

* Karol Olejniczak, European Centre for Regional and Local Studies (EUROREG), University of Warsaw, Warsaw, Mazowieckie, Poland, k.olejniczak@uw.edu.pl

Tomasz Kupiec, Evaluation for Government Organizations s.c. (EGO s.c.), Warsaw, Mazowieckie, Poland, t.kupiec@evaluation.pl

Kathryn Newcomer, The Trachtenberg School of Public Policy and Public Administration, George Washington University, Washington, DC, USA, newcomer@gwu.edu

Learning from evaluation – the knowledge users' perspective

Karol Olejniczak, Tomasz Kupiec, Kathryn Newcomer*

Abstract

Public managers require different types of knowledge to run programs successfully. This includes knowledge about the context, operational know-how, knowledge about the effects, and causal mechanisms. This knowledge comes from different sources, and evaluation studies are just one of them.

This article takes the perspective of knowledge users. It explores to what extent evaluation is a useful source of knowledge for public managers of cohesion policy. Findings are based on an extensive study of 116 Polish institutions: surveys with 945 program managers, followed by 78 interviews with key policy actors. The article concludes that: (a) utility of evaluation studies, in comparison to other sources of knowledge, is limited, (b) evaluation reports are used to some extent as a source of knowledge on effects and mechanisms, however, (c) "effects" are shallowly interpreted as smooth money spending, not socio-economic change.

In conclusion, this article offers practical ideas on what evaluation practitioners could do to make evaluation more useful for knowledge users in policy implementation.

Keywords

Evaluation use, policy implementation, cohesion policy, evidence-based policies, knowledge utilization

Funding

The data analyzed in this study were collected as a part of a more comprehensive study of the management and implementation of cohesion policy in Poland. The study was commissioned by the Polish Ministry of Regional Development – National Evaluation Unit, co-financed by the European Union – European Regional Development Fund, and executed by the research company Evaluation for Government Organizations (EGO s.c.).

1. Introduction

The ultimate goal of evaluation is “social betterment” (Henry, Mark 2003; Christie 2007). It should be achieved by providing policy actors with research-based knowledge that will provide them with better understanding and improve targeting, designing, and implementing of public policies and programs. Ultimately, such evidence-informed policies and programs should be more effective in serving citizens.

In practice, this logic tends to be challenged by the complexity of policy implementation systems in at least four ways. First, actors engaging in policy implementation are highly diverse in terms of their backgrounds and objectives (politicians, bureaucrats, NGOs, media, experts), and positions in multi-level governance system (international agencies, national and regional actors). They have different goals, and they naturally have different knowledge needs.

Second, the spectrum of knowledge types required for running successful policies and programs spans from knowledge about the context in which the program is implemented, through technical know-how, to knowledge about the effects, and explanatory knowledge about causal mechanisms of socio-economic change (Ekblom, 2002; Nutley et al., 2003). No one type of evaluation inquiry can address this broad spectrum. Gathering these varied types of information must be tackled by different research approaches and cumulative evidence (Petticrew & Roberts, 2003).

Third, evaluation is just one of the sources of knowledge that policy actors use.¹ Policy actors can gain insights from many sources such as, to name just few, controls, audits, monitoring of programs, performance analysis, and informal contacts with beneficiaries, and knowledge exchange networks of public managers. These sources sometimes complement, but often compete for the policy actors' attention (Davies et al., 2010; Newcomer & Brass, 2016; Nutley et al., 2007; Weiss, 1980).

Last, evaluative insights, even of high relevance and quality, are not always incorporated in policy learning processes. Individual and organizational actors absorb information and learn in complex, non-linear ways (Argyris, 1977; Leeuw et al., 1994; Lipshitz et al., 2007; Olejniczak, Mazur, 2014; Weiss, Bucuvalas, 1980).

The challenge of aligning the production of evaluation studies with the knowledge needs of decision-makers has been the focus of both the theory and practice of evaluation utilization. The alignment challenge has been explored for over a decade (Weiss, Bucuvalas, 1980; Shulha, Cousins, 1997; Johnson et al., 2009). However, so far there has been limited attention given to the extent to which evaluation complements or competes with other sources of knowledge in complex systems of public program implementation. Therefore, the question addressed here is:

How useful are evaluation studies as a vehicles for promoting learning for actors involved in designing and implementing complex public polices?

To address this question we take the user-centered perspective. We frame evaluation as a service provided to "knowledge users", decision-makers involved in the implementation of the public policy and programs. We begin our article by providing a framework that positions evaluation practices in the complex system of multi-level policy implementation. We discuss the main types of knowledge that can be provided to different policy actors.

Second, we present findings on the main sources of learning for the staff responsible for implementation of a complex policy. We use the case

¹ As quoted by Weiss et al. (2005) "Evaluation is fallible. Evaluation is but one source of evidence. Evidence is but one input into policy..."

of the cohesion policy implemented in Poland (2007-13 programming period). We focus on the role of evaluation in the spectrum of knowledge sources about processes, effects and mechanisms of program delivery.

