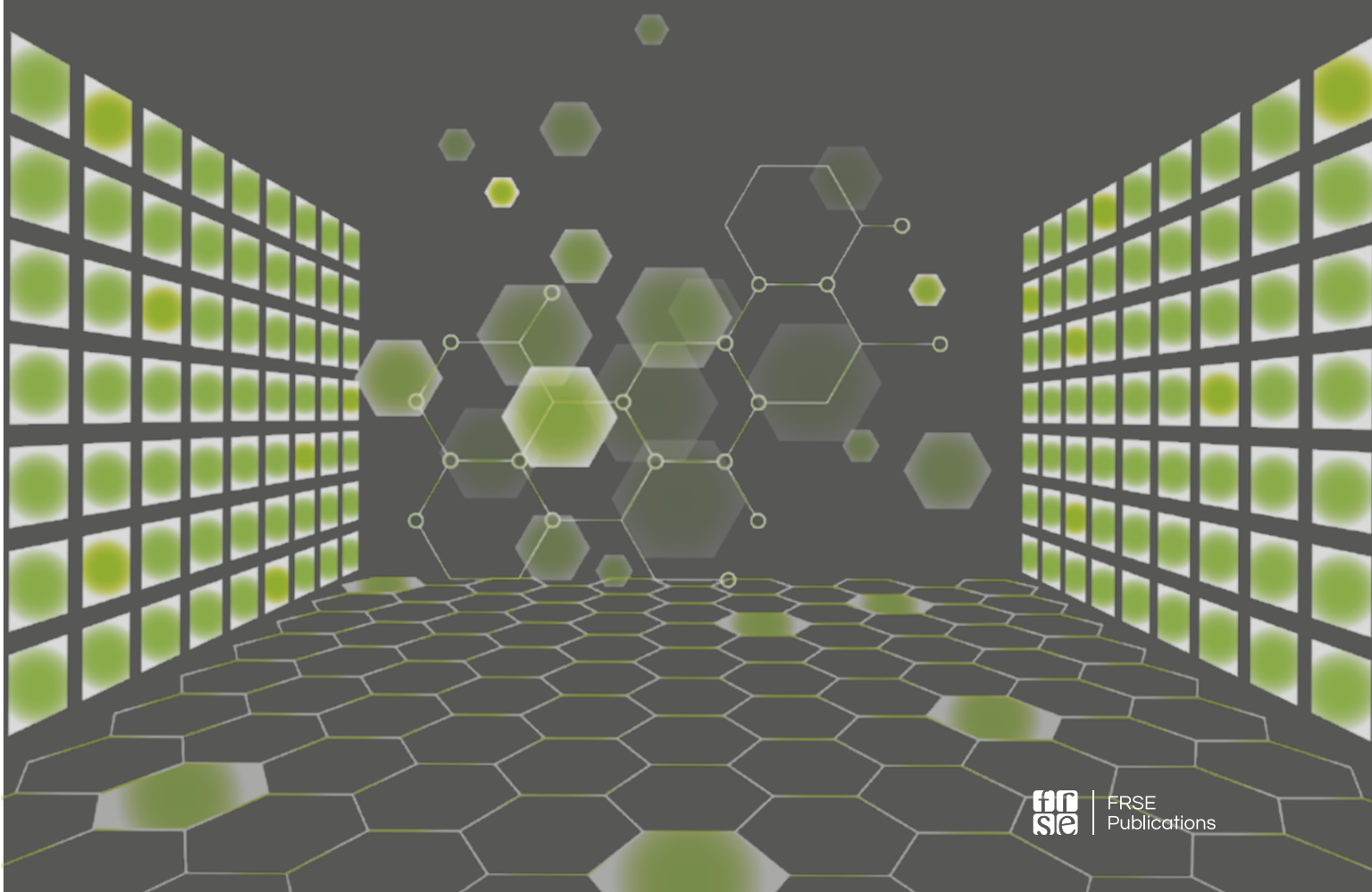
FRSE THEMATIC REPORTS

Evaluation in Educational Practice



FRSE Publications

Evaluation in Educational Practice

FRSE Publications
Warsaw 2020

FRSE THEMATIC REPORTS
Evaluation in Educational Practice

Edited by: Beata Ciężka
Agnieszka Rybińska

Translation: POLIGLOTA Biuro Tłumaczeń
Proofreading: Leila Chenoir
Production editor: Barbara Jędraszko

Design: Podpunkt
Cover design and DTP: Artur Ładno
Printed by: Drukarnia KOLUMB Chorzów

Publisher: Foundation for the Development of the Education System
Polish National Agency for the Erasmus+ Programme
Al. Jerozolimskie 142a, 02-305 Warsaw
www.frse.org.pl | kontakt@frse.org.pl
www.erasmusplus.org.pl

ISBN 978-83-66515-24-6

DOI 10.47050/66515246

This publication has been developed with the financial support from the European Commission in the framework of the Erasmus+ programme. The information and views set out in this report are those of the authors and they do not necessarily reflect the official opinion of the European Union. Neither the European Union institutions, neither the bodies nor any person acting on their behalf may be held responsible for the use which may be made of the information contained therein.

