluacyjnych prowadzonych zagranicą oraz z doświadczeń Komisji Europejskiej.

Olivier Rouland

Ekonomista, dyrektor Jednostki Ewaluacyjnej w departamencie Komisji Europejskiej - Dyrekcji Generalnej do Spraw Zatrudnienia, Spraw Społecznych i Równych Szans. Prowadził badania w brukselskim Free University w latach 1978-1985, posiada również doświadczenie w sektorze bankowym, które zdobywał w latach 1985-1991. Od roku 1991 pracuje w Komisji Europejskiej gdzie od 1995 roku w ramach Dyrekcji Generalnej do Spraw Zatrudnienia i Dyrekcji Generalnej do Spraw Budżetu zajmuje się ewaluacją.

Leszek Korporowicz

Tworzenie potencjału ewaluacyjnego

Identyfikacja zagadnień związanych z rozumieniem istoty, struktury i charakterystyką potencjału ewaluacyjnego poprzedzona musi być refleksją nad źródłami owego potencjału, jak i jego przeznaczenia, a więc motywów, dla których warto o tym mówić, a co ważniejsze, aktywnie go stymulować. On sam, podobnie jak wszelkie formy wynikających z niego działań ewaluacyjnych nie stanowią bowiem wartości samej w sobie, niezależnie od tego w jakim stopniu są one w konkretnych przedsięwzięciach obligatoryjne, bądź też jak bardzo zalecane. Z istniejących już doświadczeń w realizacji badań ewaluacyjnych w ciągu ostatnich kilku lat, a także w realiach funkcjonowania polskich instytucji i organizacji rządowych, pozarządowych i samorządowych widać wyraźnie, iż mają one często charakter nader pozorny, niemal rytualny, tylko w niektórych przypadkach wiążą się z pełnym wykorzystaniem ich funkcji i szerszą refleksją nad ich strategiczną obecnością w dynamice ewaluowanych działań. O potencjale ewaluacyjnym trzeba więc mówić na różnych płaszczyznach, w różnych miejscach procesu podejmowania decyzji oraz w sposób, który wychodzi poza ów proces tworząc elementy kultury organizacyjnej, społecznej, a także politycznej.

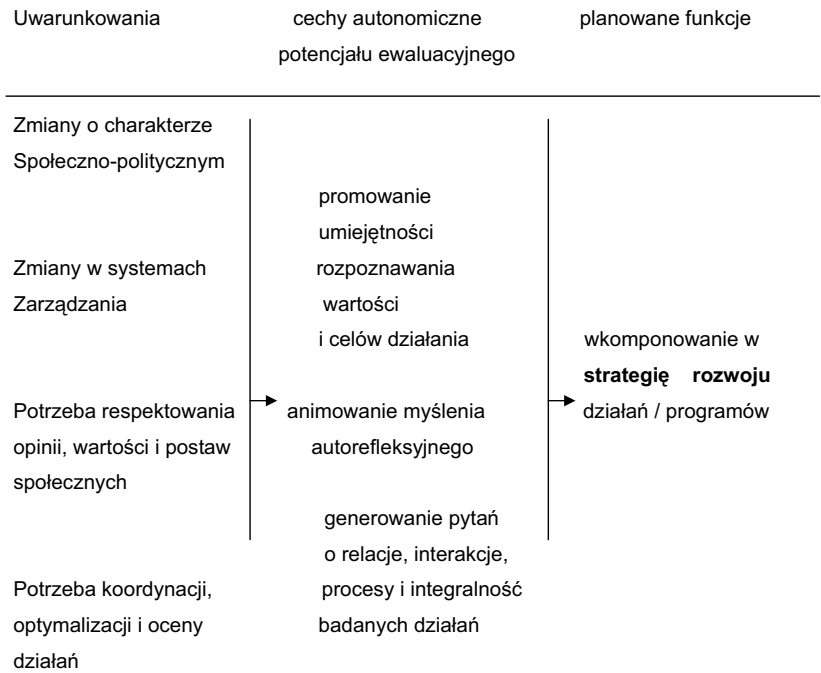
Potencjał, jak i konkretne działania ewaluacyjne są z jednej strony wynikiem przeobrażeń o charakterze **rozwojowym** w szeroko pojętym obszarze gospodarczym i społecznym, ale z drugiej strony są, a przynajmniej powinny być, tego rozwoju stymulatorem. To ów rozwój stanowi ich cel i uzasadnienie, jest miarą ich organizacyjnego i społecznego sensu, wskazuje na ich jakość i nadaje znaczenie. Czynnikiem, które potencjał ewaluacyjny wygenerowały są w pierwszej kolejności zmiany o charakterze społeczno-politycznym, które demokratyzują systemy podejmowania decyzji w oparciu o weryfikowalną publicznie wiedzę, następnie zmiany w systemach zarządzania preferujące metody diagnozy i zapewniania jakości produkcji i usług, które biorą pod uwagę postawy i oceny konsumenta, klienta i petenta. Nie bez znaczenia jest też potrzeba optymalizowania społecznych kosztów przedsięwzięć gospodarczych o różnej skali zasięgu. Na potrzebę badań ewaluacyjnych wpłynęła konieczność badania efektywności starań o koordynację i

spójność działań w różnych obszarach i na różnych poziomach struktur administracyjnych, ale także potrzeba oceny doświadczeń społecznych związanych z programami restrukturalizacji, przekształcania kompetencji zawodowych, postaw i wyborów obywatelskich, umiejętności samoorganizacji społeczności lokalnych oraz transformacji kultur organizacyjnych przedsiębiorstw, urzędów, szkół, placówek ochrony zdrowia itd.

Potencjał ewaluacyjny kryje jednak w sobie możliwości sprawcze, które nie są jedynie prostą odpowiedzią na aktualne oczekiwania. W tym zawiera się zresztą istota potencjału jako:

możliwości poznawczych i organizacyjnych, w czym kryją się zdolności koncepcyjne, komunikacyjne, badawcze i proceduralne, do rozpoznawania zakresu, stopnia i sposobu realizacji wartości społecznych zakładanych i obecnych w planowanych lub wykonywanych programach.

Świadome kreowanie zarówno kompetencji diagnostycznych i badawczych, ale w równej mierze sposobu istnienia badań ewaluacyjnych w strategiach planowania i zarządzania organizacjami, przypisany im zakres funkcji i odniesień do szerszej rozumianych wartości społecznych wynikać może z woli nadania mu konkretnych form aktywizacji. Jaka jest to forma aktywizacji, czy ma mieć funkcje jedynie technokratyczne, pomiarowe, czy też stymulować zmiany w kulturze organizacyjnej, świadomości ludzi, w sposobach pytania i weryfikowania deklarowanych wartości zależy od roli jaką chcemy nadać badaniom ewaluacyjnym. Jest to względnie autonomiczna funkcja potencjału ewaluacyjnego, która niekoniecznie wynika z przeszłych i obecnych przeobrażeń rozwojowych, ale w znacznej mierze je projektuje, wpływając na rzeczywisty kształt przyjmowanych rozwiązań organizacyjnych w realizowanych programach oraz na poziom wiedzy o ich funkcjonowaniu. Dlatego też istotne jest świadome kształtowanie polityki ewaluacyjnej, która może stanowić jeden z instrumentów wpływających na drogi rozwoju organizacyjnego wielu instytucji, ale także działań społecznych stanowiąc ich integralny komponent. Tę dwoistość potencjału ewaluacyjnego, który syntetyzuje retrospektywne i prospektywne aspekty ewaluacji obrazuje poniższy rysunek: (rys. nr 1)