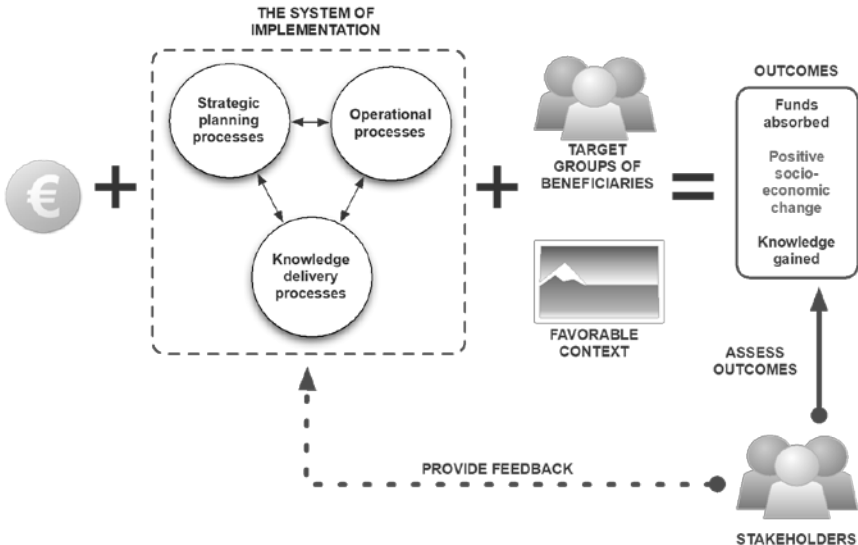
Third, we discuss the implications of our findings for evaluation practitioners who want to operate effectively in complex policy systems. We lay out key trade-offs that have to be addressed, and we draw upon experience of evaluation units in shaping the learning about cohesion policy in Poland and across the European Union.

2. Learning in complex policy and program implementation systems

2.1 Framework for understanding policy and program implementation

Program and policy implementation is a complex process that has been discussed in the management literature for many years, well before the introduction of the European Union cohesion policy (May 2003). We offer a general logic of public program and policy implementation in Figure 1.

Figure 1: The universal logic of program and policy implementation



Source: Inspired by Ostrom (2005).

The basic assumptions are that public funds are transferred, in the form of monetary aid or service activities, through a policy implementation system to certain target groups. The assumption is that if this aid is a receptive context, will trigger in a target groups desired behaviors, and the behaviors should eventually lead to a positive, sustainable socio-economic change. Thus, **the ultimate goal of policy** is desirable socio-economic change that addresses local challenges and problems, that the public funds are used to modify behaviors of those target groups that can bring this positive change.

The system of policy implementation is institutional, and involves procedural machinery of public agencies responsible for targeting promising beneficiaries and delivering aid smoothly (legally and on time). As we see in Figure 1, public institutions involved in this policy implementation system can engage in three groups of processes. First, "*strategic planning*", provides strategic documents, objectives and targets for interventions. It entails activities such as: (a) diagnosis and planning, (b) consultation and negotiations, (c) and coordination and alignment with the changing environment. In a cohesion policy terminology this is the domain of agencies assigned responsibilities of Coordinating Bodies and programming units within Managing Authorities.

Second, "*operational processes*", focus on spending and absorbing financial aid coming from the European Union. Operations cover sub-processes of (a) information and promotion given to beneficiaries – potential applicants of the projects, (b) application and selection of the most promising beneficiaries, (c) and financial management. In a cohesion policy terminology this is the domain for agencies called Financing Authorities, and Intermediate & Implementing Bodies.

Finally, "*knowledge delivery*" involves activities designed to produce knowledge to improve the system's operations (single loop learning), and to gain better understanding of socio-economic phenomena that are addressed by cohesion policy (double loop learning) (Argyris, Schon 1995; Fiol, Lyles 1985). Knowledge production encompasses evaluation, monitoring, performance auditing, and purchase and other expertise elaborated fully below. In cohesion policy it is assigned to monitoring and evaluation units, audit and control bodies.

The outcomes of policy implementation are typically measured by three indicators. The ultimate success indicator is positive socio-economic change. However, the observable effects of change are often delayed in time. Assessing policy implementation by measuring final outcomes is difficult. Thus, policy actors use more process-oriented indicators such as the level of funds absorption. They assume that the timely and legal use of public funds by beneficiaries is a proxy for successful of the policy implementation. In practice, this indicator is more likely to measure the efficiency of operational processes of the implementation system but will not measure actual rationality of strategic policy direction, nor the utility of the policy for local beneficiaries. The third indicator measures "knowledge gains" on lessons learned and mistakes that have been corrected or avoided over time. Such gains can be used in future planning for the next generation of policies and programs.

Stakeholders are expected to assess policy delivery and provide feedback to institutions of the implementation system. In the case of cohesion policy these stakeholders include the countries that are net payers of the policy, public opinion of the EU member states, media, interest groups and other European institutions, such as the European Commission and European Parliament.

2.2 Types of knowledge for public policy

In our analysis, we are especially interested in knowledge use and learning in policy delivery. Thus, we now focus on knowledge delivery processes within the system of policy implementation. In broad terms, there are five types of knowledge that may be produced in this setting (Nutley et al., 2003; Olejniczak et al., 2016):

- Knowledge about policy issues (*know-about*) – information about the spatial and temporal distribution of the socioeconomic problems, the needs, expectations and characteristics of targeted population;
- Knowledge about policy stakeholders (*know-who*) – awareness of which actors should be involved in the policy process to develop and implement solutions;

- Knowledge about effects (*know what*) – evidence on what policy approaches worked, what solutions and strategies produced desired outcomes in the past;
- Knowledge about change mechanisms (*know why*) – insights into why things work, and the causal mechanisms that lead to desired outcomes, as well as side effects;
- Knowledge on operational issues (know-how) – technical, operational knowledge about effective implementation procedures, activities and processes.

These five types of knowledge can be provided by different sources, such as evaluation studies, policy expertise, monitoring activities, and performance audits, etc. The actual producers of knowledge could be external actors to the policy implementation processes such as independent experts, research agencies, and audit companies or units within the policy implementation systems such as evaluation and monitoring units, and internal audit teams.

3. Learning in cohesion policy – empirical findings

3.1 Scope and method of the study

Our research deals with the issue of the **utility of evaluation studies as vehicles for providing learning for actors involved in designing and implementing public policy**. We focus on knowledge delivery activities that are within the system of implementation. In addition, we explore how these activities inform staff of the public institutions responsible for two other types of implementation activities – strategic planning and implementation processes (Figure 1).