Free copy

Citation: Ciężka, B. & Rybińska, A. (eds.). (2020). *Evaluation in Educational Practice.* Warsaw: FRSE Publishing House. DOI: 10.47050/66515246

Table of Contents

Introduction

Agnieszka Rybińska
Research and Analysis Department Director at the Foundation for the Development of the Education System. She has extensive experience in the coordination and monitoring of SME support programmes (Polish Agency for Enterprise Development, Bank Gospodarstwa Krajowego). In the years 2011–-2016 she was Team Manager at the Educational Research Institute. Her research covers primarily the evaluation of public programmes, including those addressing support for SMEs and education. In recent years, she has been engaged in evaluation of various aspects of education, including transit from school to work, and school assessment.

Evaluation in Poland has been developing for nearly thirty years. Although it is used in all sectors, ranging from public administration to private and non-governmental organisations, it is most often associated with projects funded by the European Union. It is on this basis that most evaluation-related experience has been gained in recent years. However, for several years, evaluation practice has no longer been limited to projects receiving EU funding. The usefulness of evaluation is also starting to be appreciated when carrying out other activities, especially educational and developmental ones.

The Polish Evaluation Society was founded 20 years ago in response to the need to consolidate the evaluation community and to create links with people interested and involved in evaluation in Poland. The Society's mission is to promote evaluation in Poland and to create a culture of evaluation perceived as a democratic and social process.

The operations of the Foundation for the Development of the Education System (FRSE) are in tune with the mission of the Polish Evaluation Society. Research conducted by the Foundation allows for a better evaluation of the outputs of the European educational programmes it manages, including Erasmus+. It also provides feedback informing both further activities and programme beneficiaries.

This publication is a product of cooperation between the two organisations. It focusses on evaluation in the field of education and features chapters presenting the current status of evaluation in Poland and methodology approaches which can support the development of evaluators' work. The 20th anniversary of the establishment of the Polish Evaluation Society is a good opportunity to take a look at evaluation practices currently used in Poland.

Edyta Boratyńska-Karpiej authored the first chapter, which serves as an introduction to evaluation and emphasises its applications and implementation opportunities. In the text entitled *Quality and usefulness of internal and external evaluation. When do we learn more: commissioning an evaluation or carrying it out ourselves?* the author attempts to answer the question whether evaluation can be a tool for providing knowledge and undertaking a critical analysis of actions taken and effects achieved, or whether its role is limited to that of an instrument used when carrying out projects or programmes. The author also reflects on the usefulness of external and internal evaluation for contracting parties.

In the second chapter entitled *Unnecessary burden, quality assurance, or opportunity for development? A study of coordinators' attitudes to project*

Beata Ciężka
Independent evaluator since 1995. She has extensive experience in the implementation of research and evaluation projects conducted, among others, for the European Commission, the European Parliament, public administration bodies, non-governmental, as well as private sector organisations. Author of training programmes in evaluation, many of which she conducted herself. She was also a lecturer, teaching postgraduate courses in evaluation at the Institute of Sociology of the University of Warsaw and at the Warsaw School of Economics. Founding member of the Polish Evaluation Society, of which she was President in the years 2004–2010 and is currently a Member of the Management Board.

evaluation, Monika Bartosiewicz-Niziołek reflects on the function and importance of evaluation in projects co-financed by the European Union. The chapter presents the results of a qualitative study which aimed to explore the attitudes of project coordinators towards evaluation. Its results are juxtaposed with data from surveys on similar topics conducted among teachers and employees of non-governmental organisations, public administration bodies and social services. These results present a picture of different ways of using evaluation by the surveyed professional groups and the evolution of their views on the importance and usefulness of evaluation, which is influenced by their accumulated evaluation experience.