Rysunek 1: Retrospektywne i prospektywne wymiary potencjału ewaluacyjnego

Niezwykle inspirujące jest z tego punktu widzenia porównanie strategii ewaluacyjnych dokonane przez Helen Simons i Jenny Hyatt jak i rezultat ich bogatych doświadczeniach w realizacji międzynarodowych projektów ewaluacyjnych. Wyodrębniają one, porównują, a jednocześnie przeciwstawiają sobie dwie orientacje, ale i cele badań ewaluacyjnych, które mogą być w warunkach polskich różniącymi się sposobami tworzenia i aktywizowania potencjału ewaluacyjnego. Pierwsza z tych orientacji projektuje funkcje ewaluacji wkomponowane w zadania kontrolne, egzekucyjne, bliskie wspomaganie potencjału i aparatu władzy. Stąd też ewaluacja pojmowana jest jako sprawozdawczości i pomiaru. Druga z orientacji nastawiona jest zdecydowanie na stymulowanie potencjałów kooperatywnej współpracy organizacji i społeczności, komunikacji społecznej, szukania kreatywności i nowych dróg definiowania i realizacji zadań. Można oczywiście przyjąć, że obie te orientacje mogą się uzupełniać i rozbudowywać potencjał ewaluacyjny o właściwe sobie funkcje. Jest jednak niezwykle istotne, aby w warunkach potrzeby odbudowywania kreatywności

społecznej i przełamywania statycznych wzorów działania organizacji, autorytaryzmu we wzorcach komunikacji i zachowań społecznych, słabo rozwiniętych praktykach interaktywnego przewodzenia, zakorzenionymi w mentalności wielu środowisk zawodowych wzorami zachowawczości, a nawet represywności działań publicznych, zaktywizować potencjał ewaluacyjny wspierający procesy rozwojowe, otwierający drogi zmiany społecznej i instytucjonalnej. Porównanie obu orientacji według przytoczonych powyżej brytyjskich Auterek, ale i praktyków badań ewaluacyjnych w różnych krajach świata przedstawia rysunek nr 2.

Ewaluacja jako sprawozdawczość	Ewaluacja jako rozwój
Nastawienie na:	
Uprawomocnienie	Zmiana
Krytyka	Procesy uczenia
Opis	Rozumienie
Dyscyplinowanie	Wspieranie
Kontrolowanie	Autorefleksja

Rysunek 2: Strategie ewaluacyjne według H.Simons i J.Hyatt ¹

Z orientacji eksponującej w ewaluacji jej cechy rozwojowe wyłania się jeszcze inna dynamika aktywizacji jej potencjału, który może mieć bardzo istotne znaczenie społeczne aktywizując z kolei inne potencjały rozwoju organizacji uczących się oraz społeczności skupionych w tych organizacjach i to w sposób, który wspiera ich autonomię, zdolności autoanalizy, umiejętności debaty wokół podzielanych systemów wartości, kierunków planowanych działań. Istotę takiej właśnie formy potencjału opisuje Walter Werner⁸ już pod koniec lat siedemdziesiątych, czyli w okresie bardzo poważnych przewartościowań teorii ewaluacyjnych, które uważnie śledziły swe

¹ Hyatt J., Simons H., *Cultural Codes – Who Holds the Keys?* Evaluation, 1999, vol 5(1)

społeczne uwikłania i krytycznie oceniały dotychczasowy dorobek. Werner odwołał się do myśli Jurgena Habermasa, który rozwijał swe koncepcje epistemologii społecznej, szczególnie w pracy *Knowledge and Human Interests* ²

Habermas wyodrębnił zasadniczo trzy rodzaje wiedzy, które spełniają odmienne funkcje, wymagają innych postaw badawczych i zróżnicowanych form ich aplikacji oraz niosą odmienne implikacje społeczne. Werner odwołując się do opisanych przez Habermasa odmienności analizuje ich konsekwencje w rozumieniu i praktycznym uprawianiu działalności ewaluacyjnej. Idąc jeszcze dalej możemy pokusić się o wynikające z tych rozważań społeczne modele ewaluacji oraz ich funkcji.

Pierwsza z wymienionych przez Habermasa form wiedzy wynika z *empiryczno-analitycznych* zainteresowań badaczy. Koncentruje się na precyzyjnym określeniu celów konkretnych działań społecznych lub gospodarczych, a następnie na arsenale realizujących je środków. Wiedza ta ma umożliwić najbardziej racjonalny sposób wykorzystania środków bez większej refleksji nad słusnością wyboru samych celów. Jest to w istocie wiedza o charakterze *technokratycznym*, dlatego cenione są umiejętności kontrolowania, przewidywania, ścisłego pomiaru, a wartościami są skuteczność, efektywność, minimalizacja kosztów.

W tak pojętym modelu wiedzy bardzo wyraźnie rysują się zadania, jakie spełniać ma funkcjonalna w stosunku do niego ewaluacja. Zasadniczym kryterium strategii ewaluacyjnych staje się stopień realizacji zamierzonych celów określany poprzez techniki badania mierzalnych osiągnięć i określania racjonalności sposobu ich realizacji. Zasadniczym celem badań ewaluacyjnych jest wytwarzanie sprawdzalnej, możliwie zobiektywizowanej wiedzy o poziomie realizacji celów ewaluowanych programów poprzez dostęp do wyspecjalizowanych danych i operacyjną umiejętność ich wykorzystania. Ewaluator jest w tej sytuacji dostarczycielem ukierunkowanej i przygotowanej według precyzyjnego wzoru informacji, która jest „narzędziem” w procesie podejmowania decyzji, optymalizowania wyborów, pomniejszania kosztów, zwiększania efektywności i skuteczności działań. Sam wzbrania się przed jakąkolwiek świadomą ingerencją w ewaluowaną rzeczywistość zakładając, że jest w stanie wystarczająco się od niej zdystansować. Na wynik jego studiów nie powinna

² Werner W., *Evaluation: sense-making of school programs*, w: Curriculum Evaluation in a New Key, Monograph series No 1, Centre for the Study of Curriculum and Instructions, University of British Columbia, Vancouver 1978.
Por. Habermas J., *Knowledge and Human Interests*, Heinemann Educational Books, London 1978.

mieć wpływu jego własna pozycja społeczna, jego własne przekonania i miejsce w analizowanej rzeczywistości, która jest zasadniczo „odrębna”, „zewnętrzna”, a on sam istnieje od niej niezależnie.

Drugi typ opisywanej przez Habermasa wiedzy określany jest przez niego jako *historyczno-hermeneutyczny* i wyrasta z potrzeby zrekonstruowania sposobu interpretacji i sposobów rozumienia rzeczywistości przez zaangażowanych w nią ludzi. Wiedza tego typu ma dać rozeznanie, jakie znaczenia nadają ludzie swych działaniom i wytworom, aby uchwycić ich świat niejako „od środka”, aby zidentyfikować rzeczywiste funkcje konkretnych zachowań. Wiele z nich nie różniąc się od siebie w swych przejawach „zewnętrznych” wyrażać może zupełnie inne intencje i wartości. Pominięcie tego prostego faktu doprowadzić może do rozminięcia się z realnością społeczną, kulturą i międzyludzka.

Zadania ewaluacji jawią się w tej perspektywie zupełnie inaczej. Ewaluacja odsłonić ma parametry „sytuacyjne”, „znaczeniowe” i motywacyjne opisując *realność doświadczeń ludzkich*. Ewaluator dokonać więc musi koniecznej interpretacji danych, w tym także zachowań indywidualnych i zbiorowych ludzi uczestniczących w ewaluowanych przedsięwzięciach, bazując nie tyle na swoim, co rekonstruowanym systemie wartości i znaczeń, jakie wyrażają ewaluowani. Ewaluacja pokazuje wówczas nie uśredniony, ale zróżnicowany system interpretacyjny uczestniczących w programie osób i grup. Celem ewaluacji nie jest wówczas ujawnienie jednej, „zobiektywizowanej” prawdy, ale raczej pokazanie społecznej percepcji analizowanej rzeczywistości, która jawi się zawsze poprzez czyjeś „prawdy”. Ewaluator nie jest zewnętrznym sędzią i nie nakłada na rzeczywistość zewnętrznie przyjętych kryteriów. Analizuje także swą własną rolę w tym, co rekonstruuje starając się mówić językiem ewaluowanych i odsłaniać ich własną perspektywę wartościowania.

Aktywizacja potencjału ewaluacyjnego jest w takim przypadku zdolnością oświetlania motywacji ludzkich, a więc źródeł swoistej energii, która może być napędem sprawczym działań objętych ewaluowanym programem.