Out of the five types of knowledge discussed above, we concentrate on three: know-what, know-how, and know-why. The rationale behind limiting attention to these three is that in theory evaluation may potentially provide these three types of knowledge. The other two (know-about and know-who) are the domain of different types of disciplined inquiry, especially policy analysis (Lincoln and Guba 1986).

As a case for our study, we have chosen the cohesion policy implementation system in Poland, in the programming period 2007 to 2013. Learning in the context of cohesion policy has been discussed extensively in number of publications (Batterbury, 2006; Rodrigues-Pose & Novak, 2013; Hojlund, 2014; Neacsu Petzold, 2015), however those studies do not explore the perspective of potential, individual users of knowledge. Cohesion policy is especially helpful to analyze knowledge use for several reasons. First, cohesion policy included number of multi-sectoral programs, ranging from labor market, trainings, institutional support, through enterprise innovation, to hard infrastructure. That broad scope makes its experience relevant to other public policies, as well as aid programs, across the world.

Second, there have been extensive evaluation activities undertaken to assess cohesion policy in Poland. A total of 976 evaluation studies have been completed through programming period (MIR 2014), making evaluation in Poland, at least in terms of volume, and an ample opportunity for policy learning.

Finally, European regulations guiding cohesion policy are standard across countries and regions of the European Union. Thus, the Poland case provides an opportunity for undertaking comparative studies to support generalization across national cases.

Our findings are based on a mixed-method descriptive research design executed on an extensive scale. The study covered all 116 institutions within the Polish cohesion policy system and it was part of the bigger ex post evaluation of cohesion policy implementation system in Poland.²

Our basic method was an online survey with public servants involved in program management. Heads of each institution received a link to the survey with a kind request to provide this link to experienced employees defined as "having at least 3 years of involvement in cohesion policy implementation, and 3 years of employment in the particular

² The full report EGO s.c. (2013) "Ocena systemu realizacji polityki spójności w Polsce w ramach perspektywy 2007-2013" is available on the National Evaluation Unit database: https://www.ewaluacja.gov.pl/media/24655/ggov_290.pdf

institution". In total, 945 responses were collected, with the majority of these from senior agency staff and heads of programs units. Referring to Figure 1, these were representatives of agencies running programs' strategic and operational processes (13 from Coordinating Bodies, 470 from Managing Authorities, 154 from Intermediate Bodies, and 308 from Implementing Authorities).

The survey respondents were asked to assess with 5-point scale (from strongly agree to strongly disagree) statements about the role of several potential knowledge sources for learning in their organization. Ten different potential sources of information were included in the survey, in addition to our subject of interest – evaluation studies. They were:

- 1) monitoring of physical progress,
- 2) monitoring of financial progress,
- 3) findings from project controls,
- 4) external controls (Supreme Court, tax office),
- 5) trainings, postgraduate studies,
- 6) conferences related to the area of respondents' work,
- 7) everyday contacts with program beneficiaries,
- 8) cooperation with other entities in National Strategic Reference Framework system (NSRF),
- 9) cooperation with national and international actors outside NSRF system, and
- 10) press articles.

Separate answers were given for each source with respect to each of three types of knowledge: implementation processes, program impact, and mechanisms of change (Questions are provided in Appendix A). To calculate the utility of each source we summed up the number selecting "strongly agree" and "agree" for each source.

The survey was complemented by a series of interviews (n=78) with key players in the system (usually department directors of leading institutions), and experts (mostly professors dealing with cohesion policy system or public administration). Referring to Figure 1, these were actors involved in strategic planning processes of the policy implementation and directors of departments involved in operational processes.

The interviews were designed to examine further the role of evaluation in learning. We asked interviewees about their main sources of information during different aspects of implementation processes (strategic, operational). We inquired if they remember particular studies that helped them in decision-making. All interviewees were asked to assess, on the scale 2 to 5 (the old Polish school grade system in which 2=unsatisfactory, 5=excellent), the utility of evaluation, monitoring, audits and controls for decision-making. Apart from grading, they provided justifications for their assessments and illustrated them with examples. We also asked about perceived improvement of knowledge delivery activities between two programming periods (2004–06 vs. 2007–13). For interviewees' answers, we have applied the magnitude coding method, to reduce qualitative data, and represent the interview data quantitatively (Saldaña, 2013).

The survey and interviews were conducted between April and June 2013 as part of more comprehensive study of the management and implementation of cohesion policy in Poland commissioned by the Polish Ministry of Regional Development.

3.2 Findings

In the EU programming period 2007–2013 Poland was the largest beneficiary of cohesion policy in the EU with an allocation of 67 bln € from a total budget of 347 bln € (19.3%). The entire cohesion policy package given to Poland was divided into five national Operational Programs and 16 regional Operational Programs. Each program had a distinctive structure of strategic goals and targeted groups of beneficiaries.

To deliver such an extensive and complex aid package to final beneficiaries, the largest national CP implementation system in Europe was established. The delivery system consisted of around 116 public institutions, and almost 12,000 civil servants involved strategic planning, operational processes and knowledge delivery (MIR 2013). In a cohesion policy terminology the implementing agencies were divided into Financing Authorities, Managing Authorities, Intermediate & Implementing Bodies.

Applying a user-center perspective, we assume that all of those 12,000 public agency staff involved in the cohesion policy implementation system could have been potential users of evaluation. The larger part of this population dealt with operational processes of Operational Programs (information and promotion given to beneficiaries, project selection and financial management), while a smaller group was responsible for strategic issues – including programs design, modification, and financial reallocations. However, since the competencies of these managers often overlap, we refer to both those groups together as “program managers”.