In the chapter entitled *Evaluation as a tool supporting the management and development of seniors' organisations on the example of the third age university movement* Jakub Wróblewski describes the development of seniors' organisations thanks to creating an evaluation culture. The author focusses on the operations of entities whose importance and role has significantly increased in recent years due to the ageing of society. Third age universities are presented as institutions facing new developmental and organisational challenges, while evaluation is included in the category of activities that can support the institutionalisation of educational activities aimed at seniors. This is possible thanks to including evaluation in the process of creating development strategies and ongoing organisation management.

Tomasz Kasprzak and Paweł Szymborski also discuss the impact of evaluation on institutional development in the chapter entitled *Implementing an evaluation culture among stakeholders of the Integrated Qualifications System*. The article presents measures supporting the development of evaluation practices and their inclusion in the functioning of institutions responsible for the Integrated Qualifications System (IQS) – especially external quality assurance entities and ministries involved in the creation and implementation of the system. The authors also stress the importance of building a culture of internal evaluation in certifying institutions. In their opinion, high standards of external and internal evaluation form an essential element of the operation of the IQS.

In the chapter entitled *Developmental capacities of an educational institution* Mirosław Warowicki discusses self-evaluation activities in the context of institutional and organisational development. The author presents, among others, the capacity development methodology used for the development of local resources and making it possible to diagnose and develop the potential of an educational institution. This approach also enables the verification of an organisation's priorities and may change their perception by employees and stakeholders. The chapter also presents the capacity self-assessment methodology (CSA) for assessing and developing an institution's resources and internal and external relations.

In the next part of the publication, readers will find articles promoting the use of methodology by evaluators. In the chapter entitled *Evaluation of support for competencies of people on the labour market and for employment promotion. The full picture of effects from the micro and macro perspective* Jacek Pokorski discusses the methodology of estimating the causal effects of support activities developed in recent years which refers to the counterfactual impact evaluation. This type of research attempts to limit the impact of external factors on interventions and to estimate the net effect. The author takes a critical look at the counterfactual methodology and considers the possibilities of applying it in the evaluation of labour market policy instruments. At the same time, he points out the need to apply methodological triangulation combining different participant-oriented (micro perspective) and market-oriented (macro perspective) research approaches.

Jadwiga Fila and Michał Pachocki, who represent the Foundation for the Development of the Education System, show *How to analyse sustainability and long-term effects of transnational learning mobility*. For several years now, the Foundation has been tracing the participants of Erasmus and Erasmus+ mobility projects with a focus on assessing the impact of these programmes on participants' further education and career paths. The results of the study not only broaden knowledge of the effects of the programmes managed by the Foundation, but are also used to promote mobility projects.

Agnieszka Skowrońska presents the Design Thinking methodology in the context of using it to support the development of educational activities and in relation to educational research, which has a similar function. In the chapter entitled *Design Thinking and evaluation – differences, similarities, and examples of their application in education* she compares the objectives and assumptions of evaluation and the concept of Design Thinking. She shows that the two approaches are similar. They aim at supporting problem solving and designing measures that meet the needs of recipients. This article can serve as an inspiration for evaluators interested in including in their research activities based on the Design Thinking concept.

In the penultimate chapter, entitled *Creative evaluators or how to make evaluation more appealing?*, Beata Ciężka presents practical guidelines for the implementation of inclusive evaluation. This article attempts to break up the monopoly of questionnaires in the evaluation of training and other educational activities. It describes concepts and techniques of evaluation focussed on its development-oriented function, which are based on a participatory approach and provide for deeper reflectiveness, dialogue and responsibility for learning outcomes.

In the chapter closing the publication, *Effectiveness of executive coaching in the light of research findings*, Seweryn Krupnik and Ewa Krupnik present inspirations and tools for designing and conducting studies on the effectiveness of executive coaching. Starting from the assumption that evaluation of the outcomes of any coaching process forms its essential element, the authors present a review of eleven studies using various research approaches, from qualitative to quasi-experimental, and examine the usefulness, quality and reliability of the results obtained.

We hope that this publication will serve as a source of inspiration on how to develop professional knowledge and skills, broaden the spectrum of applicable research methods used in the evaluation of educational programmes and activities, and contribute to further development of an evaluation culture.

Beata Ciężka
Agnieszka Rybińska

Quality and usefulness of internal and external evaluation. When do we learn more: commissioning an evaluation or carrying it out ourselves?

DOI: 10.47050/66515246.8-18

Edyta Boratyńska-Karpiej

Abstract

Ever since Poland joined the European Union, evaluation has become very popular here. When granting financial support from EU funds, the European Commission obliges Member States to provide evidence that the money has been spent effectively and that the assumed intervention objectives have been achieved. This chapter examines whether evaluation is a tool for critical analysis or whether it is nothing more than an instrument and the reports drawn up are fit only to end up in a drawer. The article attempts to answer the question whether we can learn and benefit from internal evaluation more when conducting it on our own or when outsourcing it. It also analyses whether evaluation in general provides knowledge about a project, especially about its failures, their causes and the lessons which can be learnt from them.

