Tomasz Kupiec

Program evaluation use and its mechanisms : The case of Cohesion Policy in Polish regional administration

2015

MUZEUM HISTORII POLSKI

Tomasz Kupiec

Program evaluation use and its mechanisms: The case of Cohesion Policy in Polish regional administration

This article discusses evaluation use in the area of EU operational programs implemented by Polish regional administration, which is an uncharted territory. The analysis is based on the assumption that evaluation is a long-term process producing a stream of knowledge that supports management decisions throughout a program's lifetime. Three cases of regional programs, their managing authorities and 44 evaluation studies completed by them between 2007 and 2012 were analyzed. The degree of evaluation use was found unsatisfactory and limited to minor modifications of the implementation process. The main barrier to the evaluation use was poor quality of evaluation studies, obvious and insignificant conclusions, reports missing answers to key questions. That resulted from other problems: incompetence of evaluators and inappropriate research methodologies.

Keywords: evaluation use, program evaluation, regional operational program, Poland, Cohesion Policy

Introduction

Use is the key concept in the field of evaluation. Most evaluators expect their evaluations to be used, and accept the fact that the value of their work can be judged from this perspective (Henry and Mark, 2003). The question of whether evaluations are used and how was asked almost as soon as evaluation practice emerged (Ledermann, 2011); in the late 1980s it became a "hot topic" (Alkin and Coyle, 1988).

Considerable number of research in the field has led to the development of taxonomies of use (e.g. Knorr, 1977; Weiss, 1980; Patton, 1997; Shulha and Cousins, 1997). Eventually, at the beginning of the 20th century an alternative term and theory of evaluation influence was proposed (Henry and Mark, 2003; Mark and Henry, 2004; Kirkhart, 2000).

The first set of empirical studies on evaluation use was conducted in the 1970s (e.g. Weiss, 1972; Knorr, 1977; Alkin et al., 1979; Weiss and Bucuvalas, 1980). Many more followed in the next decades. In popular literature reviews (Cousins and Leithwood, 1986; Shulha and Cousins, 1997; Johnson et al., 2009) over a hundred of empirical studies about evaluation use are analyzed. This rich body of research presents many potential factors of evaluation use. However, as Ledermann (2011) states, research on evaluation use is still disappointingly inconclusive. Therefore, there is still a need to explore the subject.

This article presents another case of evaluation use. It is based on the current state of play in the field and aims to provide another piece of evidence on evaluation use and factors influencing it. Original input that this research offers is that it presents the case of Poland – a country that so far might be perceived as a "white gap" on the map of evaluation use empirical studies[1]. Poland is an example of a country were evaluation practice in public policies was imposed by external regulation and has developed with no prior domestic tradition (Żuber, Bienias, 2008; Górniak, 2007; Olejniczak, 2009). Such a situation is typical in the European Union, where member states eligible for support from the Cohesion Policy (CP) are also obliged to evaluate effects of that assistance.

Tomasz Kupiec
EGO – Evaluation for Government Organizations S.C.
ul. Podmokła 53b, 04–819 Warszawa
t.kupiec@evaluation.pl

[1] Until 2013, when this research was conducted, the only exception was a study of Ferry and Olejniczak (2008), who analyzed the use of six CP evaluation reports.

Before CP was established, evaluation traditions were present only in some Nordic and northern European countries. Because of that, in many cases CP has become a medium transmitting evaluation practice (Bachtler, 2001; Bachtler and Wren, 2006). There were two main waves of evaluation practice diffusion through CP (European Commission, 2012; Bustelo, 2014): the first in 1980–90 to Southern Europe (e.g. Portugal, Spain, Greece) and the second from 2004 to Central and Eastern Europe (e.g. Poland, Hungary, Bulgaria), affecting in total more than 15 countries. Therefore, the case of Poland, and more specifically the evaluation of Polish regional operational programs under CP, may serve as an example of a more general phenomenon of EU member states facing "coercive adoption" of evaluation, which might influence actual evaluation use (Højlund 2014). Findings of this research may contribute to the discussion about how the evaluation systems used in EU member states and regions should be organized and operate. What is and could be the role of CP requirements in developing those systems?

EU Cohesion Policy is aimed at socio-economic development of a union of 28 countries, 274 regions and over 500 million citizens. In the programming period of 2007–2013, its budget amounted to €350 billion (35.7% of the EU budget). CP is implemented through multiannual socio-economic programs prepared and managed by member states/regions, but approved by the European Commission. Poland joined the EU in 2004 and has been the largest beneficiary of CP since 2007. Although evaluation practice in Polish public administration started only in 2004, it has developed rapidly since then[2]. The number of commissioned evaluations has risen from 5 in 2004 to more than 120 per year after 2007 (National Evaluation Unit and Skórska, 2011: 201). Much of this development is a contribution of 16 regional governments, which – due to the decentralization of CP implementation in 2007 – each became a Managing Authority for Regional Operational Program (MA ROP). The ROP 2007–2013 budget ranges from €491,000 to €1,867,000, and the scope of support in each program is fairly wide, including: SMEs, R&D,

information society, tourism, culture, environment, urban regeneration, transport, education, health. An MA's responsibilities include preparing, implementing and evaluating its program. To complete the task each MA established an evaluation unit (EvU). By the end of 2012, EvUs completed 236 evaluation studies, which is an average of 14.8 per region (Kupiec, 2014a). That proves that MA ROPs are capable of producing an evaluation report. Yet, so far there is no evidence that these reports and knowledge is actually used.

The article is divided in several parts. First, the idea of evaluation as a management tool is introduced. From this perspective, evaluation use is defined. In this part also an evaluation production and use model (which was a basis for organizing this study) is presented. The second part describes the design and methodology of this study. Part three presents the results: the degree of evaluation use and the characteristics of the evaluation process (barriers/factors influencing evaluation use).

Evaluation as a management tool

A regional development program, which involves setting goals, planning measures, allocating resources and finally implementation, is a clear example of a strategic management process. It is, therefore, reasonable to analyze program evaluation from the perspective of evaluation as a management tool.

My understanding of the link between management and evaluation can be explained best on the basis of evolutionary or learning school of strategic management, which says that the environment is too complex to formulate a strategy once and for all with a clear vision and goals. Strategies must be developed step by step, with subsequent adaptations and learning (Mintzberg et al., 1998). A tool for this learning, particularly in the public sector, may be evaluation (Dahler-Larsen, 2005; Furubo et al., 2002). Moreover, if strategic management is a process of constant learning and adaptations (Chakravarthy, 1982), evaluation should also be regarded as an ongoing activity – a sub-process of management conducted in parallel with other management activities, providing information on the quality of the implementation process (support

---

[2] However, it is still limited to the field of Cohesion Policy.

for operations management) as well as effects of intervention (support for strategic management). Such an understanding of evaluation is very close to that of Rist and Stame (2006), who argue that when talking about evaluation we should no longer refer to individual evaluation studies, but to streams of studies and data.

The use of evaluation

With evaluation defined as above, evaluation use should be regarded primarily as an overall impact of evaluation on the management process, that is the number and significance of program adaptations, modifications informed by evaluation. Both strategic and operations decisions (with a greater emphasis on the former) must be informed for evaluation use to be regarded as satisfactory. Referring to the popular taxonomies, it is rather *the use of findings* than *process use* (Alkin and Taut, 2003; Kirkhart, 2000; Cousins, 2003). Considering the trinity of instrumental, conceptual and symbolic use[3] (Patton, 1997; Shulha and Cousins, 1997; Henry and Mark, 2003; Weiss, 1980; Knorr, 1977), we should focus here on instrumental use. Conceptual use might, in a longer term, also lead to program adaptations, but this impact is indirect and hard to grasp. In this context, symbolic use, which only appears to have an impact on decision-making, is regarded as nonuse or misuse of evaluation (Alkin and Coyle, 1988).