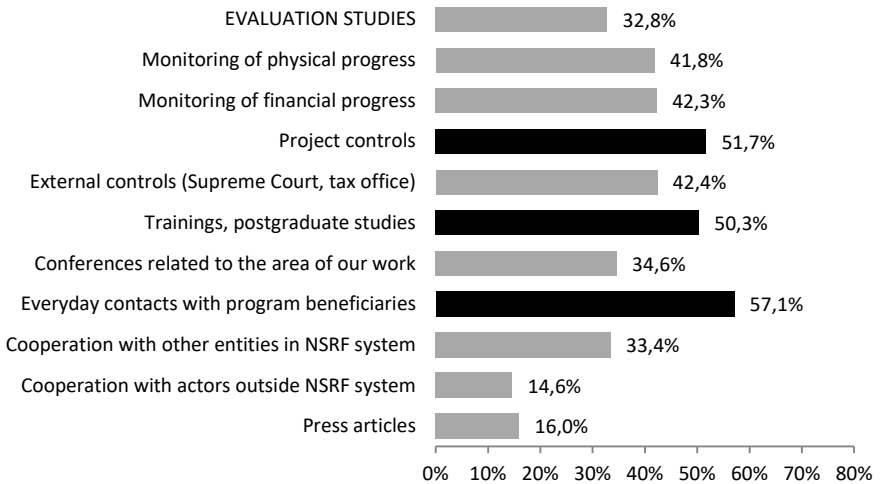
Regarding the production of evaluative knowledge, within the CP implementation system 59 evaluation units were responsible for planning and conducting evaluations (mostly commissioning execution of the studies to external contractors). Evaluation units were located in the structures of Managing Authorities, Intermediate & Implementing Bodies (National Evaluation Unit 2011). A total of 976 evaluation studies were completed through 2014, and the average number of studies completed per year in the period 2008–2014 was over 140 (MIR 2014). It is estimated that 40% of those studies were of a strategic nature, and the rest examined operational processes (EGO 2013). Given the large number of evaluations completed, users could have gained a substantial amount of useful knowledge. Yet, as we present below it was not necessarily so.

Let us now have a closer look at the users of evaluation – the program managers of cohesion policy in Poland. Out of 945 surveyed program managers, a little less than one third declared that they had learned about implementation processes from evaluation studies (Figure 2).

The most popular source of knowledge about program implementation appears to be everyday contacts with program beneficiaries and applicants. Two other sources were indicated by more than half of respondents: on-site project controls (performed by managing / implementing authority representatives), and rather surprisingly, training and post-graduate studies.

It is worth mentioning that in our survey we distinguished monitoring of physical progress from monitoring of financial progress³. Combined together they put monitoring at the top of the list of popular sources of information about program implementation process.

Figure 2: In our department/team we learn about implementation PROCESS from (% of answers “strongly agree” and “agree”)*



Note: *Marked in black three top answers

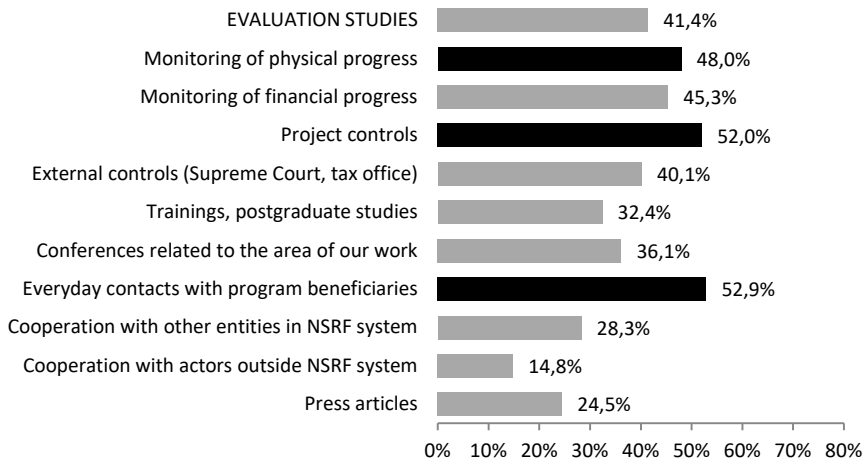
As one would have expected, the declared role of evaluation is greater when it comes to gaining knowledge about program effects, with over 41 % of respondents declaring it is useful (Figure 3). Evaluation studies

³ We believe it is justified, as such division is popular among program implementing units. It was also interesting to know which type of monitoring respondents had in their minds when they declared to learn from it about programme effects.

were only fifth most popular among 11 analyzed sources for learning about program effects.

As in the case of gaining knowledge about implementation processes, respondents most frequently report that their knowledge about program effects comes from feedback from beneficiaries. Second source indicated by more than a half of respondents are project controls, and third most used was monitoring of physical progress.

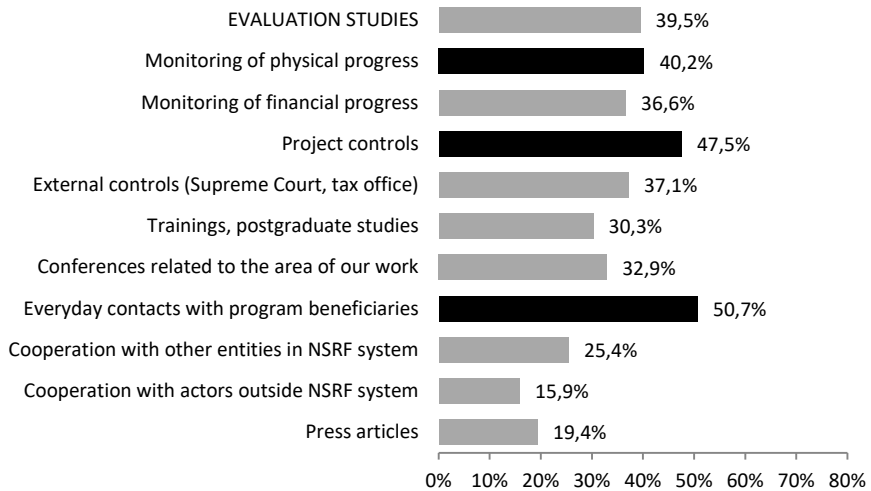
Figure 3: In our department/team we learn about program IMPACT from (% of answers “strongly agree” and “agree”)*



Note: *Marked in black three top answers

The survey results for knowledge about mechanisms of change are very much the same as in the case of knowledge about effects (98 % correlation). Evaluation studies fell 4th in terms of use, with less than 40 % of the respondents agreeing one may learn from evaluation studies about mechanisms (Figure 4). Everyday contacts with program beneficiaries were the most often used source again, and are the only option indicated by more than a half of the respondents.