Introduction

Poland, like any other European country, develops and implements certain public policies, including in the areas of transport, energy and agriculture. The authorities aim to achieve specific results, such as development, competitiveness, stability, and economic and social cohesion. However, the standards for the development and implementation of policies adopted in Poland as compared to those in the most developed Western European countries can raise certain controversy. Public policies can be defined as an area of systemic, organised actions of the state (dictated by objective, research-based knowledge) undertaken to solve key collective issues. Legal regulations, performance indicators, persuasion, policy analysis, institutions (public offices and NGOs), cooperation, debate, and research are used to formulate policies. Evaluation, which is the subject of this chapter, can also be used

Edyta Boratyńska-Karpiej
Political scientist with over 17 years of experience in the transport sector of government administration. Expert in evaluation and analysis at the Centre for EU Transport Projects. Author of articles on the functioning and economics of transport, and co-editor of publications on evaluation and analysis of the impact of transport projects on socio-economic development. Member of the Association of Information Professionals and the Polish Evaluation Society.

KEYWORDS:
INTERNAL EVALUATION,
EXTERNAL EVALUATION,
CONTRACTING PARTY,
NEEDS,
RESOURCES

for this purpose. There is a common view that in Poland most actions are undertaken at random, on the spur of the moment, based on the intuition of decision-makers and on political reasons, especially before elections. Yet public policy should be carried out on the basis of research data and feedback documenting the way the implemented solutions operate, using modern public governance methods, such as taking into account performance indicators, as well as audit, evaluation and consultation of designed activities with stakeholders. These issues are not new – they have been observed in Poland for many years. Fortunately, partially due to the requirements of the European Union, positive phenomena are emerging. Recently, numerous studies and projects which make it possible to explain or reconstruct the situation with regard to various public policies and even to develop analytical tools to improve certain state policies (e.g. transport policy) have been carried out. More and more often the question how to evaluate supplants the question whether to do it at all.

According to a common definition, evaluation is a systematic examination of the worth or characteristics of a specific programme, plan or action from the perspective of accepted criteria in order to improve, develop or better understand it[1]. It is an assessment of the value of an intervention using specific criteria undertaken to determine its effectiveness in relation to objectives, as well as to analyse its impact on specific structural problems. This definition is fairly precise. Doubts arise when the essence of this phenomenon is explained, along with its real significance for the way in which intervention is made, the evolution of political culture and public affairs governance. In other words, here we mean evaluation from the perspective of a modern management model. Unfortunately, Polish publications in the field of evaluation most often refer to analyses of its use in the context of carrying out programmes funded by the EU. Decidedly, evaluation measures should be extended to cover all interventions, regardless of the source of investment financing. This is due to the fact that an evaluator is a broker of knowledge which should be used by everyone at different stages of developing and implementing public policies.

What is evaluation?

The term evaluation (Polish: *ewaluacja,* French: *évaluation*) literally means determining value (*Oxford English Dictionary, 2002*). The tradition of evaluation research (*Ewaluacja w pracy metodą projektu...*, 2008) dates back to the early 20th century. The first evaluation studies were conducted in the 1920s in the United States, and the first attempts at documented evaluation of educational progress were made in the US by Joseph M. Rice (1857–1934), a physician by training who worked as a school superintendent. He was mainly interested in increasing the effectiveness of teaching. Evaluation theory was first formulated in 1949 by Ralph W. Tyler. The current concepts of formative and summative evaluation were introduced by Michael Scriven in 1967 in his book *The Methodology of Evaluation* (Scriven, 1967). In the 1960s, evaluation became an independent discipline of social studies.

A search of the literature aimed at issues related to monitoring and evaluation shows that there is no single, universal definition of evaluation. Many concepts used in this field are ambiguous and depend on adopted methodological assumptions and research concepts. In general, it can be said that monitoring shows "what and how much", whereas evaluation answers the question "why is this

1 See: pl.wikipedia.org/wiki/Ewaluacja [accessed: 9 January 2020].

happening?". This points to the fact that evaluation is knowledge of effects, but also of causes of their occurrence in a given dimension. If we adopt this approach, we can safely say that evaluation plays a major role in the education of decision-makers.