Trzeci rodzaj wiedzy ma według Habermasa charakter *refleksyjno-krytyczny*. Jej zadaniem jest odsłonięcie ukrytych, choć stałych zależności przejawiających się w szeroko rozumianych interakcjach społecznych oraz zainicjowanie związanej z tym „odkryciem” autorefleksji. Ewaluator ujawnia w takim przypadku niewyrażane jawnie założenia programowe działań, niedostrzegane stosunki władzy, ukryte systemy wartości. Ten typ ewaluacji spełnia funkcje *emancypujące* jej uczestników poprzez

inicjowanie autorefleksji, odsłanianie mało dostrzegalnych elementów rzeczywistości społecznej, które z jednej strony wyzwalają mogą dynamizmy rozwojowe, ale także wprost przeciwnie, przyczyniać się do powstawania mechanizmów marginalizacji osób i grup, wzmacniać dysfunkcje, tworzyć bariery. Wykorzystanie tych procesów, ale i uwolnienie się od nich, gdy przynoszą zjawiska negatywne, hamujące realizację ewaluowanego programu, to także walory potencjału ewaluacyjnego.

Ewaluacja to proces **wartościowania**, badania, opisu, ale i próba rozumienia, poszukiwania i umiejętnego ukazania wartości, jakie zaistniały w zorganizowanym działaniu ludzi. Projektując ewaluację od samego początku należy pamiętać, ale i wykorzystać dwie cechy tego, co stanowi o istocie „procesu wartościowania”. Cechy te stanowią także o potencjale ewaluacyjnym, o tym, jakie dynamizmy, ale i nowe typy wartości jest w stanie wygenerować.

Pierwsza z cech, to fakt, iż ewaluacja opiera się o wymienione powyżej *umiejętności rozpoznawania wartości*, które daleko wychodzą poza techniki pomiaru i skalowania. Druga cecha, to *dynamiczny, interaktywny, rozwojowy* charakter samego procesu. Te dwie cechy stanowią mogą same w sobie bardzo istotne wartości nowoczesnie i społecznie zorientowanej ewaluacji.

Pierwsza z nich dalece odbiega od traktowania ewaluacji jedynie jako technokratycznej procedury będącej elementem procesu zarządzania, koniecznego składnika administrowania projektem. Jest działaniem zmierzającym do pobudzenia refleksji o zasadniczych celach i wartościach projektu, jest animowaniem świadomości organizacyjnej, ale i społecznej uczestników projektu. W tym znaczeniu jest ona swoistym „zaproszeniem do rozwoju” poprzez umiejętnie stawiane pytań dotyczących celów projektowanego przedsięwzięcia, sposobu w jaki realizowane mają być istotne społecznie potrzeby oraz poprzez inicjowanie umiejętności autorefleksyjnych na wszystkich szczeblach organizacji projektu.

Druga cecha skłania do uspołecznienia zarówno planowania, ale i realizacji ewaluacji poprzez zaangażowanie różnych „aktorów sceny ewaluacyjnej”, a więc organizacji podejmującej przedsięwzięcie, liderów projektu, jego beneficjentów, osoby i organizacje zainteresowane, przedstawicieli środowiska społecznego. W ten sposób refleksja i badanie wartości programu jest stymulowaniem dialogu, wzajemnego uczenia się, względnie szybkiej korekty celów, sposobu i środków realizacji zaplanowanych działań. Jest też zabezpieczeniem, przed społeczną alienacją projektu.

Wychodząc z tych właśnie przesłanek ewaluacja staje się integralnym elementem zarówno procesu planowania, jaki i wykonania programu, stanowiąc jego **wspomagający komponent systemowy**, który skłania do uspołecznienia wielu działań ze względu na wymogi demokratycznie pojmowanej ewaluacji. Wspomaga ona wówczas projekt w tej części, która czyni z niego system uczący się, interaktywny społecznie, poprzez uspołecznienie procedur ustalania celów i obszarów ewaluacji, pytań badawczych, poprzez możliwości, jakie niesie ze sobą każda jej faza, zaczynając od ewaluacji ex-ante skoncentrowanej na wartościowaniu projektu działań. Zadaniem tego typu ewaluacji jest możliwie szerokie i kompetentne zaangażowanie uczestników projektu oraz samych ewaluatorów w poszczególne fazy procesu ewaluacyjnego, jak dookreślanie kryteriów oceny, kategorii interpretacji danych, projektowanie celów i struktury, formy i języka raportów ewaluacyjnych.

Realizacja wspomnianych powyżej walorów i możliwości procesu ewaluacyjnego stanowić może o licznych animacyjnych potencjałach ewaluacji, przybierającej obecnie w opinii ekspertów postać „czwartej generacji”, która dziedziczy, ale i pokonuje ograniczenia pierwszej, skoncentrowanej na pomiarze, drugiej, rozwijającej techniki opisu wartości, trzeciej, podporządkowanej procesom decyzyjnym, dochodząc do realizacji najbardziej uspołecznionych, interaktywnych i stymulujących funkcji.³ Ich urzeczywistnienie możliwe jest poprzez odpowiednią organizację procesu ewaluacyjnego, który dokonuje się w ścisłej i ciągłej współpracy z kierownictwem programu, poprzez wypracowanie procedur i harmonogramu spotkań warsztatowych zarówno z liderami, jak i innymi uczestnikami projektu, zaplanowanych jako część responsywnej metodologii badań. Metodologia badań traktowana jest w tym przypadku nie tylko jako funkcja logicznej poprawności i wiarygodności wnioskowania, ale jako **instrument społecznej organizacji projektu**. Daje to możliwość nie tylko odpowiedniego zbierania danych, ale również artykulacji potrzeb i formowania świadomości uczestników poprzez generowanie kluczowych pytań do oceny poszczególnych działań. Efekty takich czynności możliwe są już w fazie oceny samego projektu przedsięwzięcia, a więc wspomnianej ewaluacji ex-ante, jak również w czasie jego realizacji, której odpowiada ewaluacja formatywna oraz po zakończeniu, gdy możliwe staje się podjęcie tzw. ewaluacji ex-post.

³ Guba E. G., Lincoln Y. S., *Fourth Generation Evaluation*, Sage Publ., Newbury Park 1989
Weiss C., *Evaluation*, Prentice Hall, Inc., Harvard University, New Jersey 1998
Simons H., *Specyfika i rozwój studiów ewaluacyjnych*, w: *Ewaluacja w edukacji*, red. Korporowicz L., Oficyna Naukowa, Warszawa 1997

Metodologia responsywna w ewaluacji charakteryzuje się też wrażliwością na kontekst kulturowy, organizacyjny czy też lokalny ewaluowanych działań. Staże się też elementem wspomaganie procesu zarządzania projektem, wzmacnia zainteresowanie projektem ze strony jego uczestników, dociera do zmarginalizowanych często doświadczeń i perspektyw widzenia projektu.

Istnieje wiele powodów, dla których współczesne projektowanie ewaluacji programów społecznych nie może już pominąć perspektywy jej uspołecznienia. Jest ona szczególnie przydatna, gdy kryterium oceny wychodzi poza wąsko rozumiane parametry efektywności, wewnętrznej spójności i kosztowej racjonalności.⁴

Motywy uspołecznienia wiążą się też z istotnymi pytaniami określającymi strategię ewaluacji i niosą społecznie doniosłe walory procesu ewaluacyjnego. Są to:

1. Redukcja obaw i ryzyka związanego z procedurami ewaluacyjnymi i udostępnianiem danych. Konkretnie rozwiązania, jakie służą temu celowi wiążą się z odpowiedziami na pytania:

- Kto ustala kryteria ewaluacji?
- Kto określa procedurę?
- Kto interpretuje wyniki?
- Kto jest kim w ewaluacji?

2. Unikanie marginalizacji konkretnych grup, czy punktów widzenia w procesie pozyskiwania i zarządzania danymi. Istotne są tu następujące pytania:

- Jakież są „stosunki władzy” w procesie planowania, pozyskiwania i gospodarowania informacją?
- Kto narażony jest na osłabienie i pomniejszenie znaczenia w wyniku redukcji i wybiórczego gospodarowania informacją?
- Jakież są mechanizmy konstruowania wizerunku działań i grup jako skutek odpowiedniego doboru i zagregowania informacji?

⁴ Pawson R., Tilley N., *Realistic Evaluation*, Sage Publ., Newbury Park, London, Thousand Oaks, New Delhi, 1997.

Cronbach L., *Designing Evaluations of Educational and Social Programs*, Jossey-Bass, San Francisco 1982.