Figure 4: In our department/team we learn about MECHANISM of change from (% of answers “strongly agree” and “agree”)*



Note: *Marked in black three top answers

The findings from interviews provide slightly more favorable pictures of evaluation than the survey. Our interviewees were asked to assess overall utility of evaluation on the old Polish school grade scale (from 2 to 5, where 2 is unacceptable and 5 is excellent). The average of this assessment was 3.6, which means evaluation as a source of knowledge "passed the utility exam", but only slightly above the acceptable minimum. This score is comparable to audit and control, but visibly lower than monitoring (4.0)⁴.

Evaluation was the most often mentioned source of knowledge about effects, and the majority of our respondents noticed improvements in evaluation over time (in comparison to 2004–06 period). Yet objections concerning the quality of evaluation reports are still frequent. Interviewees specifically mentioned the lack of analytical added value, simple repeating of monitoring data, ignoring organizational and legal limitations and obligations in policy system. Conclusions were often perceived as trivial, and recommendations deemed hardly useful, and often not meeting information needs.

⁴ Only these three sources of information were discussed during interviews.

Some respondents identified variations in evaluation utility depending on the type of the evaluation study. The most useful studies were those of diagnostic character, and involving program managers⁵. However, mid-term studies were considered more as routine obligations than responding to actual information needs. Most respondents could not recall any ex-post evaluations studies.

The usefulness of contacts with beneficiaries as a leading source of knowledge for program managers is interesting when combined with the finding that cooperation with actors outside of the cohesion policy system was the least useful source in all three cases, with evaluation studies falling in the middle. The managers in cohesion policy system seem to be inward looking, and rely on simple feedback signals generated from beneficiaries. Lack of interest in contacting and sharing knowledge with academia or officials involved in other public policies might suggest that actors in the cohesion policy system are not much interested in the impact of their programs in a wider socio-economic perspective. That observation corresponds with the fact that the leading sources are quite similar for all three types of knowledge. Managers tend to use information on implementations and even when asked about learning on program effects, respondents interpreted them as financial matters, implementation barriers, compliance with rules, and very basic outcomes.

4. Discussion and implication for evaluation practice

4.1 Discussion

For knowledge users, defined as staff of the agencies responsible for implementation of cohesion policy in Poland, evaluation studies were viewed as a limited source of knowledge. The main source for operational knowledge, as well as knowledge on what works and why, were everyday contacts with beneficiaries and project controls. These findings are not unique to cohesion policy. Others have found that everyday unstructured contact with beneficiaries provides preferred feedback on performance for public managers (Kroll, 2013). However, this

⁵ In fact those studies might not fit into the definition of evaluation and resemble more policy analysis.

source brings the risk of "availability heuristic," when managers build their understanding of the overall situation of the program on vivid stories of outspoken beneficiaries.

The fact that evaluation studies do not constitute one of the top choices when information is needed, may explain the recently observed phenomenon that, despite the large number of studies conducted, evaluation studies have almost no influence on decision-making process (Kupiec, 2016). If most decisions are not informed by evaluation findings, we may assume they are supported by other sources of knowledge.

The findings of this study also correspond well with other research on the cohesion policy evaluation system in Poland. One of the reasons why evaluation studies are not strong sources of learning for program managers may be the amount of time it takes to complete an evaluation study. Based on a sample of 235 studies Kupiec (2015) calculated that it takes seven months on average from data collection to the completion of an evaluation study. Program managers interviewed as part of that research reported:

- "evaluation takes too long to have real impact. When it comes to operational management, evaluation is useless, because it only repeats what we have already known, or what we have already changed. Waiting for evaluation recommendations is worse than making even bad decisions"
- "if there is a problem, and the answer is needed immediately, it is not possible to get it from evaluation, for procedural reasons."

As one can see the time, pressure is most evident in the case of informing operational decision-making. That is probably why the utility of evaluation studies was lowest in providing knowledge about implementation processes. However, it also raises a question why the majority of commissioned studies examined implementation issues (over 60 %). This question becomes even more intriguing when we realize that in the other 40 % of studies, which were supposed to deal with strategic problems, the vast majority of recommendations also focused on implementation processes (Kupiec, 2014). The lack of strategic recommendations may also account for the fact that evaluation studies are not found among

the more useful sources of knowledge about program effects and mechanisms leading to those effects.