Factors influencing the use of evaluation

A rich body of literature on evaluation use (mostly from the United States) offers a number of classifications of factors influencing the use of evaluation (e.g. Patton et al., 1977; Alkin et al., 1979; Alkin, 1985; Cousins and Leithwood, 1986; Shulha and Cousins, 1997; Johnson et al., 2009; Owen, 2007; Balthasar, 2006, 2009; Saunders, 2012) as well as models explaining evaluation use

(Johnson, 1980; Cousins and Leithwood, 1986; Johnson, 1998; Cousins, 2003; Mark and Henry, 2004). If less numerous, European sources also offer a number of potential factors influencing evaluation use. For example, studies of the European Commission's evaluations (Williams et al. 2002; EPEC 2005) suggested that evaluation use depends on: timing, the goal of evaluation, support from the management, quality of the research process and results, monitoring of the recommendations' implementation. Balthasar (2006, 2009) drew attention to the importance of the institutional context: the relation and distance between the potential user of the research and the evaluator. Dahler-Larsen (2012) and Højlund (2014) stressed the dependence of evaluation on its social and organizational context, and the latter developed a classification showing how the external pressure and internal propensity to evaluate impact the level and type of the research's use.

One of the general conclusions that can be drawn from the above is that there are two main sets of factors influencing evaluation use. The first is labeled *usability* (Saunders, 2012), *evaluation implementation* (Cousins and Leithwood, 1986; Johnson et al., 2009) or *characteristics of the evaluation* (Owen, 2007). It refers to the evaluation study, e.g. organization of the research process, quality, relevance and the reliability of the findings. The second is named *use* (Saunders, 2012), *decision or policy settings* (Cousins and Leithwood, 1986; Johnson et al., 2009), *characteristics of the settings* (Owen, 2007). It relates to the evaluation context, e.g. institutional capacity, organizational structure, the practice of using evidence and the political climate. Another common feature of the relevant papers is that they seem to focus on a single evaluation study and present the issue of evaluation production and use from the perspective of the researcher (who is responsible for the identification of the potential user, their engagement in the research process, and the quality of communication).

The model of evaluation production and use applied in this paper assumes the opposite: evaluation is a process producing a stream of studies and responsibility for its quality and evaluation use lies on the shoulders of the commissioning institution (or the evaluation unit within its

---

[3] Instrumental – decision makers use the evaluation findings to modify the evaluated; conceptual – the evaluation findings help the program staff understand the program in a new way; symbolic – when an evaluation is conducted to to legitimize a decision that has already been made.

structure). Still, the model is deeply rooted in the existing literature on the subject. The idea was to present a logical sequence between successive groups of explanatory variables, i.e. evaluation use factors, and the dependent variable, the actual use of evaluation. In this sense, the model draws on the work of Cousins and Leithwood (1993) and Johnson (1998). The model uses a process approach, i.e. subsequent modules correspond to the stages of the evaluation production process. This approach is adapted from the models of Mark and Henry (2004) and Olejniczak (2008). The difference, however, lies in the identification of the stages of the process, due to the understanding of evaluation as a holistic process, whose product is not a single study but a stream of them, as outlined above. The identification of factors included in each module was inspired by Cousins and Leithwood (1986, updated by Johnson et al., 2009) and Owen (2007). However, as a consequence of the model's assumptions, some factors relevant only in the context of a single study being developed were abandoned or reformulated to correspond to the idea of evaluation being a pro-

cess. Additionally, the idea of evaluation's dependence on the organizational context (Dahler-Larsen, 2012) was reflected in the model, e.g. by including factors concerning the characteristics of EvUs and the ability of an organization to learn.

The logic of the model assumes that evaluation must start with identifying goals, verbalizing what we expect from this tool, how it is to help us and with what. Next, the evaluation process must be planned, which includes laying down procedures, organizing structures (EvU), providing resources (human, financial). During the implementation, subsequent studies are planned, commissioned, conducted and approved. Later, their findings are disseminated, debated, accumulated and hopefully used. The functioning of the whole process needs to be controlled and modified in case of deviations from the expected effects. External determinants influence the way the process is planned, implemented and the way findings are used. On the other hand, the actual utility of the evaluation impacts the environment (e.g. managers' attitude toward evaluation).
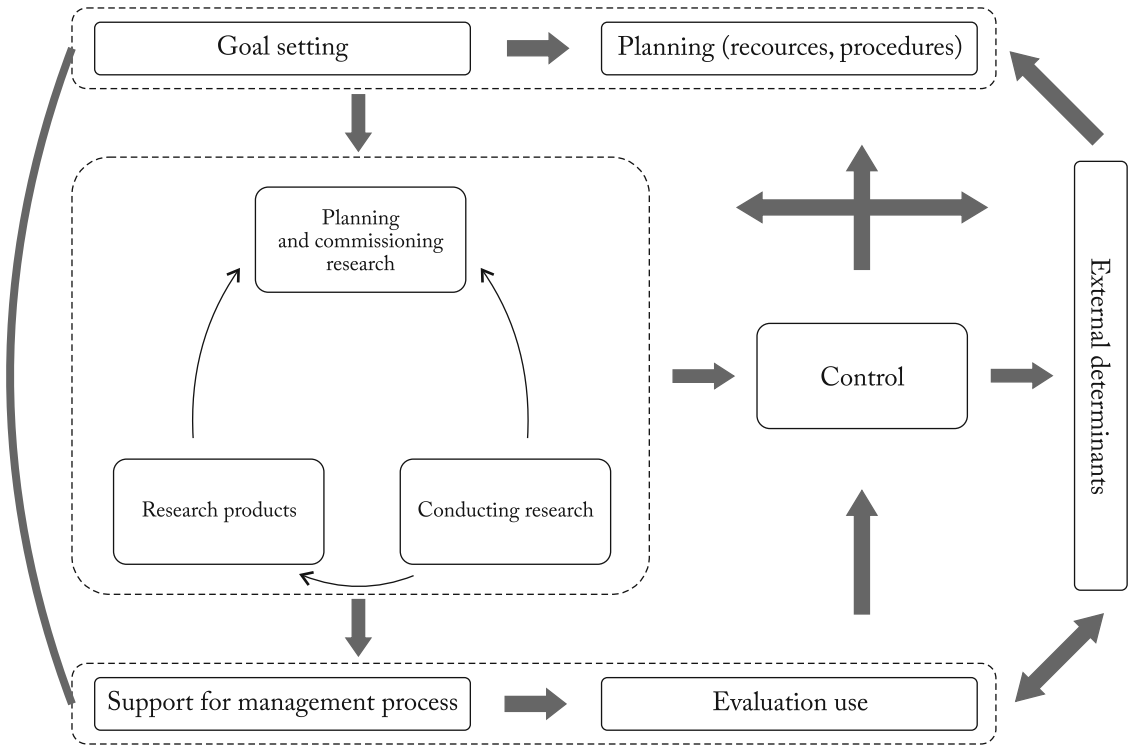

Figure 1 A model of evaluation production and use

The model includes:

- External determinants: the EvU's wider environment (decision-making characteristics, political climate, program flexibility, competing pieces of information) and the EvU's immediate environment (managers' approach (to evaluation), programming & implementing staff's approach, public procurement unit's approach);
- Goal-setting: Are there any objectives of the evaluation process as a whole? What are they? Who identifies them? Who is aware of them? (one factor)
- Planning of the evaluation process: characteristics of the EvU (capacity, place in the organizational structure, responsibilities), evaluation budget, evaluation procedures;
- Conducting of the evaluation process:
  - o planning and commissioning of the studies: relevance, methodological assumptions, characteristics of the contractor;
  - o conducting of the studies: quality of the research process, stakeholders' engagement, supervising contractor;
  - o outcomes of the studies: report quality (complete, credible, relevant), type of knowledge (revealing/obvious, controversial/acceptable, important/insignificant, diagnostic/prescriptive);

The evaluation process also includes efforts to build an evaluation culture:

- Support for the management process: dissemination of the findings, consideration of the findings, knowledge accumulation, managing and monitoring of the recommendations;
- Evaluation use – as explained above;
- Control of the evaluation process: the procedure and practice of carrying out cyclical analyses to check if the evaluation process's goals are met, and introducing modifications if needed (one factor).

It is important to realize that evaluation is a sub-process of a wider policy cycle (e.g. Bingham, Felbinger, 2002; Górniak, 2007) or (as in the cases described here) a process of managing public intervention/program (Haber, Szałaj, 2010). This relationship is reflected in some features of the model, although it intentionally focuses on a zoomed-in, detailed picture of the evaluation process. First of all, the evaluation use module refers to decisions that are in fact outside of the evaluation process but are informed by it. A number of factors comprising the model reflect how evaluation is organized within a wider management process and depends on it, e.g.: external determinants – decision-making in the organization, political climate, ROP flexibility; research planning – matching topics with information needs, research timing (especially delivering knowledge at a right time); support for the management process – consideration of findings (that is, the ability of the organization to learn). Some of those elements could be considered external to the evaluation process, yet they were included in the model because its basic function was not to precisely delineate the process but to identify factors enhancing and restricting evaluation use.

Design & Methodology

Out of 16 Polish regions managing regional operational programs under the EU Cohesion Policy, three cases were chosen. The strategy of "maximum case variation" (Flyvbjerg 2006) based on three criteria was applied (see Table 1).

All ROP evaluations in selected regions (region I – 20 studies, region II – 10 studies, region III – 14 studies) were analyzed in terms of evaluation use. A total of 440 recommendations arising from those studies were traced to check whether they had led to any modifications in the ROP or its management system. Modifications that could be linked to evaluation[4] were labeled "operational" (referring to improvements of the implementation process) or "strategic" (referring to the ROP's scope, its goals, measures, financial allocations, supported types of projects), and classified as: significant, medium, or minor. The basis for this analysis was desk research covering data from the recommendation management system and subsequent versions of the ROP and the ROP supplement. Additional questions about conceptual use were obtained from interviews.

---

[4] A modification was defined as resulting from evaluation if: 1) we could find a recommendation indicating the need for such a change, 2) the recommendation had been issued before the modification was introduced, 3) the recommendation was specific and clear enough to prove such a causal link.

Table 1 Characteristics of voivodeships selected for case studies

Voivodeship	I	II	III
Number of completed evaluation studies	20	10	14
Location of EvU within MA ROP's organization structure	Inside ROP monitoring unit; ROP implementation in different department	Inside ROP programming unit; ROP implementation in different unit of the same department	Inside ROP monitoring unit; ROP implementation in separate institution.
Relation between ROP and HCOP evaluation	ROP and HCOP evaluation in separate departments	ROP and HCOP evaluation in separate units of the same department	ROP and HCOP evaluation in the same unit

Source: own elaboration.

Four semi-structured, in-depth interviews were conducted in every region with: the chief of the department responsible for managing the ROP, individuals from the ROP managing and implementation units, a representative of the EvU[5]. Such a selection of interviewees allowed us to grasp the perspectives of knowledge producers and potential users. The structure of the interview protocol reflected the modules and detailed factors of the research model described above. The interviews were what we primarily used to characterize the evaluation process, with desk research (DR) playing a supplementary role. However, for some modules (research planning & commissioning, research outcomes – quality of the report) DR was the basic source of information. For others,(resources and procedures, type of knowledge) interviews and DR were equally important.

The analysis covered evaluation studies completed between 2007 and 2012. The interviews and the measurement of evaluation use were conducted between September and November 2013.

Results

Evaluation use

The MA ROPs in the three analyzed regions conducted 10 to 20 evaluation studies, costing PLN from 516,000 to 1,328,000 (EUR 123,000 – 316,000) in the period 2007–2012. In two regions the evaluation reports contained around 170, and in the third one – 107 recommendations addressed to the institutions involved in the implementation of the ROP. The highest share of officially implemented recommendations was 58% (95 out of 163), and in the other two regions it did not exceed 40% (42/107 and 40/170). An additional 16% (26/163) to 35% (60/170) of the recommendations were suggestions formulated for the future operational programs, nevertheless only part of them is possible to implement. In other words, evaluators in the analyzed regions proposed on average around 50 recommendations that would be implemented and another 25 which may be used in the future.

As suggested above, it is not recommendations but actual adaptations of the program and its implementation system that indicate the degree of evaluation use. Although one might expect that each implemented recommendation is equivalent to a modification, the reality in the analyzed regions is quite different. A large number of recommendations marked as implemented did not in fact result in any change. In most of those cases the recommended actions had been taken before the evaluation study was completed, with no relation to its conclusions. Another group of recommendations leading to no modification are very general ones, not indicating any specific action. For instance, "a reallocation of funds should be considered to support measures of higher importance for the beneficiaries" is a very general statement, and although in subsequent years the MA ROP decided to transfer funds between

---

[5] Respondents were contacted by mail and phone, and interviews were organized in the particular ROP MA's premises. Each of them lasted from 1h30m to 2h30m.

Table 2 Evaluation studies in analyzed regions

Region	I	II	III
No. of evaluation studies, including:	20	10	14
studies containing recommendations	19	7	9
cost of studies (PLN)	1 328 190	516 108	1 275 636
No. of recommendations, including:	170	163	107
implemented recommendations, including:	40	95	42
partly implemented recommendations	15	19	14
recommendations not implemented	70	42	28
recommendations to be implemented in the future, including:	60	26	37
recommendations possible to implement	30	19	33
Modifications/decisions supported by evaluation	20–40	30–50	20–30

Source: own elaboration.

program measures, it cannot be concluded that they were a consequence of this recommendation.

The number of adaptations that were introduced as a consequence of evaluation conclusions and recommendations amounted to 30 in region II and 20 in the other two regions. In the case of 10–20 modifications in each region we could not determine whether they resulted from evaluation or not. The number of adaptations does not seem too small, but their scope is a matter of concern. In fact, **all modifications implemented in the three analyzed regions as a result of evaluation have an operational character**, i.e. they focus on improvements of the ROP implementation system. The majority of them were also of minor or at most moderate importance. The subjects of the implemented modifications can be grouped into several areas:

- trainings for beneficiaries, e.g. focus on practical examples, reducing the size of training groups, different layout of information materials;
- information activities, e.g. creating a new tab on the website, creating FAQ, providing a tracked changes version when updating documents, providing information about typical errors in funding applications received during previous calls;
- promotional activities, e.g. use of ROP mascots, abandoning of short commercials in favor of longer ones, focusing on effective media channels;

- information points, e.g. clear marking of the offices, monitoring of activities, implementation of customer service standards;
- organization of work related to payment applications, e.g. shifting of tasks between employees, launching an application generator;
- monitoring indicators, e.g. changing the catalogue of indicators, clarifying indicator definitions, changing the target values.