Patton M. Q., *Utilization – Focused Evaluation*, Sage Publ., London 1997

Korporowicz L., *Nowe twarze ewaluacji; w kierunku interwencji, analizy, czy inspiracji*, w: Socjotechnika: kontrowersje, rozwój, perspektywy, red. Kubin J., Kwaśniewski J., IPSiR UW, Warszawa 2000

3. Zwiększenie możliwości użycia rezultatów ewaluacji poprzez różne podmioty procesu ewaluacyjnego. Istotę i znaczenie zawartych w nim możliwości uczenia ujawniają pytania:

- W jaki sposób i kto może wykorzystać rezultaty ewaluacji?
- Kto i w jaki sposób korzystać może z samego procesu ewaluacyjnego?
- Kto z kim negocjuje „obrazy/konstrukcje” konkretnych problemów w procesie i raporcie ewaluacyjnym?

4. Wzajemna edukacja społecznych podmiotów procesu ewaluacji. Walory tego wymiaru uspołecznienia eksponują następujące pytania:

- Jak przebiega proces negocjacji w fazie planowania celów ewaluacji?
- Kto komu i co komunikuje?
- Kto poznaje, analizuje i ujawnia potrzeby, interesy, zainteresowania, stanowiska oraz poglądy zaangażowanych w ewaluację podmiotów?
- Co staje się przedmiotem wspólnych ustaleń?
- Jak zmieniają się stanowiska poszczególnych uczestników?

Uspołecznione traktowanie ewaluacji doprowadzić ma do aktualizowania następujących wartości potencjału ewaluacyjnego :

- zredukowania arbitralności przyjmowanych dla ewaluacji kryteriów z racji uwzględnienia perspektywy różnych podmiotów procesu ewaluacyjnego, a co za tym idzie różnych punktów widzenia, potrzeb i wymogów,
- podniesienie prawdopodobieństwa „trafienia” studiów ewaluacyjnych w oczekiwania odbiorców,
- uczynienia z ewaluacji procedury wzmagającej zainteresowanie jej efektami,
- stymulowanie procesu uczenia się, zbliżania i uzgadniania perspektyw pomiędzy jego uczestnikami, zarówno indywidualnymi, jak też instytucjonalnymi,
- spotkania się różnych, społecznych „aktorów sceny ewaluacyjnej” w sposób wybiegający poza codzienną rutynę, który daje możliwości zaprezentowania własnej wizji problemów oraz uświadomienia istotnych różnic w postrzeganiu celów, kryteriów i sposobu zastosowania badań ewaluacyjnych,
- ujawnienia potrzeb informacyjnych, kompetencyjnych, analitycznych, technicznych i organizacyjnych uczestników ewaluacji oraz samego procesu jej realizacji.

Wymienione wyżej kwestie odpowiadają wymiarom, poprzez które rozpatrywać możemy funkcjonowanie, możliwości i procedury ewaluacji realizowanej w środowisku demokratycznym, nawet wtedy, gdy poddane jest ono sformalizowanym kanonom systemów biurokratycznych i gdy może być ona nazwana „ewaluacją demokratyczną”. W warunkach polskich taka właśnie strategia ewaluacji, która jest czymś znacznie więcej niż instrumentem zarządzania i pozyskiwania informacji może być istotnym elementem inspirującym i wspomagającym uspołecznienie systemów podejmowania decyzji, kultury organizacyjnej, strategii optymalizowania gospodarowania zasobami społecznymi, a więc może być wkomponowana w istotne procesy rozwojowe. Ewaluacja może być stymulatorem poważnych zmian społecznych, wprowadzać nowe standardy kultury pracy, komunikacji organizacyjnej.⁵

Nie bez powodu podkreśla się jednak, że otwarcie na zmiany, które wywołać może ewaluacja zależne jest od charakteru samej ewaluacji, od tego, jakiej filozofii rozwoju społecznego jest ona komponentem. Najgorszym skojarzeniem, ale i wariantem jej realizacji jest utożsamianie jej z formą kontroli. Budzi to daleko posuniętą niechęć ze strony ewaluowanych, liczne działania obronne, pozorne lub wręcz rytualne. W krajach o rozwiniętych systemach demokratyzacji, ale i profesjonalizacji życia publicznego już od lat sześćdziesiątych odchodzi się od ewaluacji jako technokratycznego instrumentu nadzoru, inspekcji ukrytej w procedurach badawczych. Odchodzi ona nawet od jednoznacznych form oceny na rzecz analitycznych form doradzania, metod rozbudzających refleksyjność uczestników procesu, troszcząc się o możliwości działania wspierającego, emancypującego.

Praktykowane w Polsce systemy zarządcze, które mają nader skromne doświadczenie z tak praktykowaną ewaluacją i które narażane były na wielce dyrektywne sposoby podejmowania decyzji, długo jeszcze dochodzić będą do wypracowania właściwej kultury ewaluacyjnej.

Podkreślanie rozwojowego znaczenia nie tylko finalnego efektu procesów ewaluacji, ale także procesu dochodzenia do jej ustaleń wynika z możliwości adekwatnego zaprojektowania celów i wkomponowania przedmiotu ewaluacji w potrzeby ewaluowanego programu. Stąd wielka rola pracy zespołu ewaluacyjnego, przyjętych metod działania, komunikacji z odbiorcami ewaluacji, procedur i umiejętności

⁵ Korporowicz L., *Ewaluacja i społeczeństwo*, w: Rozumienie zmian społecznych, Towarzystwo Naukowe KUL, red. Hałas E., Lublin 2001

konstruowania wspólnego języka, szczególnie standaryzacji rozumienia głównych pojęć.

Proces ewaluacyjny ma kilka następujących po sobie etapów, niezależnie, czy jest to ewaluacja ex-ante, ewaluacja formatywna, czy też ewaluacja ex-post. Trzy pierwsze etapy dotyczą przygotowania zasadniczych czynności badawczych, kolejne są już realizacją przyjętych celów, choć ustalenia ich dotyczące muszą być określone w fazie projektowania. Są to:

1. **Rozpoznanie zasadniczych podmiotów i celów ewaluacji.** Jest to jeden z kluczowych etapów diagnozy i dialogu „aktorów sceny ewaluacyjnej”, uzgadniania podstawowych wartości stanowiących punkt odniesienia dla dalszych czynności badawczych.
2. **Określenie planu ewaluacji.** Poprzez analizę strategicznych celów ewaluacji, zasadniczych wartości i sposobu interakcji głównych podmiotów procesu, konieczne jest dookreślenie procedur badawczych, formy i modelu ewaluacji, wskazań dotyczących przyszłego raportu lub form raportowania.
3. **Konstrukcja projektu ewaluacji.** Obejmuje on kilka doprecyzowanych ostatecznie elementów takich, jak:
 - a) cele, funkcje i obiekt ewaluacji,
 - b) obszary ewaluacji (zakresy cząstkowe),
 - c) pytania kluczowe
 - d) kryteria ewaluacji,
 - e) metodologia w poszczególnej fazie badań,
 - f) próba badawcza
 - g) harmonogram
 - h) zespół badawczy
 - i) zasoby materiałowe i finansowe
4. **Generowanie i zbieranie danych.** Pojęcie „generowania” odróżniane jest od pojęcia „zbierania” z powodu, zawartego w tym pierwszym, zadania polegającego nie tylko na biernym odtworzeniu danych, ale na skomponowaniu swego rodzaju klucza ich porządkowania, następnie selektywnego często wyboru, a co najważniejsze, wytworzenia według przyjętego **wzoru analizy**, który stawia właściwe sobie pytania, przyjmuje własne metody porównań, prowadząc pośrednio do określonego wariantu wniosków.

5. **Analiza danych.** Warto zauważyć, iż analizy ewaluacyjne typu ex-ante odnoszą się przede wszystkim do samego projektu działań i to on stanowi przedmiot wartościowania, a nie rzeczywistość społeczna, czy gospodarcza w którą ma on być wkomponowany. Studia eksperckie o charakterze diagnostycznym, które koncentrują się na badaniu samej rzeczywistości są zazwyczaj podejmowane przed budową konkretnego programu działania będącego przedmiotem ewaluacji, jako konieczny element rozpoznania rzeczowego, kontekstualnego, historycznego.
6. **Tworzenie raportu końcowego.**
7. **Komunikowanie wniosków i rekomendacji**

W tworzeniu i aktualizacji potencjału ewaluacyjnego istotne jest zaprojektowanie ewaluacji tak aby w sposób wyraźny obecne były wszystkie wymienione komponenty, niezależnie od tego jakie realizować mają wymogi konkretnego projektu, jakie specyficzne oczekiwania chcemy im przypisać.