4.2 Implications

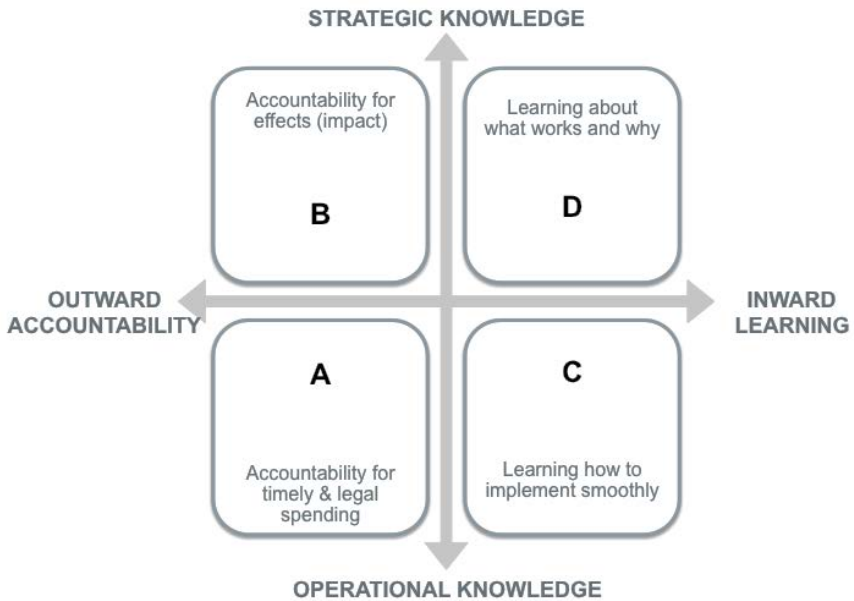
These rather sobering findings urge us to ask what evaluation practitioners can do to make evaluation more visible and useful for knowledge users in the policy and programs implementation processes. In order to address this challenge we think it would be beneficial to identify trade-offs that evaluation practitioners should address in order to provide useful information to decision-makers. In this section we especially focus on evaluation units, since they are central agents of evaluation activities in cohesion policy. However, our discussion may be also relevant for analysts and analytical units that serve other public policies.

Knowledge utilization might be usefully differentiated along two dimensions. First, there are different types of knowledge. On the one hand, evaluation studies may focus on bringing more strategic, broader knowledge about effects of policy and programs, and the mechanisms that produce successful outcomes. On the other hand, studies may focus on very technical, procedural and processual issues, thus providing policy practitioners with fine-grained operational knowledge.

The second dimension relates to the primary audience and evaluation objectives. Evaluation can be intended to inform the actors in the implementation system. In that case its primary function is learning, understood as improving strategic and operational activities over time. Or, evaluation can be intended to inform external audiences such as policy stakeholders. In the case of cohesion policy these are the European Commission, EU net payers, public opinion and the media.⁶ In that situation, evaluation may be used to hold policy implementation staff accountable to the stakeholders.

Those dimensions create four clear options for the evaluation units (Figure 5). Let us consider how the findings discussed above relate the framework.

⁶ It may as well be representatives of domestic authorities of particular country, if they are not involved in managing the programme being subject of evaluation.

Figure 5: Different uses of knowledge learned from evaluation

Source: own work

In Figure 5 studies in cell A focus on accountability for timely and legal spending. We believe that evaluation studies offer little additional value here because this area is well covered by control activities, as well as extensive monitoring systems developed over at the regional, national and European level of cohesion policy.

Studies in cell B focus on accountability for effects. Ex post evaluations are undertaken by European Commission. There is opportunity here for activities of the national evaluation units, as they could be aimed at showing to the public and main stakeholders the value for money of EU co-financed interventions. However, two issues could potentially limit evaluation units' actions in that area. First, stakeholders (especially the media and the public) could perceive the units as not fully independent and therefore not impartial, since they are located within the implementation system producing outcomes they try to measure. That, in turn, could render evaluation studies less credible in the eyes of the knowledge users. Second, assessing long-term effects means the studies must extend beyond one programming period. Longer-term

evaluations require institutional continuity of evaluation units. This is often not the case for cohesion policy since the units are parts of Managing Authorities assigned to the particular Operational Program. With new programming periods, new implementation structures are introduced.

Studies in cell C are promising for evaluation units as they can provide managers with balanced and objective views on on-going program implementation. Evaluation can help in tackling the managers' heuristics of availability – which means not making assumptions based on single stories from beneficiaries, but creating a more balanced and representative picture of reality. In that case, evaluation units could also, using a spectrum of organizational learning tools, analyze data to inform systematic data-driven reviews (Hatry, Davies, 2011; Olejniczak, 2015). During such sessions, held regularly, evaluation officers inform program managers, raise explanatory questions, and search for mechanisms that explain current implementation bottlenecks.

Finally, cell D is, in our view, the most promising for evaluation units. Evaluation studies could provide program managers – both strategic and operational staff, with insight on the actual effectiveness of theories of change that underlay certain interventions. And that in turn, would allow managers to correct interventions "on the go", providing them with data on target populations and change mechanisms so they could improve programs. For this purposes only evaluation undertaken by national and regional evaluation units could do the job, because those units are close enough to program managers to provide timely input.

However, evaluations in cell D also require evaluation units to tackle additional challenges. First, evaluation units need to educate their users in the implementation system. Our research shows that program managers frequently confuse products with effects. Evaluation units need to explain the difference, and convince managers of usefulness of looking beyond checklist of products, at the strategic goal of social change. In addition, evaluation units need to raise awareness among managers on the importance of knowing why and how interventions work – the mechanisms that change beneficiaries' reactions to the provided aid. This knowledge is crucial for the eventual success of implemented programs.

Secondly, evaluation units will have to work on the timing of their evaluations. They need to deliver timely explanations of mechanisms, and findings of first effects of interventions, in order to provide enough time for managers to react and incorporate the data to improve programming.

In the reality of cohesion policy, evaluation units will try to cover more than one option. However, it is important to be aware of the trade-offs and potential tensions since each of these options requires different levels of certain resources and skills, as well as demanding different roles from evaluators and their supervisors in evaluation units. Therefore, we encourage evaluation units to undertake strategic reflection and choose their primary focus. This would allow units to be more effective in their support of learning.