Most of the modifications resulting from evaluation in all three regions refer to information and promotion activities. Interestingly, recommendations on this subject originate not only from studies evaluating information and promotion, but frequently appear also in evaluations of different issues e.g. the potential of local governments, local governments' activity related to applying for ROP funds, barriers to ROP implementation, ROP effects. Other recommendations, concerning monitoring indicators, handling payment or funding applications, project selection procedures, are also part of the organization of the ROP implementation process.

Evaluations carried out in the period studied, had virtually **no impact on strategic decisions, concerning the assumptions behind ROP intervention, defined problems and ways to solve them, identified objectives, priorities, financial allocations**. The only exception (out of 120 analyzed) – an adaptation which can be classified as strategic – took place in region III and concerned increasing preference in the selection

criteria for clusters that are at an early stage of development. It was, nevertheless, a minor modification with low impact on ROP objectives achievement[6].

By no means did non-use of evaluation in the process of strategic decision-making result from a lack of such decisions. The orientation of ROP support and financial reallocation were modified in all three regions, but the corresponding decisions were not informed by evaluation. Adaptations were introduced in ROP documents themselves and, in most cases, in ROP supplements (UROP), which contain detailed provisions on the scope of the support. ROP documents were modified once, in 2011, mainly in order to enter into the budget additional funds from the so-called National Financial Reserve and Technical Adjustment. Therefore, it was necessary to decide to what end the additional funds should be allocated. On this occasion some previous fund allocations were also withdrawn or reduced. UROPs were modified several times a year, and the total number of versions of that document adopted by the end of 2012 exceeded 20 in every region. The modifications covered: the maximum level of financial support, minimum level of beneficiaries' contribution, financial reallocations between measures, adding/removing/modifying types of supported projects or whole measures, establishing preferences for certain types of projects, changes in project selection criteria and in project selection procedures.

---

[6] The responsibility for that situation lies both with the evaluators and the EvUs/other officials with the MA ROP. Evaluators tend to suggest minor operational recommendations as a remedy for strategic issues (which may result from insufficient competences but also the fact that, as evaluators argue, such reports are more easily approved by EvUs). EvUs often suggest to evaluators that strategic modifications should have the implementation time set for the future perspective, as it is too late to implment them in the current program (the problem of timing is mentioned in the "evaluation use barriers" section). That results in an only around 15% share of strategic recommendations in a total number of them. One third of strategic recommendations had been postponed to ROP 2007-2013, but only 8% got actually implemented. Two thirds of strategic recommendations postponed to the next programming perspective, but only 29% are possible to implement (the rest is too deeply embedded in the context of ROP 2007-2013).

Apart from the large volume of decisions not informed by evaluation, there were also cases of evaluation misuse (e.g. the MA ROP in region I claimed that its decision was grounded in evaluation, although conclusions from the study were opposite to the modification made) or symbolic use (the MA in region II commissioned a study to evaluate transport, and then used it to justify its decision to increase support for transport, although the study focused on transport only and offered no conclusions to the effect that transport should be preferred over other support areas).

Searching for the cases of evaluation impacting strategic decisions, we analyzed potential modifications resulting from recommendations for future programs. Surprisingly, more than a half of them (and over 80% in region I) is again of an operational character and refer to the same areas as recommendations proposed for the present programs, e.g.: reshaping the system of monitoring indicators and selection criteria, creating procedures for partner projects, modifying the content of trainings for beneficiaries, improving the payment application generator, clarifying website content, or systemizing FAQ. One may suspect that at least part of the recommendations listed above was initially intended for the present implementation system, but it was easier for the MA to postpone implementation to the future than to officially reject them.

Strategic proposals for future programs are not numerous. Moreover, a vast majority of them seem to be of minor importance. They relate, for example, to technical issues (focus on integrated projects coordinated by different local governments, encouraging demonstration of coherence between ROP projects and the EU2020 strategy). Other suggestions offer solutions that are typical and obvious (support for subregional centers, differentiation of support for urban and rural areas, support for entrepreneurship), resulting from EU regulations (integrated, territorially oriented interventions). Some other proposals carry no potential for change since they focus on areas that had already been supported, and the intention behind the recommendations was just to give more preference to them when it comes to selection criteria (education and health in region II).

Examples of valuable and potentially useful strategic recommendations are so scarce that they can be all quoted here:

- region I – the list of 31 potential key projects worth funding in 2014–2020,
- region II – a detailed prescription regarding the support for transport (e.g. a project improving traffic safety, the concept of building a suburban railway network, a list of complementary investments to key rail projects),
- region III – a classification of municipalities into three types (converging, diverging and stable) and the suggestion to differentiate support depending on the type of municipality.

Based on the presented facts, it is reasonable to conclude that the degree of instrumental use and evaluation impact on decisions concerning management and implementation of ROPs is clearly unsatisfactory. Evaluation does not influence strategic modifications and the resultant operational adaptations are at most of medium importance. Findings from desk research are consistent with the views of IDI respondents – MA representatives, who are the primary users of evaluation. According to the director of the ROP management department in region II, the only outcome of evaluation are small improvements of minor issues in the implementation process. A representative of the programming unit said that evaluation findings are too general and shallow to be translated into decisions, and evaluation utility to date has not justified the level of spending on research. A management representative in region I also admitted that the use of evaluation generally comes down to unimportant recommendations and minor adjustments. When asked whether any significant adaptations were made based on the knowledge provided by evaluation, respondents said: "nothing comes to my mind", "not in the last two years when I was a manager", "I can't remember".

A relatively large number of operational adaptations is not sufficient for a positive assessment of instrumental evaluation use. An efficient system for ROP management and implementation is one in which it is more important where we head (strategic function) than how fast we go (implementation-absorption function) (Olejniczak, 2013). Therefore, an evaluation process which serves only to improve the implementation procedures is not enough. Moreover, in the case of operational recommendations related to the institutions of the ROP implementation system, their structure and procedures, we are facing a situation where evaluation users are much more familiar with the object of the study than researchers. As a result, a specific "deadweight" effect may occur. Evaluation users are aware of at least part of the problems that the study "identifies" before it actually begins, and then during interviews they provide solutions that are later presented to them in the evaluation reports. At least some of the modifications would therefore be introduced even without evaluation, and for the remaining, evaluation is only a general impulse, but not the source of specific solution proposals. Thus, the importance of evaluation is even lower than it is indicated by the simple sum of recommendations leading to modifications.

Conceptual use, or the evaluation's contribution to MA representatives' state of knowledge and perception of issues related to the ROP is also insignificant. In all analyzed regions respondents argued that evaluation had not offered them new insights into regional development and the way the ROP impact it. Only in region II respondents were able to point to examples of evaluation studies that they had learnt anything from (e.g. how ROP implementation looks from the beneficiaries' perspective and how small the impact of ROP investments is on healthcare compared to decisions of the National Health Fund). Again, only in region II evaluation results are sometimes used in discussions and debates as an argument to justify decisions. Just as in the case of instrumental use, conceptual use of evaluation should as well be considered as inadequate.