Powstaje oczywiście pytanie: Co jest warunkiem zaistnienia takiego potencjału? W jaki sposób można go zaktualizować? Odpowiedzi na te pytania wymagają obszernych, wieloaspektowych analiz. Zanim jednak powstaną systemowe, zintegrowane programy tworzenia kadr, działań, wspierania i rozwoju ewaluacji dla potrzeb współczesnej polskiej gospodarki i zinstytucjonalizowanych form życia społecznego, konieczne jest wykorzystanie ogromnego dorobku w krystalizowaniu strategicznych celów studiów i praktyk ewaluacyjnych, jakie powstały w rozwiniętych państwach gospodarek rynkowych, ale i demokracji parlamentarnych tak, aby wkomponować je, a nawet pobudzić procesy rozwojowe społeczeństwa i państwa polskiego, jego struktur organizacyjnych, wartości społecznych, a nawet wzorów kultury.

Leszek Korporowicz

Dr hab., profesor Uniwersytetu Warszawskiego, socjolog, prorektor Szkoły Wyższej im. Bogdana Jańskiego, współzałożyciel i pierwszy prezes Polskiego Towarzystwa Ewaluacyjnego, w latach dziewięćdziesiątych uczestnik i przewodniczący zespołów ewaluacyjnych programów PHARE przy Ministerstwie Edukacji Narodowej, ekspert Centralnej Komisji Egzaminacyjnej w okresie tworzenia i wdrażania programów ewaluacji, w latach 2001-2003 przedstawiciel MEN w OECD (Education Committee), członek Państwowej Komisji Akredytacyjnej w pierwszej kadencji jej prac

2004), członek Zespołu Ewaluacyjnego Narodowego Planu Rozwoju na lata 2007-2013, wykładowca i współorganizator kilku studiów podyplomowych w dziedzinie ewaluacji, redaktor i autor prac z tego zakresu, jak "Ewaluacja w edukacji", ale także w dziedzinie socjologii kultury ("Tworzenie sensu. Język, kultura, komunikacja", "Osobowość i komunikacja w społeczeństwie transformacji"). Od kilku lat prowadzi w Instytucie Socjologii UW seminarium "Strategie i techniki ewaluacji" oraz Zespół Studiów Ewaluacyjnych.

Elżbieta Opałka

Plan działań ewaluacyjnych NPR/PWW na lata 2004-2006

Wprowadzenie

Na tle państw Piętnastki, w których ewaluacja stała się nieodzownym elementem procesu zarządzania programami o charakterze społeczno-gospodarczym, polskie doświadczenia w tym zakresie są niewielkie. Wynika to m.in. w dużej mierze z niestosowania w przeszłości na większą skalę ewaluacji jako instrumentu oceny efektów i skutków wydatków publicznych i braku uregulowań prawnych w tym zakresie.

Pomimo wielu zmian dokonujących się w ostatnim czasie w systemie prawnym, między innymi w kontekście członkostwa w Unii Europejskiej, wykorzystanie ewaluacji w codziennej pracy administracji wszystkich szczebli, jest niedostateczne. Sytuacja ta jednak ulega stopniowo zmianie w miarę stosowania przepisów Ustawy z dnia 20 kwietnia 2004 r. o Narodowym Planie Rozwoju, która uwzględniając wymogi europejskie wyróżnia i definiuje typy ewaluacji, której podlegają wszystkie dokumenty programowe współfinansowane ze środków wspólnotowych⁴⁷.

Nie zmienia to jednak faktu, że poziom wiedzy na temat celów, zakresu stosowania oraz efektów, jakie może przynieść ewaluacja jest niedostateczny i powoduje, że narzędzie to nie jest w pełni wykorzystywane w planowaniu i zarządzaniu programami pomocowymi. Sytuacja taka wymusza potrzebę koncentracji działań wspierających budowę potencjału ewaluacyjnego struktur administracji publicznej szczebla centralnego i regionalnego, które w efekcie przyczynią się do poprawy skuteczności i efektywności realizowanych programów finansowanych z funduszy strukturalnych. Z drugiej strony dynamika rozwoju rynku ewaluatorów powinna być dostosowana do potrzeb zamówień administracji do wykonywania dobrych jakościowo badań.

⁴⁷ Ustawa definiuje: ocenę szacunkową (ewaluację ex-ante) - przed rozpoczęciem realizacji programu, ocenę w połowie okresu realizacji (mid-term), oceną pełną (ex-post) - po zakończeniu realizacji oraz wprowadza termin oceny uzupełniającej (ewaluacji on-going).

System ewaluacji w Polsce

Ewaluacja funduszy strukturalnych w Polsce prowadzona jest na dwóch poziomach: Narodowego Planu Rozwoju/Podstaw Wsparcia Wspólnoty i programów operacyjnych. Za ewaluację NPR odpowiada Krajowa Jednostka Oceny, natomiast za organizację systemu i przeprowadzanie procesu ewaluacji poszczególnych programów operacyjnych odpowiadają zespoły ewaluacyjne (6 zespołów) ulokowane w strukturach instytucji zarządzających tymi programami.

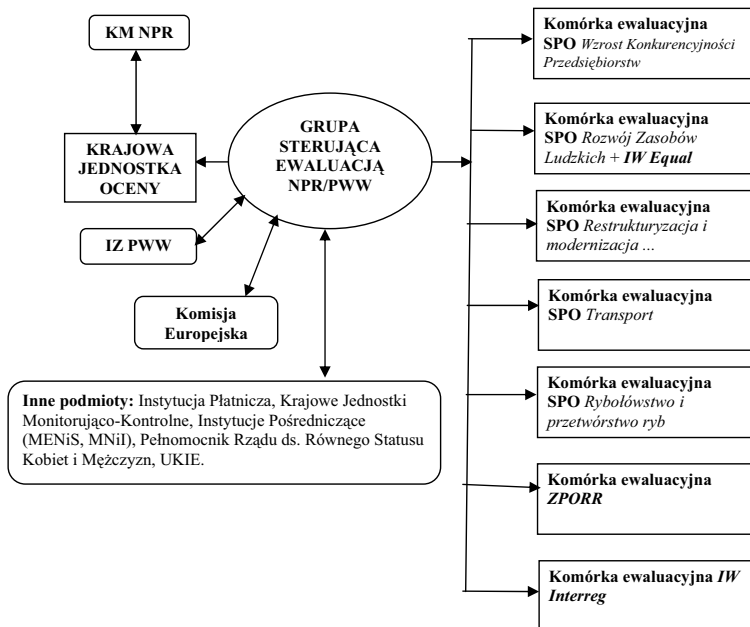
Wszystkie ewaluacje w ramach NPR i programów operacyjnych są zasadniczo zlecane niezależnym evaluatorom zewnętrznym wybieranym w procedurze zgodnej z ustawą Prawo Zamówień Publicznych. Zakłada się jednak, że wstępne analizy danych z systemu monitorowania będą wykonywane wewnętrznie przez instytucje zarządzające, w tym zespoły ewaluacyjne.

Do głównych zadań Krajowej Jednostki Oceny (KJO), powołanej zgodnie z Ustawą o Narodowym Planie Rozwoju w kwietniu 2004 r. jako Zespół ds. Oceny w ramach Departamentu Koordynacji Polityki Strukturalnej Ministerstwa Gospodarki i Pracy, należą:

- rozpowszechnianie wiedzy i informacji na temat stosowanych w Unii Europejskiej standardów ewaluacji funduszy strukturalnych,
- koordynacja procesu realizacji celów *Planu Oceny NPR 2004-2006* i współpraca w tym zakresie z zespołami ewaluacyjnymi (w ramach prac Grupy Sterującej Ewaluacją NPR/PWW),
- upowszechnianiu wyników (rezultatów) wykonanych ewaluacji NPR, w tym współpraca ze środowiskami eksperckimi w celu podnoszenia jakości badań.