In a knowledge- and information-based economy, the importance of evaluation is great. The results of evaluation studies are used to introduce changes in the programming of social and economic development, as well as to improve the effectiveness of public interventions and organisation management. We learn not only by doing, but also by evaluating. This is especially true when we carry out evaluation by ourselves as part of internal evaluation or a hybrid model (discussed later in this chapter).

In Poland, in the mid-1990s, when evaluation started to be used, it was underestimated. This was due to the distortion of the concept itself, which was presented as "assessment"[2]. Even today, in many institutions, evaluation is associated with audit or control. Yet evaluation is a process in which assessment is only one of the elements. It is important to explain the reasons for the failure of a project (if it did fail), identify intangible effects and make recommendations for the future. Evaluation is also an emerging discipline, a field of knowledge that links applied science with public policy practice.

In the operations of international organisations and public sector institutions, evaluation is defined by its practical purpose, which is to improve the quality and value of public interventions (policies, programmes and projects). As a rule, it is as an element of effective public sector management or good governance (*dobre rządzenie*), which is also a new term in Polish (*White Paper...*, 2001; *Koncepcja good governance...*, 2008; *Governance and Development*, 1992).

The World Bank (*Independent Evaluation Group...*, 2008) defines evaluation as: useful, reliable and transparent research, founded on an objective basis, which assesses the implementation and effects of policies, programmes and projects, contributes to strengthening the accountability of the Bank's activities, and helps to draw conclusions and improve future interventions. This definition underlines the accounting and summative function, as well as indicates the possibility of learning from evaluation.

In documents published by the European Union (*Focus on Results...*, 2000; *Project Cycle...*, 2001; *Evaluating EU activities...*, 2004; *Project Cycle...*, 2004; *Working Document...*, 2006) evaluation is defined in a slightly different way: as a judgement of the value of a public intervention in relation to specific criteria and standards or as an assessment of public actions depending on their results, impact and needs they aim to meet. It should be emphasised that the understanding of evaluation by European institutions has changed in successive periods of budget programming. Before 2000, evaluation focussed on providing support in planning. In the subsequent period (transition period for Poland, i.e. before and after its accession to the EU), there was a tendency to present evaluation as supporting supervision. It was also intended to increase the effectiveness of spending EU funds. In the 2007–2013 financial perspective, evaluation was intended to improve quality, effectiveness and consistency of the funds' assistance in the context of addressing specific EU structural problems. Finally, in the period 2014–2020, the role of evaluation in the management of operational programmes implemented under the cohesion policy increased. This is reflected in formal provisions of legal acts constituting the basis

2 Mistranslation made in the mid-1990s (in a translation undertaken as part of the World Bank grant for the development of the evaluation system in Poland), duplicated and used in several legal instruments (e.g. the Act on the National Development Plan 2004–2006; public tendering procedures organised by ministries, translations of the EU Structural Funds Regulation, etc.). However, in daily practice – also that of government institutions – the original term "evaluation" is used (e.g. evaluation conferences organised by the Polish Agency for Enterprise Development and the Ministry of Regional Development, evaluation guides, annual evaluation plans presented by ministries and carried out by their evaluation units). As a result, terminological confusion persists. There is little hope for putting things in order – in the translations of EU regulations binding in the years 2007–2013, the mistranslation of the term 'evaluation' was copied from the documents applicable in the years 2004–2006 and the term 'assessment' (*ocena*) was again used instead of evaluation.

for implementation of EU funds and guidelines formulated by the European Commission. Growing importance of evaluation in the implementation process also imposes additional obligations on EU Member States, thus making it necessary to reorganise and adapt evaluation systems to new requirements

Can evaluation be a tool?

On the basis of the above considerations, it can be concluded that evaluation is a mechanism for collecting knowledge about whether or not set objectives have been achieved, as well as which factors have proven effective and which have been unreliable (Zybała, 2013). Maybe it's even more important to find out why they failed. A specific feature of evaluation is searching for cause and effect relationships between individual activities and results.

In order for evaluation to provide accurate answers and reliable knowledge, it should be carried out conscientiously. The conducted evaluations should help understand not only whether activities are effective, but also why. Such knowledge cannot be obtained by simply determining and reporting the outcome of an evaluated intervention calculated on the basis of a statistical average. Hence the statement given by the Network of Networks of Impact Evaluation (NONIE):

> *Applying a theory-based approach means that well-designed impact evaluation includes both questions on the process and magnitude of achieved outcomes. The relevance of a policy is highlighted if the study is not only limited to indicating whether or not the intervention assessed has had the intended impact, but also answers the question why this has or has not happened* (Obarymska-Dzierzgwa In: Obarymska--Dzierzgwa & Boratyńska-Karpiej, 2017).