Evaluation use barriers

With such a low degree of evaluation use in all three observed regions, discussion about factors influencing this phenomenon can be reduced to the presentation of key barriers to evaluation use. Listed below are those elements of evaluation production and use that were assessed negatively in all three regions:

- Knowledge resulting from evaluation studies concerns mostly insignificant matters, or

"trifles, improving trifles in the implementation process", as one respondent called them. Most of the recommendations concern minor and technical issues.

– Conclusions following from evaluation are not revealing and explorative, but obvious and most often in line with the MA's earlier own observations. Conclusions and recommendations are often literal quotations from MA representatives' observations expressed during interviews. Respondents admitted that: "the majority of conclusions from evaluation can be expected before the study begins; evaluation reports often discover issues that have already been discovered and are obvious".

– Reports do not meet users' information needs. MA representatives admit that "evaluators often do not go in the direction we expected"; that the quality of some reports "was actually low or all we could do was to store them on the shelf". The analysis of evaluation reports shows that they often do not meet objectives set out in TOR and the methodological report, or meet them only partially. In the observed cases, conclusions and recommendations frequently focused on simple technical issues and the implementation process, although the evaluation objectives and questions were oriented at the program outcomes, and its impact on regional development and different sectors of regional economy.

– Many reports are incomplete, they do not provide answers to some key research questions. The frequently missing issue is ROP impact on the development of the region, and the situation of the different groups of beneficiaries, e.g. the competitiveness of enterprises. Instead of actual answers to these questions, reports often present less valuable content: a description of changes that have occurred in the region without demonstrating their connection to ROP intervention, subjective opinions of the beneficiaries on the effects of their projects, expert opinion not supported by any data.

– Poor quality of data analysis and conclusions. In most studies there are significant errors in data analysis, e.g. instead of the required quantitative analysis only selected cases are discussed, the research subject is only described and not evaluated, overall changes in regional economy are confused with ROP impact, opinions derived from interviews are not confronted with or supplemented by other sources, the effectiveness of different support forms is confused with the preferences of the beneficiaries.

– Scarce and inadequate data, particularly the shortage of official statistics. Data needed for ROP evaluation are often not collected at all, collected only partially, or provided with a very long delay.

– Low competence of evaluators. Respondents in region II argued that research teams often did not have enough capacity to prepare a valuable evaluation report. In region I it was said that the problem of insufficient competence and low commitment of evaluators occurred not in individual cases but, to a lesser or greater extent, virtually all studies.

– Research methodologies are not relevant to evaluation goals. Noted errors include: estimating ROP impact on the development of tourism based on the CATI with tourist projects beneficiaries, estimating additional complementarity effects based on the ITI with beneficiaries, no analysis of application funds, too small samples in quantitative analyses, missing counterfactual approach, replacing data analysis with survey questions.

– Difficulties with the study's timing. In many cases, evaluation findings were available too late to be used. On the other hand, a major portion of research on ROP effects was conducted too early to capture any outcomes of the intervention[7].

– Lack of process approach to ROP evaluation. In none of the analyzed regions there are any control procedures, indicators to measure the quality of implementation and effects of the evaluation process (goal achievement), and procedures to introduce modifications in case of unsatisfactory results of the control.

---

[7] An important factor is the durability of the research planning and conducting process. Five months on average pass between the commissioning of a study and receiving the final report, and that does not include the time needed to identify the information needs, transforming them into a research plan and preparing TOR.

Numerous and repeating divergences between evaluation reports and the expectations stated in the terms of reference raise the question why no effective measures were taken to solve that problem. Firstly, respondents from the EvU responsible for commissioning studies and approving reports do not have the impression that their quality is that low. Those who complain most are people from the ROP programming and implementation unit, but they are not sufficiently engaged in the research process to react. Secondly, as a part of a bureaucratic administration the EvU concentrates on the legality of the process more than on its utility. In most cases, reports meet the formal requirements, because the missing answers are not marked with blank pages in the reports. The fact that they are missing is hidden in pages full of obvious, irrelevant information. EvUs are overloaded with work, so they concentrate on what they are accountable for – producing a certain number of reports in accordance with the procedures[8].

Based on the research model, barriers described above can be arranged in a logical sequence of problems resulting from each other, which ultimately leads to the unsatisfactory level of evaluation utilization (Figure 2). Lack of significant and revealing conclusions is the result of incomplete reports, which, in most cases, miss answers to the most interesting but, at the same time, most difficult questions about the effects of ROP implementation. Incomplete reports stem from the lack of relevant data and a low quality of data analysis. Poor quality of data analysis may result from both the incompetence of the evaluators and the mismatch of methodologies and research objectives.

Along with the similarities – problems in the evaluation process which are present in all analyzed regions – there we also noted some differences that may affect the degree of evaluation use:

– A positive attitude toward evaluation distinguishes department director in region II. This person believes in the value of evaluation, which is confirmed by his involvement in the preparation, implementation and dissemination of findings of particular studies. The department director is a support for the EvU, and
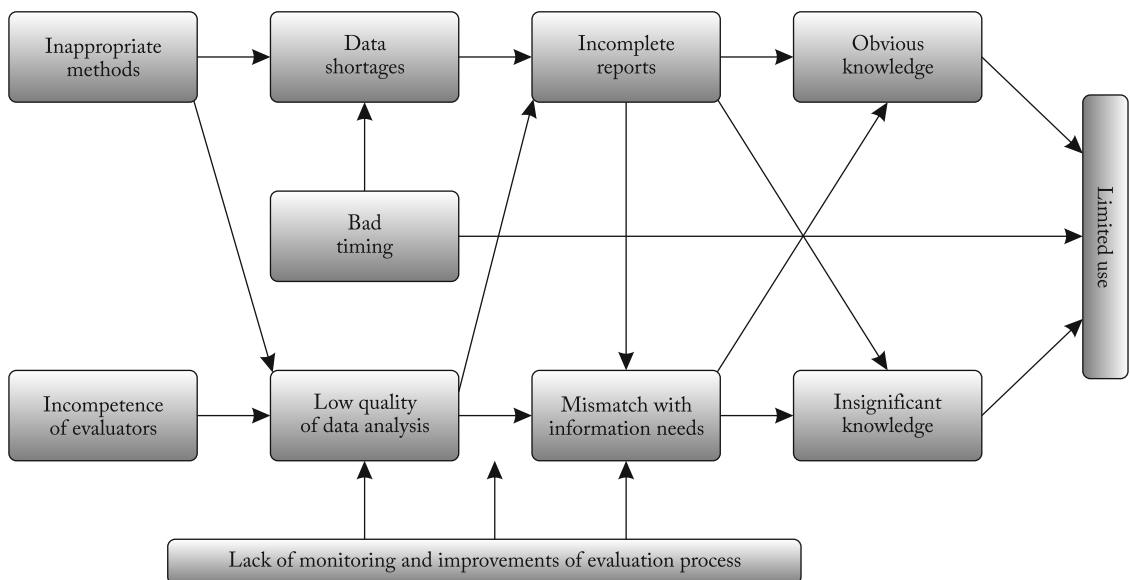

Figure 2 Barriers of evaluation use existing in all three regions

---

[8] It is important to note th vat the deficiencies of ROP evaluation do not result from a limited budget. All analyzed regions allocated substantial amounts for evaluation, of which less than a half was actually spent. In most cases, due to the competition between evaluators, the final price of the study was far below the limit set initially by the commissioning institution.

as its representatives declare "he never had to be convinced that evaluation in important, it is rather him who convinced others". In the other two regions directors rather demonstrate a lack of interest and awareness of the subject, or even are skeptical, and consider evaluation just a formal obligation.