5. Conclusions

We have applied user-centered perspectives to the analysis of evaluation as a vehicle for promoting learning for actors involved in designing and implementing complex public policies. This means we relied on the declaration (surveys and interviews) of staff of public agencies and key actors involved the implementation of cohesion policy in Poland. The collected data shows that: (a) utility of evaluation studies, in comparison to other sources of knowledge, is limited, (b) evaluation reports are used to some extent as a source of knowledge on effects and mechanisms, however, (c) "effects" are shallowly interpreted as smooth money spending, not socio-economic change.

In our opinion, the crowded landscape of evidence sources as discussed above can be treated not only as a challenge but also as an opportunity for evaluation. Evaluation units across the cohesion policy system have experience, due to the scope of their work, with understanding social research, and speaking the languages of both policy and research. This comparative advantage gives them a unique opportunity to evolve from being mere contractors of isolated reports into real knowledge brokers – providing information to lead reflexive policy learning among decision-makers of cohesion policy.

We suggested that the limited resources of evaluation units in complex policy delivery system should be primarily focused on serving knowledge users who are responsible for policy implementation – both strategic and operational activities. An especially promising role would be increasing knowledge on mechanisms that drive programs' performance (what works and why). In terms of improving operational knowledge, evaluation units could support learning sessions that are based on monitoring data. Finally, evaluation units could explore more the synergies with other evidence-based sources of information, by synthesizing different knowledge sources and building policy arguments based on evidence. Such reorientation could hopefully lead to a situation where evaluation and other sources of knowledge complement each other, while conclusions from evaluation studies have visible utility in policy decision-making process.

Acknowledgements

The authors would like to express their gratitude to the members of the research team involved in data collection: Bartosz Ledzion, Andrzej Krzewski, Anna Borowczak, Marek Kozak, Paweł Kościelecki, Katarzyna Seferyńska, Paweł Śliwowski, Anna Domaradzka and Łukasz Widła. The authors would also like to thank Piotr Strzeboszewski and Stanisław Bienias from Polish National Evaluation Unit, and two anonymous reviewers of this article for their critical comments.

Appendix 1 – survey questions

The following sets of survey questions used to measure evaluation use has been an excerpt from a bigger survey that explored all three aspects of Cohesion Policy implementation system (strategic, operational and knowledge delivery). Survey has been administered on-line.

In our department/team we learn about implementation process from:	Strongly agree	Agree	Neither agree nor disagree	Disagree	Strongly disagree
Evaluation studies					
Monitoring of physical progress					
Monitoring of financial progress					
Project controls					
External controls (Supreme Court, tax office)					
Trainings, postgraduate studies					
Conferences related to the area of our work					
Everyday contacts with program beneficiaries					
Cooperation with other entities in NSRF system					
Cooperation with (inter)national actors outside NSRF system					
Press articles					
In our department/team we learn about program impact from:	Strongly agree	Agree	Neither agree nor disagree	Disagree	Strongly disagree
Evaluation studies					
Monitoring of physical progress					
Monitoring of financial progress					
Project controls					
External controls (Supreme Court, tax office)					
Trainings, postgraduate studies					
Conferences related to the area of our work					
Everyday contacts with program beneficiaries					
Cooperation with other entities in NSRF system					
Cooperation with (inter)national actors outside NSRF system					
Press articles					

In our department/team we learn about mechanism of change from:	Strongly agree	Agree	Neither agree nor disagree	Disagree	Strongly disagree
Evaluation studies					
Monitoring of physical progress					
Monitoring of financial progress					
Project controls					
External controls (Supreme Court, tax office)					
Trainings, postgraduate studies					
Conferences related to the area of our work					
Everyday contacts with program beneficiaries					
Cooperation with other entities in NSRF system					
Cooperation with (inter)national actors outside NSRF system					
Press articles					

Bibliography

- [1] ARGYRIS, C. and D. A. SCHON. *Organizational Learning II: Theory, Method, and Practice*. Reading, MA: FT Press, 1995.
- [2] ARGYRIS, C. Double-loop learning in organizations. *Harvard Business Review*, Vol. 55, No. 5, pp. 115-125.
- [3] BATTERBURY, S. Principles and Purposes of European Union Cohesion policy Evaluation. *Regional Studies*, Vol. 40, No. 2, pp. 179-188.
- [4] CHRISTIE, C. A. Reported Influence of Evaluation Data on Decision Makers' Actions: An Empirical Examination. *American Journal of Evaluation*, Vol. 28, No. 1, pp. 8-25.
- [5] DAVIES, H., S. NUTLEY and I. WALTER. Using evidence: how social research could be better used to improve public service performance. In: WALSHE, K., G. HARVEY and P. JAS (ed.). *Connecting Knowledge and Performance in Public Services: From Knowing to Doing*. Cambridge: Cambridge University Press, 2010. pp. 199-225.
- [6] EGO. *Ocena systemu realizacji polityki spójności w Polsce w ramach perspektywy 2007-2013*. Warszawa: Ministerstwo Rozwoju Regionalnego, 2013.
- [7] EKBLÖM, P. From the Source to the Mainstream is Uphill: The Challenge of Transferring Knowledge of Crime Prevention Through Replication, Innovation and Anticipation. *Crime Prevention Studies*, Vol. 13, pp. 131-203.
- [8] FIOL, M. and M. LYLES. Organizational learning. *Academy of Management Review*, Vol. 10, No. 4, pp. 803-813.
- [9] HATRY, H. and E. DAVIES. *A Guide to Data-Driven Performance Reviews*. Washington D.C.: IBM Center for The Business of Government, 2011.
- [10] HENRY, G. T. and M. M. MARK. Beyond Use: Understanding Evaluation's Influence on Attitudes and Actions. *American Journal of Evaluation*, Vol. 24, No. 3, pp. 293-314.
- [11] HOJLUND, S. Evaluation use in the organizational context - changing focus to improve theory. *Evaluation*, Vol. 20, No. 1, pp. 26-43.
- [12] JOHNSON, K., L. GREENSEID, S. TOAL, J. KING, F. LAWERNZ and B. VOLKOV. Research on Evaluation Use: A Review of the Empirical Literature from 1986 to 2005. *American Journal of Evaluation*, Vol. 30, No. 3, pp. 377-410.
- [13] KROLL, A. The Other Type of Performance Information: Nonroutine Feedback, its Relevance and Use. *Public Administration Review*, Vol. 73, No. 2, pp. 265-276.