Jak wynika z powyższego opisu w system ewaluacji programów finansowanych z funduszy strukturalnych w Polsce zaangażowanych jest jednocześnie kilka instytucji, gdzie każda z nich zajmuje się ściśle określonym obszarem, danym programem operacyjnym. Z tego powodu konieczna jest stała komunikacja między tymi instytucjami. Taką funkcję spełnia Grupa Sterująca Ewaluacją NPR/PWW. Dzięki działaniu Grupy możliwa jest stała i szybka wymiana informacji, wiedzy i doświadczeń pomiędzy wszystkimi instytucjami. Możliwe jest również zapewnienie koordynacji prac prowadzonych w KJO i poszczególnych zespołach ewaluacyjnych w ramach ewaluacji NPR/PWW i poszczególnych programów operacyjnych. Do zadań Grupy należy także monitorowanie i koordynacja działań informacyjno-szkoleniowych dot. ewaluacji funduszy

strukturalnych. Schemat poniżej prezentuje listę instytucji biorących udział w pracach Grupy⁴⁸.



Schemat 1: Instytucje zaangażowane w prace Grupy Sterującej Ewaluacją NPR/PWW.

W zależności od potrzeb powoływane są grupy sterujące ewaluacją w ramach programów operacyjnych albo zespoły zadaniowe do wykonania pewnych typów ewaluacji, np. ex-ante.

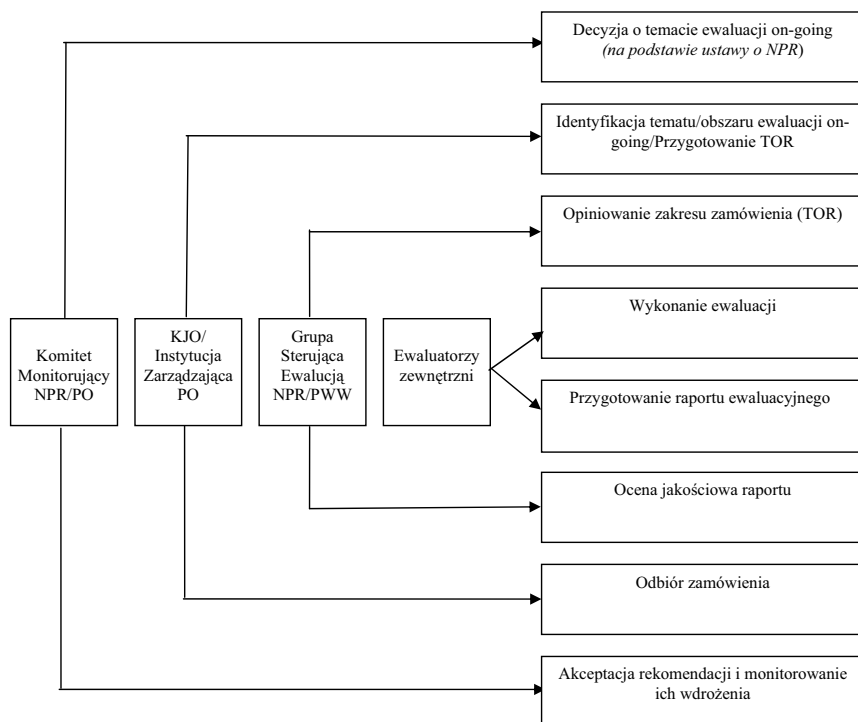
W zakresie ewaluacji NPR/PWW Krajowa Jednostka Oceny i zespoły ewaluacyjne współpracują z Komitetami monitorującymi NPR i programy operacyjne. Do głównych zadań komitetów należy:

- definiowanie tematów ocen uzupełniających (ewaluacji on-going) dotyczących NPR/PWW,
- rozpatrywanie i zatwierdzanie rekomendacji sformułowanych na podstawie zaleceń zawartych w raportach z przeprowadzonych ocen,

⁴⁸ Więcej informacji o zasadach funkcjonowania Grupy sterującej ewaluacją NPR/PWW odnaleźć można w poradniku pt. „Ewaluacja Narodowego Planu Rozwoju i programów operacyjnych w Polsce” oraz na stronie internetowej MGIP: www.funduszestrakturalne.gov.pl/Ewaluacja.

- monitorowanie postępów w realizacji zaleceń z ewaluacji (na podstawie sprawozdań przedkładanych przez KJO/instytucje zarządzające programami operacyjnymi).

Schemat poniżej prezentuje w sposób uproszczony system ewaluacji on-going programów finansowanych z funduszy strukturalnych w Polsce w latach 2004-2006.



Schemat 2. System ewaluacji on-going programów finansowanych z funduszy strukturalnych – schemat uproszczony.

Plan Oceny Narodowego Planu Rozwoju na lata 2004-2006

Zgodnie z zapisami *Podstaw Wsparcia Wspólnoty na lata 2004-2006* głównym celem działań ewaluacyjnych w obecnym okresie programowym jest wzmocnienie zdolności ewaluacyjnej polskiej administracji. Jest to niezbędny proces, który pozwoli na lepsze

wykorzystanie ewaluacji już przeprowadzonych, poprawę interwencji w okresie programowania 2004-2006 poprzez zlecenie ad hoc ewaluacji on-going/tematycznych, jak też sprawne przeprowadzenie ewaluacji ex-ante dla przyszłych interwencji funduszy strukturalnych po roku 2006 oraz przygotowanie, we współpracy z Komisją Europejską ewaluacji ex-post.

Dyskusję nad kształtem systemu ewaluacji w Polsce rozpoczęto spotkaniem zorganizowanym przez KJO (w maju 2004r.) dla instytucji zarządzających programami operacyjnymi, do udziału w którym zaproszono ekspertów mających doświadczenie w przeprowadzaniu ewaluacji programów i projektów współfinansowanych z funduszy UE. Zadania poszczególnych podmiotów zaangażowanych w proces ewaluacji NPR/PWW i programów operacyjnych były przedmiotem zorganizowanych w lipcu 2004 r. dwudniowych warsztatów konsultacyjno-szkoleniowych. Dodatkowo zlecono serię ekspertyz, służących zarówno sprawom organizacyjnym (porównanie systemów organizacji ewaluacji w Irlandii i we Włoszech) oraz kwestiom *stricte* metodologicznym prowadzenia badań ewaluacyjnych w Polsce.

Wyniki ustaleń podjętych podczas warsztatów, jak też uzgodnień z przedstawicielami instytucji zarządzających programami operacyjnymi wykorzystano przy opracowywaniu *Planu oceny Narodowego Planu Rozwoju na lata 2004-2006*, będącego narzędziem do planowania i realizacji działań ewaluacyjnych w obecnym okresie programowym, czyli do końca 2008 roku. W dokumencie tym zdefiniowano cele ewaluacji Narodowego Planu Rozwoju, w tym Podstaw Wsparcia 2004-2006 oraz wskazano zadania do realizacji w ramach budowy potencjału ewaluacyjnego i przeprowadzenia ewaluacji w ramach NPR, jak też określono wstępną listę głównych tematów ocen uzupełniających (ewaluacji on-going) przewidzianych do wykonania w ramach NPR/PWW i poszczególnych programów operacyjnych w latach 2005-2006.

Głównym źródłem finansowania przedsięwzięć z zakresu ewaluacji NPR/PWW jest Program Operacyjny Pomocy Technicznej na lata 2004-2006.

Zgodnie z Planem Oceny NPR 2004-2006 nadrzędnym celem przeprowadzania oceny w ramach Narodowego Planu Rozwoju jest stałe ulepszanie skuteczności i efektywności interwencji publicznej, rozumiane nie tylko w kategoriach pozytywnych skutków społecznych lub gospodarczych związanych bezpośrednio z Planem, lecz także zwiększania przejrzystości i promowania działań podejmowanych przez władze publiczne.

Realizacja tego celu odbywa się poprzez osiągnięcie celów głównych obejmujących:

1. budowę potencjału ewaluacyjnego polskiej administracji (w tym rozbudowę struktur, skoordynowanie działań z zakresu ewaluacji oraz rozpowszechniania wiedzy i informacji na temat sposobu przeprowadzania i wyników badań ewaluacyjnych);
2. organizowanie ocen (ewaluacji) w ramach NPR 2004-2006 i 2007-2013.