– Representatives of the ROP management and implementation units as well as the EvU in region II demonstrate a better understanding of a wide range of goals that ROP evaluation may serve and are aware that evaluation is actually a management tool.

– The explored evaluation subjects fit best the information needs of the unit managing ROP in region II. The EvU in that region is also directly involved in managing ROP, so their research proposals are actually in line with the management unit's needs. As declared by respondents, the research subjects are always chosen in close cooperation with the department director and the ROP management unit manager. In the other two regions EvUs are separated from the management unit and the latter demonstrates a lack of interest in identifying evaluation topics. In most cases in the regions evaluation subjects are selected independently by the EvU.

The above facts interestingly correspond with our other observations. The EvU in region II is the only one that complains about the attitude of the public procurement unit in their organization, claiming that when the terms of reference are framed their proposals are often rejected or modified in a way that does not fit their needs. Respondents from the other regions do not feel that their needs are not met, but that might actually mean that their (knowledge) needs are not clearly specified or do not exist. Region I, where the process of identifying information needs was assessed the worst, is also the one with the most evaluation reports containing no recommendations, which may even intensify the impression that from the very beginning no one expected those studies to be used.

– Regions differ significantly in terms of the reflection practices – discussions and debates about evaluation findings, developing specific solutions and decisions on their basis. In re-

gion II such actions are taken at several stages: when the preliminary results are received and discussed with ROP managers, at the meeting of the Evaluation Steering Group, and when the recommendation table – stating which recommendations are to be adopted, how, when, and who is responsible for that – is prepared. At the same time, in region I evaluation users admit that they do not discuss implications of the evaluation findings.

– Similar differences can be observed with regard to recommendation management. A good practice of region II comprises determining how a recommendation can be used via dialogue between the EvU and evaluation users. Next, the regional board adopts a formal decision on the recommendation's use which is binding for the whole institution. The recommendation implementation process is monitored, and information is submitted to the ROP Monitoring Committee and the Evaluation Steering Group. At the other extreme is region I, where recommendation management in fact does not exist. Use depends on an individual decision of a particular user. Neither the department director nor the regional board are interested in this process or informed about it. An interviewed EvU representative admitted that he was not in favor of monitoring recommendation implementation.

– Collaboration and supervision of the research team in regions II and III take the form of: meetings prior to the preparation of the methodology report, control of the reliability of the data collection process (including contacting random respondents, observations of FGI). In region I an EvU representative admitted that he did not control contractors in any way because of time shortages. This difference in approach corresponds with staffing shortages, which are most noticeable in region I.

– Activities building the evaluation culture are conducted only in region II. Trainings on evaluation were organized there for all MA employees, MC and beneficiaries, at the beginning of the ROP implementation cycle. The ESG operating only in this region is also a means of developing evaluation capacity. Some of the evaluation findings were presented to a regional parliament commission.
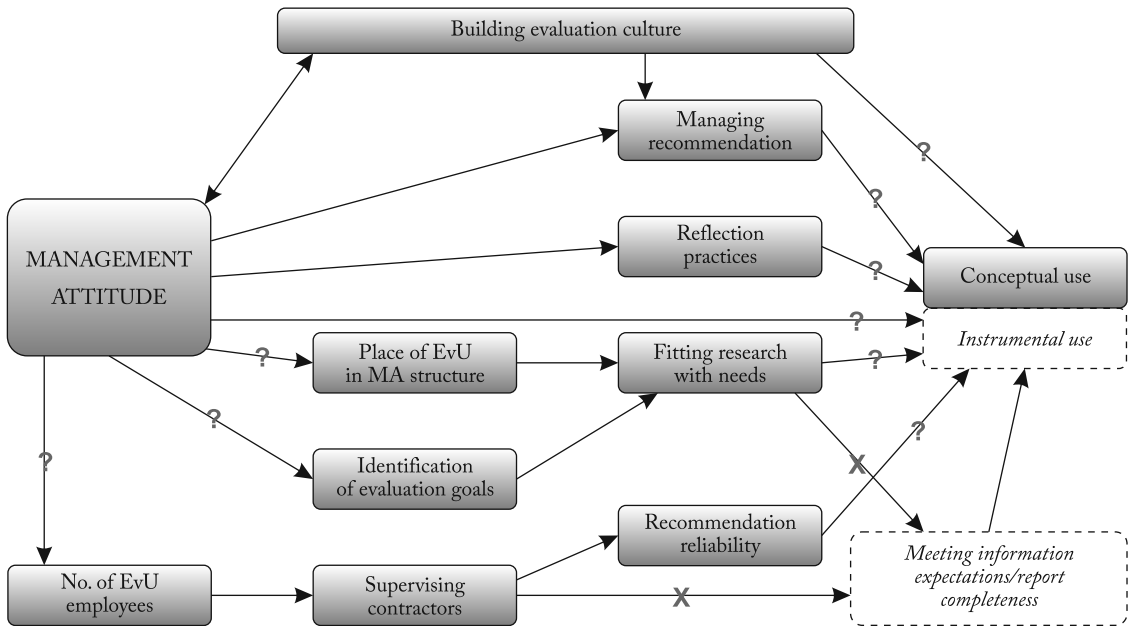
* Factors in grey boxes with solid contour were assessed positively in region II and negatively in the two others. Factors in white boxes with intermittent contour were assessed negatively in all three regions.

Figure 3 Factors supporting evaluation use in region II

– The credibility of recommendations and their coherence with evaluation conclusions are again the highest rated in region II. A negative view of this aspect in region I resulted from data analysis as well as subjective feelings of the respondents.

The fact that all differentiating elements have been positively evaluated in one region, and negatively in the others, leads to the supposition that these factors may be linked. The positive attitude of the management seems to improve the quality of the evaluation process. It ensures a better match between research subjects and the MA's information needs, facilitates a debate on the possible use of the findings and recommendation management (Figure 3). On the other hand, it has to be stressed that **coherence between information needs and evaluation subjects and solid supervision of the contractors does not guarantee a decent quality of the final reports** (crossed arrows in Figure 3). Reports do not meet evaluation users' expectations to more or less the same extent in all three regions.

It is not clear how the favorable attitude of the management and its positive consequences influ-ence the actual use of evaluation (question marks in Figure 3). As mentioned above, evaluation use is insufficient in all three regions. However, in the case of region II, where the evaluation process is plainly of higher quality, evaluation use is better. This applies particularly to the conceptual use. It can, therefore, be assumed that a positive attitude toward evaluation, commitment to the process of analyzing the potential use of findings (and, above all, the awareness of those findings), and systematic monitoring of the recommendation implementation process, facilitates or is even a necessary condition of taking advantage of valuable evaluation reports when they are provided.

Conclusions

Evaluation use in the process of managing ROPs in three analyzed Polish regions is clearly disappointing. A large number of conducted studies led only to minor modifications of the implementation process. Evaluation does not support any strategic decisions, although such decisions are made. As some of the MA repre-

sentatives admit, current benefits of evaluation do not justify the costs associated with it.

A direct and common cause of this situation is the poor quality of evaluation studies. Knowledge resulting from evaluation is usually obvious, insignificant, concerning minor issues that users are familiar with before the study begins. Reports are incomplete, they lack answers to some of the key research questions. Low quality of evaluation studies results from other problems, such as the incompetence of evaluators and inappropriate research methodologies.