Similarly, the 3ie Working Paper entitled Theory-Based Impact Evaluation: Principles and Practice (White, 2009) and devoted to impact assessment practices concludes that:

> *Research should clearly indicate how an intervention (in the form of specific inputs) has influenced the final outcomes. Research should also test all the assumptions made for a given intervention with regard to the relationship between inputs and outputs (sometimes referred to as programme theory)* (ibidem)

Evaluation effects

In 2014, Tomasz Kupiec carried out an analysis of the use of recommendations from evaluations completed in the years 2007–2012 in three intentionally selected provinces. In total, 44 evaluations of Regional Operational Programmes were carried out, 35 of which contained a total of 440 recommendations (about 12.5 recommendations per report). 40% of them were implemented, but the initiated solutions mainly concerned technical and organisational issues and the improvement of their implementation (weight of implemented recommendations: low or medium). Many of them concerned promotion and information.

The main conclusions of the meta-evaluation were that:

- → Only 20% of evaluations contain recommendations that are fully consistent with research objectives – in the case of 30% of the studies recommendations are completely inconsistent with these objectives;
- → 77% of recommendations are consistent with research conclusions (assessed as reliable);
- → 37% of recommendations are assessed as significant (linked to conclusions, clear and implementable, i.e. having the potential to bring about significant positive changes or solve the problem);
- → 7% of recommendations are of a strategic nature (concern the directions of support, scope of activities, project types and financial allocations);
- → 72% of evaluation studies do not provide answers to all research questions (often answers are simulated – they are just digressions on a given topic);
- → Most often there are no answers to questions concerning the impact of implemented interventions on social and economic development, estimation of impact, and satisfaction of social needs.

According to Tomasz Kupiec, the results of the above meta-evaluation are due to the fact that:

- → Some evaluation studies were conducted at too early a stage of the programme implementation;
- → The scale of evaluated projects was too small for their impact to be judged;
- → Evaluation methods and available budget did not make it possible to measure the impact of the projects.

According to the author, the main cause was too many research questions – 36 per study on average. On the other hand, answers to such a large number of questions are required in additional cross-sectional studies and this is why some questions are very extensive. In addition, in offer selection criteria, contracting parties award bonuses for additional research questions. As a result, the number of objectives and questions may increase several times.

The above issues result in evaluation being perceived in an unfavourable light. When performing evaluation studies or collecting information for analysis, the main problem experts encounter is that evaluation is associated with either audit or control. Additionally, in the opinion of the main recipients, evaluations provide insignificant and not very useful conclusions and recommendations (Kupiec, 2014). This may be due to the fact that evaluation is usually commissioned to people who are familiar with (sometimes even very sophisticated) methodologies but are not experts in the particular field and do not understand the specificity of the sector they are studying. This results in the production of reports that are not fit for purpose. Consequently, they end up in a drawer and are of no use in current operations of the given unit. Such reports, and underlying evaluations, are not an adequate tool for developing and implementing an intervention. They do not offer much support to the learning process of institutions and are rarely used when discussing an intervention.

Who is an evaluator?

Taking into account the above considerations, it is worth reflecting on who a good evaluator is to make sure that evaluation supports the institution in its efforts to improve. In this context,

we can consider employing a researcher and analyst or an expert in a given field. An evaluator who is a researcher is usually able to operationalise a given problem and formulate research questions, select tools, analyse obtained research results and suggest recommendations. An expert evaluator, on the other hand, has extensive knowledge of the research subject matter (e.g. about particular branches of transport or regional development) and thus it is easier for them to channel information collection accordingly. Among external evaluators, people with research competence predominate. Therefore, they emphasise the methodological correctness of studies, sampling, inferences, etc. Unfortunately, it seems research companies do not pay enough attention to their employees' expert knowledge. This is connected with lack of sufficient understanding of the subject matter of the study or to contracting parties approaching experts' competences in an academic way, i.e. verifying their knowledge on the basis of required academic degrees or number of relevant publications. When analysing the requirements of specific terms of reference, a contracting party usually places great emphasis on the competence of the researcher – i.e. the operationalisation of research questions, development of tools, conducting the research and analysing the results. It seems that not enough importance is attached to expert competence and understanding of the subject matter of research. Such situations are not encountered in internal evaluation, on which we will later extrapolate.