Budowa potencjału ewaluacyjnego administracji

W zakresie budowy potencjału ewaluacyjnego wzmacniana jest systematycznie zdolność administracyjna Krajowej Jednostki Oceny do prowadzenia efektywnej koordynacji nad funkcjonowaniem systemu ewaluacji w Polsce oraz przeprowadzania ewaluacji NPR/PWW. Do końca roku 2005 planowano zwiększyć stan kadrowy Krajowej Jednostki Oceny do 5 osób. Obecnie w KJO pracują jednak tylko 3 osoby. Dodatkowo zatrudniono pracownika na umowę-zlecenie. Rozbudowywane są także zespoły ewaluacyjne ulokowane w instytucjach zarządzających programami operacyjnymi (obecnie w tych jednostkach zatrudnionych jest łącznie 11 osób).

W ramach prac związanych z upowszechnianiem informacji o ewaluacji, przygotowano i wydano drukiem poradnik pt. *„Ewaluacja Narodowego Planu Rozwoju i programów operacyjnych w Polsce”*, stanowiący wykładnię obowiązujących procedur i zasad przy przeprowadzaniu ewaluacji, uaktualniono zawartość i rozbudowano zakładkę „Ewaluacja” na stronie MGIP poświęconej funduszom strukturalnym⁴⁹. Ponadto stworzona na serwerze MGIP biblioteka opracowań z zakresu ewaluacji - raporty ewaluacyjne, dokumenty robocze i wytyczne z krajów UE, opracowania naukowe, *best practice* - powinna przyczynić się do wymiany informacji o pracach z zakresu ewaluacji prowadzonych w różnych krajach.

Powołano w grudniu 2004 r. Grupę Sterującą Ewaluacją NPR/PWW, która obok wymiany wiedzy i informacji wynikających z prac prowadzonych we wszystkich jednostkach ewaluacyjnych ulokowanych w instytucjach zarządzających funduszami Unii Europejskiej w Polsce, zapewnia również promowanie stosowania jednolitych standardów w zakresie ewaluacji programów współfinansowanych z funduszy strukturalnych UE. Krajowa Jednostka Oceny przewodniczy pracom Grupy. Wszyscy członkowie Grupy zostali przeszkoleni w kwestiach podstawowych związanych z ewaluacją.

⁴⁹ www.fundusze-strukturalne.gov.pl

W zależności od potrzeb powoływane są grupy sterujące ewaluacją w ramach poszczególnych programów operacyjnych, do których udziału zapraszana jest Krajowa Jednostka Oceny. Funkcjonują już grupy w ramach programów operacyjnych Rozwoju Zasobów Ludzkich i Wrostku Konkurencyjności Przedsiębiorstw oraz Inicjatywy Wspólnotowej Equal. Regularnie odbywają się spotkania tych grup.

Ewaluacje Narodowego Planu Rozwoju i programów operacyjnych w 2005 roku

Ewaluacje on-going

W I połowie 2005 roku prace koncentrowały się głównie na przygotowaniu i uruchomieniu zamówień publicznych na realizację ocen uzupełniających (ewaluacji on-going) przewidzianych do wykonania w *Planie Oceny NPR* do końca 2005 roku. Obecnie realizowane badania dotyczą procesu wykorzystania środków funduszy strukturalnych i problemów, które pojawiają się w trakcie wdrażania. Są to:

- ✓ Ocena przewidywanego wpływu wybieranych projektów służących modernizacji i rozbudowie infrastruktury społecznej na rozwój społeczno – gospodarczy kraju, w tym na osiąganie celów zapisanych w dokumentach programowych;
- ✓ Funkcjonowanie instytucji we wdrażaniu funduszy UE na poziomie regionalnym pod kątem określenia ich potencjału dla wdrożenia zdecentralizowanego systemu zarządzania regionalnymi programami operacyjnymi w okresie programowym 2007-2013;
- ✓ Funkcjonowanie wdrażania NPR/PWW 2004-2006 i poszczególnych programów operacyjnych⁵⁰.

Ponadto, na potrzeby konferencji z beneficjentami wsparcia z funduszy strukturalnych UE, organizowanej przez MGIP na początku października 2005 r. zlecono wewnętrznie przygotowanie 4 ocen uzupełniających wybranych obszarów, które obejmują:

- charakterystykę projektów szkoleniowych realizowanych w ramach Europejskiego Funduszu Społecznego (dot. SPO RZL);
- typy projektów realizowanych przez organizacje pozarządowe w ramach Europejskiego Funduszu Społecznego (dot. SPO RZL). Przyczyny różnicowania zainteresowania organizacji pozarządowych poszczególnymi działaniami SPO RZL;

⁵⁰ Większość instytucji zarządzających rozpoczęła już prace nad tym badaniem, a niektóre są już w trakcie realizacji.

- analizę rodzajów inwestycji realizowanych w ramach SPO „*Rybołówstwo i przetwórstwo ryb 2004-2006*” oraz przyczyny zróżnicowanego zainteresowania beneficjentów poszczególnymi typami projektów;
- rodzaje projektów realizowanych w ramach działania 1.3 *Tworzenie korzystnych warunków dla rozwoju firm SPO „Wzrost konkurencyjności przedsiębiorstw, lata 2004-2006”* oraz ocena ich wpływu na konkurencyjność gospodarki.

Wyniki uzyskane z tych badań zostaną wydane w formie raportu i zaprezentowane podczas konferencji.

Zakresy wszystkich przygotowanych ocen zostały skonsultowane na posiedzeniach Grupy sterującej ewaluacją NPR/PWW.

Zgodnie z zapisami *Planu Oceny NPR* do głównych zadań planowanych do realizacji w 2005 roku należą:

- Wykonanie ocen uzupełniających w ramach NPR/PWW i programów operacyjnych, dotyczących:
 - ✓ skuteczności i efektywności (wydajności) podejmowanych działań i ich oddziaływania na sytuację społeczno-gospodarczą, w tym na rynek pracy;
 - ✓ oceny przygotowania i realizacji dużych projektów infrastrukturalnych;
 - ✓ analizy postępów w realizacji NPR/PWW, spójności wybieranych projektów z celami NPR i programów operacyjnych oraz oceny wpływu funduszy UE na sytuację społeczno-gospodarczą;
 - ✓ analizy efektywności projektów z zakresu infrastruktury drogowej;
 - ✓ analizy postępów, efektywności i oddziaływania realizacji programów operacyjnych na sytuację społeczno-gospodarczą;
 - ✓ identyfikacji barier we wdrażaniu innowacyjnych i kompleksowych projektów;
 - ✓ wzrostu poziomu innowacyjności polskich przedsiębiorstw i rozwój współpracy ze sferą badawczo-rozwojową;
 - ✓ analiza poziomu i dostępności usług instytucji otoczenia biznesu;
 - ✓ zasada równości szans w instytucjach zaangażowanych we wdrażanie funduszy strukturalnych na różnych poziomach administracji.

Wnioski i rekomendacje z tych ocen postępują instytucjom zarządzającym w usprawnieniu koordynacji i spójności działań podejmowanych na różnych szczeblach zarządzania NPR/PWW i programów operacyjnych. Wyniki zostaną również wykorzystane na potrzeby raportowania do Komisji Europejskiej o osiągniętych rezultatach i szacowanym

wpływie realizacji funduszy strukturalnych w Polsce (*up-date of mid term evaluation*) oraz oceny ex-ante NPR 2007-13 i PO, jak też ex-post NPR 2004-2006 i PO.

- Wykonanie oceny szacunkowej (ewaluacji ex-ante) projektu *Narodowych Strategicznych Ram Odniesienia* (NSRO) oraz uruchomienie prac nad przygotowaniem ewaluacji ex-ante projektów programów operacyjnych, opracowywanych w ramach NPR na lata 2007-2013 (w ramach prac Grupy sterującej ewaluacją dokumentów unijnych na lata 2007-13);
- Zlecenie analiz i ekspertyz służących rozbudowie systemu wskaźników monitorowania NPR/PWW na potrzeby przyszłego okresu programowego 2007-2013;
- Zapewnienie odpowiedniej koordynacji prac w obszarze ewaluacji poprzez regularne spotkania Grupy Sterującej Ewaluacją NPR/PWW (w 2005 zaplanowano 6 spotkań) i konsultacje na forum Grupy konsultacyjnej ds. systemu oceny funduszy Unii Europejskiej w Polsce (zaplanowano 2 spotkania w roku). Dodatkowo zorganizowanie konferencji międzynarodowej poświęconej dyskusji nad metodami i technikami ewaluacji różnych typów interwencji społeczno-gospodarczych posłużą upowszechnieniu wyników przeprowadzonych badań ewaluacyjnych;
- W celu upowszechniania wiedzy i informacji nt. ewaluacji planowane jest przygotowanie i wydanie publikacji z zakresu ewaluacji z udziałem ekspertów zagranicznych oraz uruchomienie długofalowego programu szkoleń specjalistycznych dla pracowników komórek zaangażowanych w zlecenie ewaluacji programów o charakterze społeczno - gospodarczym finansowanych z funduszy strukturalnych UE.