Importantly, the problems described above – evaluation studies not fulfilling information needs, unreliable (due to inappropriate methodologies), obvious and insignificant conclusions – exist in all 16 Polish regions, which has been revealed by an analysis of a representative sample of 71 ROP evaluation studies (30% of studies completed in 2007–2012) (Kupiec, 2014b). This is a reason to suspect that insufficient evaluation use is very likely the case in all regions.

This research dealt with the details of the evaluation production process. As a consequence, it did not focus on a wider political and administrative context, which is ridden with a number of problems that may be behind the barriers to evaluation use identified above:

1. The first is the typical organizational culture of Polish administration in which still a perception of evaluation as a formal obligation is popular. It is notable that more than 10 years after joining the EU there are still no signs of disseminating the evaluation practice from CP to any domestic policy domain. Instead of evidence-based decisions, intuitive and arbitrary decisions dominate (Zybała, 2012). Two case studies showed that important decisions concerning ROP budget allocation were made against evaluation findings or with evaluation used only in a symbolic way[9].

2. The administration is interested in preserving the status quo (Mazur, 2014). Decisions once taken stiffen (Mazur, Olejniczak, 2012) and evaluation is rather used to justify and confirm than to question them.

3. Another reason is the lack of acceptance for failures in the EU and Polish administration. The pressure is on accountability, narrowly defined efficiency and predictability (Mazur, Olejniczak, Płoszaj 2013). Leaving no space for errors means also no space for experimenting, testing, evaluating, learning.

4. In the context of CP, this leads to the situation where regional authorities have an interest in presenting the ROP programming and implementation process in as positive a light as possible. The national government and the European Commission have no reason to deny it and question the success of structural funds. Therefore, there is little motivation for an evaluation process to be effective, able to highlight serious problems and propose significant corrections.

The the political and administrative context outlined leaves one with little hope that regional authorities might become interested in improving the ROP evaluation system. Yet the question about how it could be done is still worth consideration. The easiest and the most straightforward answer would be removing the identified evaluation use barriers. EvUs must refocus from report production to valuable knowledge production. One of the options to consider is reducing the number of commissioned studies. EvUs could then expect more from themselves (spend more time on identifying information needs, deciding on the research scope and methodology, supervising the contractor and disseminating and debating evaluation findings) and evaluators (not accepting reports based on poor data analysis or missing answers to key questions). A promising alternative for the "production of new studies" is meta-evaluation. It helps to develop evaluation capacity by identifying strengths and weaknesses in the current evaluation practice (Stufflebeam, 2001). By combining information from multiple studies, meta-evaluations may also serve as a more reliable and convincing base for decisions than a single evaluation (Cooksy, Caracelli, 2005). The broader and long-term goal should be to develop a comprehensive evidence-based policy approach of which evaluation is only one of core competencies along with: forward & outward looking, innovative and creative, using evidence, inclusive, joined up, reviews, lessons learnt (Office, 1999).

---

[9] In region II a study was conducted only to justify earlier decisions.

Funding

References

Alkin M.C. (1985). *A Guide for Evaluation Decision Makers*. Beverly Hills, CA: Sage.

Alkin, M.C., Coyle, K. (1988). "Thoughts on evaluation utilization, misutilization and non-utilization", *Studies in Educational Evaluation*, 14(3): 331–340.

Alkin, M.C., Daillak, R., White, P. (1979). *Using Evaluations. Does Evaluation Make a Difference?* Beverly Hills, CA: Sage.

Alkin, M.C., Taut, S.M. (2003). "Unbundling evaluation use", *Studies in Educational Evaluation*, 29: 1–12.

Bachtler, J. (2001). "Quod erat demonstrandum? Ewaluacja polityki regionalnej", *Studia Regionalne i Lokalne*, 7(4): 41–59.

Bachtler, J., Wren, C. (2006). "Evaluation of European Union Cohesion Policy: Research questions and policy challenges", *Regional Studies*, 40(2): 143–153.

Balthasar, A. (2006). "The effects of Institutional Design on the Utilization of Evaluation. Evidenced using Qualitative Comparative Analysis (QCA)", *Evaluation*, 12(3): 353–371.

Balthasar, A. (2009). "Institutional design and utilization of evaluation: A contribution to a theory of evaluation influence based on Swiss experience", *Evaluation Review*, 33(3): 226–256.

Bingham, R.D., Felbinger, C.L. (2002). *Evaluation in Practice: A Methodological Approach*. New York: Seven Bridges Press.

Bustelo, M. (2014). "Presidential editorial", *Evaluation*, 20(1): 3–9.

Chakravarthy, B.S. (1982). "Adaptation: A promising metaphor for strategic management", *The Academy of Management Review*, 7(1): 35–49.

Cooksy, L.J., Caracelli, V.J. (2005). "Quality, context, and use: Issues in achieving the goals of metaevaluation", *American Journal of Evaluation*, 26(1): 31–42.

Cousins, J.B. (2003). "Utilization effects of participatory evaluation", in: T. Kellaghan, D.L. Stufflebeam, L.A. Wingate (eds.), *International Handbook of Educational Evaluation*. Boston: Kluwer, pp. 245–265.

Cousins, J.B., Leithwood, K.A. (1986). "Current empirical research on evaluation utilization", *Review of Educational Research*, 56(3): 331–364.

Cousins, J.B., Leithwood, K.A. (1993). "Enhancing knowledge utilization as a strategy for school improvement". *Knowledge: Creation, Diffusion, Utilization*, 14: 305–333.

Dahler-Larsen, P. (2005). "Evaluation in public management", in: E. Ferlie, L.E. Lynn, C. Pollitt (eds.), *The Oxford Handbook of Public Management*. Oxford: Oxford University Press, pp. 615–639.

Dahler-Larsen, P. (2012). *The Evaluation Society*. Palo Alto, CA: Stanford Press.

EPEC (2005). *Study on the Use of Evaluation Results in the Commission*. Final report, Paris: European Policy Evaluation Consortium.

European Commission (2012) *EVALSED: The resource for the evaluation of Socio-Economic Development*, Guide, http://ec.europa.eu/regional_policy/sources/ docgener/evaluation/evalsed/guide/index_en.htm (access: 30.01.2013).

Ferry, M., Olejniczak, K. (2008). *Wykorzystanie ewaluacji w zarządzaniu programami unijnymi w Polsce*. Warszawa: Ernst & Young Polska.

Flyvbjerg B. (2006). "Five misunderstandings about case-study research", *Qualitative Inquiry*, 12(2): 219–245.

Furubo, J.E., Rist, R.C., Sandahl, R. (2002). *International Atlas of Evaluation*. New Brunswick: Transaction Publishers.

Górniak, J. (2007). "Ewaluacja w cyklu polityk publicznych", in: S. Mazur (ed.), *Ewaluacja funduszy strukturalnych – perspektywa regionalna*. Kraków, UE Katowice, MSAP, pp. 11–28.

Haber, A., Szałaj, M. (eds.) (2010). *Ewaluacja w strategicznym zarządzaniu publicznym*. Warszawa: Polska Agencja Rozwoju Przedsiębiorczości.

Henry, G.T., Mark, M.M. (2003). "Beyond use: Understanding evaluation's influence on attitudes and actions", *American Journal of Evaluation*, 24(3): 293–314.

Højlund, S. (2014). "Evaluation use in the organizational context – changing focus to improve theory", *Evaluation*, 20(1): 26–43.