- [14] KUPIEC, T. Użyteczność ewaluacji jako narzędzia zarządzania regionalnymi programami operacyjnymi. *Studia Regionalne i Lokalne*, Vol. 56, No. 2, pp. 52-67.
- [15] KUPIEC, T. Ewaluacja regionalnych programów operacyjnych w warunkach prawa zamówień publicznych i finansów publicznych, Samorząd Terytorialny, 10/2015, pp. 27-39.
- [16] KUPIEC, T. Program evaluation use and its mechanisms: The case of Cohesion Policy in Polish regional administration. *Zarządzanie Publiczne*, Vol. 33, No. 3, pp. 67-83.
- [17] LEEUW, F. L., R. C. RIST and R. C. SONNICHSEN. *Can governments learn?: comparative perspectives on evaluation & organizational learning*. New Brunswick; London: Transaction Publishers, 1994.
- [18] LINCOLN, Y. S. and E. G. GUBA. Research, Evaluation, and Policy Analysis: Heuristics for Disciplined Inquiry. *Policy Studies Review*, Vol. 5, No. 3, pp. 546-565.
- [19] LIPSHITZ, R., V. J. FRIEDMAN and M. POPPER. *Demystifying Organizational Learning*. Thousand Oaks: Sage Publications, Inc, 2007.
- [20] MAY, P. Policy Design and Implementation. In: PETERS, B. G. and J. PIERRE, eds. *Handbook of Public Administration*. London: Sage Publications, 2003.
- [21] MIR. *Potencjał administracyjny system instytucjonalnego Narodowych Strategicznych Ram Odniesienia na lata 2007-2013 (stan na 30 czerwca 2013 r.)*. Warszawa: Ministerstwo Infrastruktury i Rozwoju, 2013.
- [22] MIR. *Process of evaluation of the Cohesion Policy in Poland 2004-2014*. Warsaw: Ministry of Infrastructure and Development, 2014.
- [23] National Evaluation Unit. *Process of Cohesion Policy Evaluation in Poland*. Warsaw: Ministry of Regional Development, 2011.
- [24] NEACSU, M. and W. PETZOLD. Policy learning and transfer in EU Cohesion Policy: the impact of events. Paper presented at the Regional Studies Association Conference "Cross-national policy transfer in regional and urban policy", 19 January, Delft, The Netherlands.
- [25] NEWCOMER, K. and C. BRASS. Forging a Strategic and Comprehensive Approach to Evaluation Within Public and Nonprofit Organizations: Integrating Measurement and Analytics Within Evaluation. *American Journal of Evaluation*, Vol. 37, No. 1, pp. 80-99.
- [26] NUTLEY, S., I. WALTER and H. T. O. DAVIES. From Knowing to Doing. A Framework for Understanding the Evidence-Into-Practice Agenda. *Evaluation*, Vol. 9, No. 2, pp. 125-148.

- [27] NUTLEY, S. M., I. Walter and H. T. O. Davies. *Using Evidence: How research can inform public services*. Bristol: Policy Press, 2007.
- [28] OLEJNICZAK, K. and S. MAZUR (ed.). *Organizational Learning. A Framework for Public Administration*. Warsaw: Scholar Publishing House, 2014.
- [29] OLEJNICZAK, K. Focusing on Success: A Review of Everyday Practices of Organizational Learning in Public Administration. In: BOHNI NIELSEN, S., R. TURKSEMA. and P. van der KNAAP (ed.). *Success in Evaluation*. New Brunswick: Transaction Publishers. 2015. pp.99-124..
- [30] OLEJNICZAK, K., E. RAIMONDO and T. KUPIEC. Evaluation units as knowledge brokers: Testing and calibrating an innovative framework. *Evaluation*, Vol. 22, No. 2, pp. 168-189.
- [31] OSTROM, E. *Understanding institutional diversity*. Princeton, N.J.; Woodstock: Princeton University Press, 2005.
- [32] PETTICREW, M. and H. ROBERTS. Evidence, hierarchies, and typologies: horses for courses. *Journal of Epidemiology and Community Health*, Vol. 57, No. 7, pp. 527-529.
- [33] RODRÍGUEZ-POSE, A. and K. NOVAK. Learning processes and economic returns in European Cohesion policy. *Investigaciones regionales*, Vol. 25, pp. 7-26.
- [34] SALDAÑA, J. *The coding manual for qualitative researchers*. London-Singapore: Sage Publications, 2013.
- [35] SHULHA, L. M. and B. J. COUSINS. Evaluation Use: Theory, Research, and Practice Since 1986. *Evaluation Practice*, Vol. 18, No. 3, pp. 195-208.
- [36] WEISS, C. H., E. MURPHY-GRAHAM and S. BIRKELAND. An Alternate Route to Policy Influence: How Evaluations Affect D.A.R.E. *American Journal of Evaluation*, Vol. 26, No. 1, pp. 12-30.
- [37] WEISS, C. H. and M. J. BUCUVALAS. Truth Tests and Utility Tests: decision-makers' frame of reference for social science research. *American Sociological Review*, Vol. 45, no. 2, pp. 302-313.
- [38] WEISS, C. H. Knowledge Creep and Decision Accretion. *Science Communication*, Vol. 1, No. 3, pp. 381-404.