Unfortunately, employing too many researchers and too few experts in a research team created for the purposes of a given evaluation can lead to activities which, from the view of the contracting party, are feigned, ineffective or even useless. The opposite situation, i.e. too many experts and too few researchers, is also not favourable. In both cases, the contracting party is glad to have ticked all evaluation boxes, and the contractor is glad to have finished the task and received payment. On the other hand, the problem with evaluations (e.g. in the transport sector) is that the contractor market is very limited, which also adversely affects the quality of services and their prices. This makes it necessary for companies conducting high-grade research to cut costs and reduce quality. The answer to this problem can be to carry out internal evaluations using experts employed in the institution, who will be guaranteed independence by the employer (see below).

Should we commission an evaluation or carry it out ourselves?

Based on information presented above we can say exactly what evaluation is and what we can expect from it. Let us now consider whether it would be better to outsource it, or whether it would be more beneficial if it is carried out by independent experts working within the organisation. Opinions vary. To my mind, it is worthwhile to try at least once to carry out an evaluation study using the resources of a given unit and only then decide whether it would be better to commission evaluation or carry it out on our own.

However, the decision has to be made already at the first stage of evaluation, i.e. when planning it. It is necessary to consider research objectives understood as information needs that are indispensable for a given unit. It is fundamental to determine the proper way of carrying out the evaluation. At the initial stage, we have to choose the subject of our research, i.e. determine what we want to evaluate. We have to know what needs to be analysed in order to obtain the information indispensable for evaluation. The achievement of general objectives can be examined on the basis of the project's overall impact, specific objectives can be analysed on the basis of obtained results, and operational objectives can be investigated on the basis of obtained outputs.

However, the most important thing at the initial stage of evaluation design is to establish for whom the study is conducted and who will carry it out:

- → Will it be an internal evaluation whose results will be used by the organisation? Or maybe the results will be used by a grantor, e.g. the European Commission?
- → Who will perform the evaluation? An internal evaluator (from within the organisation) or an external one (e.g. representing the grantor or selected in a tender process)?

When selecting the type of evaluation, a prior detailed analysis of needed and available resources is necessary: time, personnel (experts) and funds (resources needed for external evaluation).

Evaluation forming part of a programme can be planned as external and/or internal. When carrying out internal evaluation, a great deal of responsibility falls on the project coordinator, who must ensure that a detailed evaluation plan is prepared at the very beginning of project implementation. In the case of this kind of evaluation, we should remember that experts responsible for it should be independent and should not be involved in intervention programming and implementation. Therefore, it would be necessary to appoint them from among project team members who do not perform substantive tasks within the programme. The organisational structure of the project/programme should provide for an employee – a monitoring and evaluation specialist – responsible for all activities related to this area of project/programme management. On the other hand, however, we must remember that the evaluator themselves, without active cooperation with project managers, are not able to develop evaluation assumptions. It is necessary to hold working meetings in order to collect answers to relevant questions concerning evaluation study methodology.

External evaluation also requires time and commitment on the part of project managers. In addition, provision should be made for the selection of an entity responsible for carrying out the evaluation. Contracting an evaluation usually requires taking into account procedures resulting from the public procurement law and compliance with the rules of competition, so that expenditure incurred in a project is competitive and effective.

On the basis of my professional experience, I can say that internal evaluation has many advantages. First of all, it meets current needs of the unit, and does not merely implement the commissioned plan. Internal experts can start the research process when a unit manager decides they need to obtain information. We do not have to waste time on procedures, both internal and external (under applicable public procurement law). Internal evaluators are guided by the objective of solving a current problem of a given unit. An unquestionable advantage of internal evaluation is that it can provide information quickly, when the research problem is still topical. The obtained evaluation results, verified by expert opinions, can be available in a relatively short time, thanks to which the developed recommendations provide proposals for improving project implementation or management processes[3]. Outsourced evaluation usually provides information too late for it to be used. This is so because it is necessary to apply internal procedures and all formalities relating to the award of a public contract. All this takes so long that the subsequent examination may no longer be useful.

It should be mentioned that, as time is of the essence, the internal evaluation process should not be too lengthy or complex. Therefore, evaluation carried out by the organisation itself, due to limited human resources, should not be too complicated as far as methodology is concerned. It is also worth emphasising the great advantage of internal evaluation, i.e. that it allows for saving financial resources.

3 Statement by Adrian Mazur, Director of the Department of Transport Strategy at the Ministry of Infrastructure, during a conference organised by the Centre for EU Transport Projects on 12 December 2018, www.youtube.com/watch?v=0lyBkUKKbgg [accessed: 9 January 2020].

Internal evaluation should be carried out as part of the duties – as far as time allows – as well as the skills and experience of staff employed by the institution.