Johnson, K.W. (1980). "Stimulating evaluation use by integrating academia and practice", *Science Communication*, 2(2): 237–262.

Johnson, R.B. (1998). "Toward a theoretical model of evaluation utilization", *Evaluation and Program Planning*, 21(4): 93–110.

Johnson, K., Greenseid, L.O., Toal, S.A., King, J.A., Lawrenz, F., Volkov, B. (2009). "Research on evaluation use: A review of the empirical literature from 1986 to 2005", *American Journal of Evaluation*, 30(3): 377–410.

Kirkhart, K.E. (2000). "Reconceptualizing evaluation use: An integrated theory of influence", *New Directions for Evaluation*, 88: 5–25.

Knorr, K.D. (1977). "Policymakers' use of social science knowledge: Symbolic or instrumental?", in: C.H. Weiss (ed.), *Using Social Research in Public Policy Making*. Lexington, MA: Lexington Books, pp. 165–182.

Kupiec, T. (2014a). "Evaluation practice of Regional Operational Programmes in Poland", *Management and Business Administration. Central Europe*, 126(3): 135–151.

Kupiec, T. (2014b). "Użyteczność ewaluacji jako narzędzia zarządzania regionalnymi programami operacyjnymi", *Studia Regionalne i Lokalne*, 56(2): 52–67.

Ledermann, S. (2011). "Exploring the necessary conditions for evaluation use in program change", *American Journal of Evaluation*, 33(2): 159–178.

Mark, M.M., Henry, G.T. (2004). "The mechanisms and outcomes of evaluation influence", *Evaluation* 10(1): 35–57.

Mazur, S. (2014). "Dowody naukowe w zarządzaniu publicznym", in: A. Haber, K. Olejniczak (eds.), *(R)ewaluacja 2. Wiedza w działaniu*. Warszawa: Polska Agencja Rozwoju Przedsiębiorczości, pp. 155–160.

Mazur, S., Olejniczak, K. (2012). "Rola organizacyjnego uczenia się we współczesnym zarządzaniu publicznym", in: K. Olejniczak (ed.), *Organizacje uczące się. Model dla administracji publicznej*. Warszawa: Wydawnictwo Naukowe Scholar, pp. 25–60.

Mazur, S., Olejniczak, K., Płoszaj, A. (2013). "Organizacyjne uczenie się w administracji publicznej – wnioski z porównania międzynarodowego, in: S. Mazur, A. Płoszaj (eds.) (2013), *Zarządzanie wiedzą w organizacjach publicznych. Doświadczenia międzynarodowe*. Warszawa: Wydawnictwo Naukowe Scholar, pp. 295–309.

Mintzberg, H., Ahlstrand, B., Lampel, J. (1998). *Strategy Safari. A Guided Tour Through the Wilds of Strategic Management*. New York: The Free Press.

National Evaluation Unit and Skórska, P. (2011). "Poland", in: K. Olejniczak, P. Strzęboszewski, S. Bienias (eds.), *Evaluation of Cohesion Policy: Overview of Practices*. Draft Report for Polish Presidency. Warsaw: Ministry of Regional Development, pp. 196–206.

Office, S.P.M.T.C. (1999). Professional Policy Making for the Twenty First Century. London: Cabinet Office.

Olejniczak, K. (2008). *Mechanizmy wykorzystania ewaluacji. Studium ewaluacji średniookresowych INTERREG III*. Warszawa: Wydawnictwo Naukowe Scholar.

Olejniczak, K. (2009). *Praktyka ewaluacji efektów programów rozwoju regionalnego – studium porównawcze*. Raporty i analizy EUROREG, nr 2/2009, Warszawa.

Olejniczak, K. (2013). "System zarządzania polityką – propozycja uniwersalnych ram koncepcyjnych". *Zoon Politikon*, 4: 41–55.

Owen, J.M. (2007). *Program Evaluation: Forms and Approaches*. London: The Guilford Press.

Patton, M.Q. (1997). *Utilization-focused Evaluation: The New Century Text* (3rd ed.). Thousand Oaks, CA: Sage.

Patton, M.Q., Grimes, P.S., Guthrie, K.M., Brennan, N.J., French, B.D., Blyth, D.A. (1977). "In search of impact: An analysis of the utilization of federal health evaluation research", in: C.H. Weiss (ed.), *Using Social Research in Public Policy Making*. Lexington, MA: Heath.

Rist, R., Stame, N. (eds.) (2006). *From Studies to Streams: Managing Evaluative Systems*. London: Transaction Publishers.

Saunders, M. (2012). "The use and usability of evaluation outputs: A social practice approach", *Evaluation*, 18(4): 421–436.

Shulha, L.M., Cousins, J.B. (1997). "Evaluation use: Theory, research, and practice since 1986", *Evaluation Practice*, 18(3): 195–208.

Stufflebeam, D.L. (2001). "The metaevaluation imperative", *American Journal of Evaluation*, 22(2): 183–209.

Weiss, C.H. (1972). "Utilization of evaluation: Toward comparative study", in C.H. Weiss (red.), *Evaluating Action Programs: Readings in Social Action and Education*. Boston, MA: Allyn & Bacon.

Weiss, C.H. (1980). "Knowledge creep and decision accretion", *Knowledge: Creation, Utilization, Diffusion*, 1: 381–404.

Weiss, C.H., Bucuvalas, M.D. (1980). *Social Science Research and Decision-Making*. New York: Columbia University Press.

Williams, K., de Laat, B., Stern, E. (2002). *The Use of Evaluation in the Commission Services. Final Report*. Paris: Technopolis France.

Zybała, A. (2012). "O lepszą jakość polityk publicznych", *Analizy i Opinie*, 127, http://www.isp.org.pl/uploads/analyses/2048186281.pdf (access: 12.10.2015).

Żuber, P., Bienias S. (2008). "System ewaluacji w Polsce – dotychczasowe doświadczenia i wyzwania na przyszłość", in: K. Olejniczak, M. Kozak, B. Ledzion (eds.), *Teoria i praktyka ewaluacji interwencji publicznych. Podręcznik akademicki*. Warszawa: Wydawnictwa Akademickie i Profesjonalne, pp. 84–98.

Wykorzystanie ewaluacji i jego mechanizmy w zarządzaniu regionalnymi programami operacyjnymi w polskich województwach

Artykuł omawia słabo dotychczas rozpoznaną kwestię wykorzystania ewaluacji w regionalnych programach operacyjnych wdrażanych przez polskie województwa. Analizę problemu poprzedziło założenie, iż ewaluacja jest długofalowym procesem produkującym strumień wiedzy wspierającej zarządzanie programem na wszystkich jego etapach. Na potrzeby badania zrealizowano trzy studia przypadków – programów regionalnych, instytucji nimi zarządzających i 44 ewaluacji zrealizowanych przez nie w latach 2007–2012. Wyniki wskazują na niski poziom wykorzystania ewaluacji, skutkującego jedynie drobnymi zmianami w procesie wdrażania. Bezpośrednimi barierami wykorzystania ewaluacji okazały się: niska jakość badań, oczywiste i mało znaczące wnioski, brak odpowiedzi na kluczowe pytania badawcze. Te z kolei wynikały z: braku kompetencji ewaluatorów oraz niewłaściwej metodyki badań.

Słowa kluczowe: wykorzystanie ewaluacji, ewaluacja programu, regionalne programy operacyjne, polityka spójności.