Ewaluacja ex-ante NPR 2007-2013

W czerwcu 2005 r. zakończono prace nad wykonaniem oceny szacunkowej (ewaluacji ex-ante) *Wstępnego Projektu Narodowego Planu Rozwoju na lata 2007-2013*. Uwagi, sugestie i zalecenia ekspertów – członków specjalnie powołanego Podzespołu ds. Ewaluacji NPR 2007-2013 – zostały uwzględnione podczas opracowywania Projektu Narodowego Planu Rozwoju. Raport końcowy z tej ewaluacji wydrukowano i udostępniono na stronie internetowej Ministerstwa Gospodarki i Pracy.

Wraz z zakończeniem tego etapu ewaluacji ex-ante, rozpoczęto prace zmierzające do przeprowadzenia ocen szacunkowych na poziomie programów operacyjnych (sektorowych i regionalnych). Rozpoczęto również przygotowania do wykonania ocen

oddziaływania na środowisko dokumentów programowych. Opracowana na potrzeby wstępnego projektu NPR 2007-2013 *prognoza oddziaływania na środowisko* stanowi pierwszy etap prac nad dokumentami programowymi na lata 2007-2013. Wypełniając zapisy ustawy Prawo Ochrony Środowiska oceny oddziaływania zostaną przygotowane dla dokumentów wymaganych procesem negocjacji z Komisją Europejską, tj. Narodowych Strategicznych Ram Odniesienia, Narodowej Strategii Rozwoju Obszarów Wiejskich i programów operacyjnych wynikających z tych dokumentów.

Kolejną oceną wykonaną w ramach ewaluacji ex-ante NPR, była ocena oddziaływania makroekonomicznego tego dokumentu. Wykonano ją modelem Hermin. Równocześnie zainicjowano prace nad regionalizacją modelu Hermin i przygotowaniem 16 baz danych.

Problemy i zagrożenia w procesie zarządzania ewaluacją

Do głównych, zidentyfikowanych już na obecnym etapie prac, problemów i zagrożeń procesu zarządzania ewaluacją NPR/PWW należą:

- niedostateczny potencjał kadrowy jednostek administracyjnych odpowiedzialnych za realizację ewaluacji funduszy strukturalnych na poziomie NPR/PWW i programów operacyjnych (niski stan kadr, brak wykwalifikowanego personelu),
- skomplikowane i czasochłonne procedury zamówień publicznych,
- niska jakość i niedostępność danych monitoringowych o realizowanych projektach,
- ograniczony rynek ewaluatorów zewnętrznych (niezadowolająca jakość pracy ewaluatorów).

W celu przeciwdziałania tym zagrożeniom zaplanowano stosowne działania w *Planie oceny NPR 2004-2006*. Do głównych działań należą:

- udział pracowników administracji, w tym Krajowej Jednostki Oceny i zespołów ewaluacyjnych ulokowanych w strukturach instytucji zarządzających programami operacyjnymi w seminariach i spotkaniach służących poznaniu doświadczeń krajów UE w zakresie organizacji i sposobów przeprowadzania badań ewaluacyjnych zapewni podnoszenie wiedzy o ewaluacji, a więc wpłynie na jakość wykonywanych przez administrację zadań, w tym na sprawne zlecenie badań ewaluacyjnych w ramach Narodowego Planu Rozwoju i programów operacyjnych;
- w celu zapewnienia odpowiednich danych do przeprowadzania badań ewaluacyjnych prowadzona będzie przez KJO stała aktualizacja bazy wskaźników monitorowania (poprzez udział w pracach Grupy roboczej ds. wskaźników), jak też stosowanie przez

instytucje zarządzające mechanizmów weryfikacji danych na miejscu (wizyty kontrolne).

- prace Grupy konsultacyjnej ds. systemu ewaluacji funduszy strukturalnych w Polsce oraz organizowanie seminariów/konferencji zapewnią wsparcie umiejętności (zdolności) do przeprowadzania ewaluacji wśród administracji publicznej i potencjalnych ewaluatorów zewnętrznych

Podsumowanie

Budowie potencjału do prowadzenia ewaluacji funduszy strukturalnych sprzyjają wszelkie działania służące wymianie wiedzy, informacji i doświadczeń z Komisją Europejską, bliźniaczymi instytucjami administracji publicznej z innych krajów Unii Europejskiej (np. na forum *Evaluation Network*⁵¹) oraz instytucjami krajowymi zaangażowanymi w programowanie i zarządzanie funduszami strukturalnymi (w ramach prac grup sterujących ewaluacją). Wymiana informacji i doświadczeń sprzyjać będzie nie tylko podnoszeniu umiejętności do zlecania ewaluacji, ale również przyczyni się do budowania zaufania wobec wyników badań ewaluacyjnych i wykorzystania ich w praktyce administracyjnej.

Konsekwentna realizacja działań zidentyfikowanych w *Planie oceny Narodowego Planu Rozwoju na lata 2004-2006* w zakresie budowania potencjału ewaluacyjnego administracji, obejmujących wzrost liczebności kadr, stworzenie instrumentów koordynacji, tj. powołanie Grupy sterującej ewaluacją NPR/PWW, Grupy konsultacyjnej ds. systemu oceny funduszy UE w Polsce (złożonej z ekspertów zewnętrznych) oraz przygotowanie poradnika stanowiącego wykładnię obowiązujących procedur i zasad przy przeprowadzaniu ewaluacji, przyczyni się do ukształtowanie profesjonalnego i skutecznego modelu wykonywania zadań przez polską administrację. Będzie to miało także niewątpliwy wpływ na sposób przygotowania i realizacji programów korzystających wyłącznie z krajowych źródeł finansowania zwłaszcza, iż zagadnienia związane z ewaluacją stają się coraz bardziej powszechne w tych programach.

⁵¹ Jest to Grupa techniczna ds. Ewaluacji zorganizowana przez DG ds. Polityki Regionalnej Komisji Europejskiej, w której pracach uczestniczą przedstawiciele wszystkich krajów członkowskich UE.

Ukończyła studia na Wydziale Geografii i Studiów Regionalnych na Uniwersytecie Warszawskim. Od lutego 2004 r. doktorantka w studium Modernizacja struktur społeczno-gospodarczych w Polsce w warunkach członkostwa w Unii Europejskiej w Szkole Głównej Handlowej w Warszawie. Od 2001 r. zatrudniona w MGiP, a od kwietnia 2004 r. kieruje Zespołem ds. Oceny (zwanym Krajową Jednostką Oceny) w Departamencie Koordynacji Polityki Strukturalnej MGiP, gdzie m.in. odpowiada za: funkcjonowanie systemu ewaluacji skuteczności i efektywności realizacji Narodowego Planu Rozwoju/Podstaw Wsparcia Wspólnoty, wypracowanie standardów prowadzenia badań ewaluacyjnych w Polsce oraz koordynuje pracami nad wykonywaniem ewaluacji programów interwencji społeczno-gospodarczych finansowanych z funduszy strukturalnych UE realizowanych w ramach NPR/PWW. Przewodniczy pracom Grupy sterującej ewaluacją NPR/PWW oraz reprezentuje Polskę w pracach Grupy Technicznej ds. Ewaluacji w DG ds. Polityki Regionalnej Komisji Europejskiej (Evaluation Network). W okresie od stycznia do czerwca 2005 r. prowadziła sekretariat Zespołu Ewaluacyjnego, w ramach którego wykonała ewaluację ex-ante projektu NPR 2007-13. Współpracuje z nowymi krajami członkowskimi UE w zakresie wymiany informacji i doświadczeń ewaluacyjnych na forum Grupy Wszehradzkiej oraz ze środowiskiem polskich ewaluatorów w ramach Grupy konsultacyjnej ds. systemu oceny funduszy UE w Polsce. Jest współautorką wydanego w 2005 roku poradnika „Ewaluacja Narodowego Planu Rozwoju i programów operacyjnych w Polsce”.