It is also important that information obtained for the purpose of the study is collected using methods that are both accessible to researchers and understandable to the recipient. This makes it possible to carry out evaluation that is useful for the institution and whose results are relevant to needs as well as interpretable by decision-makers.

In the case of more complex research problems, it is worth using a hybrid model of evaluation. In this case evaluation is carried out by internal experts supported by external specialists. This way, internal evaluation responding to more complicated needs can be conducted and can make it possible to solve more complex research problems. On the other hand, internal evaluation can also be chosen in an emergency situation. It is still a cheaper method to collect data and interpret it for decision-making purposes. Both internal and hybrid evaluations should be carried out with respect towards the people who are being interviewed. The needs, expectations and opinions of all stakeholders should be taken into account, even if they contradict one another. What is more, it is easier for internal experts with knowledge of the subject matter of evaluation to verify the authenticity of the positions presented and look for sources to confirm them. On the other hand, internal evaluation can be hampered by information providers' limited confidence in experts working in the same institution. People who are interviewed or surveyed may find it harder to believe that critical opinions will not affect relations within the institution and that those providing information will remain anonymous. At the same time, it is worth remembering that it will be easier for internal evaluators to work out compromise solutions.

Lessons learnt from evaluation

Evaluation is not conducted solely for the sake of evaluation. It can, therefore, be concluded that it is the starting point of activities aimed at improving the functioning of an institution. It educates or provides a basis for educating employees and decision-makers. It is essential that the results of evaluation and its recommendations are disseminated as widely as possible within the institution, that they are discussed and interpreted, conclusions are drawn and, consequently, further action is planned on their basis.

It is worth involving in the process of improving the institution/organisation as broad a range of stakeholders as possible, including managers and other staff. These groups should take part in the process of discussing the results of evaluation and making recommendations for further actions. This participatory approach serves the purpose of building consensus on the introduction of change, as well as responsibility for this change.

In the publication of the Polish Agency for Enterprise Development entitled *(R)ewaluacja 2. Wiedza w działaniu*, edited by Agnieszka Haber and Karol Olejniczak, evaluators were described as knowledge brokers. This is an appropriate approach if we consider the role of internal evaluators. *Thinking in terms of 'knowledge brokerage' helps to better describe, organise and strengthen the role of evaluation units in providing reliable, useful knowledge to its users – decision-makers, designers of interventions, and public managers* (Haber & Olejniczak, 2014). It would be worth disseminating this approach. The aim of knowledge brokers should be to help decision-makers or professionals close to them to acquire and use sound knowledge to better design and lead public interventions.

> *In other words, the better the knowledge broker's actions, the more likely it is that decision-makers will properly direct the intervention and that it will serve citizens well in meeting their needs. The assumption about how a knowledge broker would influence the reality of public action can be written down in the form of a logical sequence – the theory of change. It is a tool commonly used to analyse and evaluate public interventions. After all, the activity of knowledge brokers is also an intervention, as it aims at changing the current state of affairs* (ibidem).

An appropriately conducted (according to the needs of the unit) evaluation process gives the opportunity to collect a large amount of data which, when properly interpreted, provide knowledge that the unit needs. For this to happen, however, evaluators should originate from the given unit. Only then will they become brokers of knowledge which is relevant to the organisation. And only with this assumption will they be able to disseminate this knowledge in a reliable way, reaching the right people. An undeniable advantage of such an approach is the curtailment of costs, which in the era of expenditure reduction (due to Brexit, among others), is of great importance.

Bibliography

Evaluating EU activities – A practical guide for the Commission services,(2004). Brussels: European Commission Directorate General for Budget (2001), Luxembourg: Office for Official Publications of the European Communities.

Ewaluacja w pracy metodą projektu. Pakiet Edukacyjny Pozaformalnej Akademii Jakości Projektu Część 4 (2008). Warsaw: Foundation for the Development of the Education System.

Fernández-Ramírez, B. & Rebolloso, E. (2006). *Evaluation in Spain: Concepts, Contexts, and Networks.* bit.ly/2s7ITEv [accessed: 9 January 2020].

Focus on Results: Strengthening Evaluation of Commission Activities (2000). Brussels: European Commission. SEC2000/1051.

Governance and Development (1992). Washington: World Bank.

Haber, A. & Olejniczak, K. (2014). *(R)ewaluacja 2. Wiedza w działaniu.* Warsaw: Polska Agencja Rozwoju Przedsiębiorczości.

Hausner, J. (2007). Polityka a polityka publiczna. In: J. Hausner (ed.), *Zarządzanie publiczne.* Warsaw: Wydawnictwo Naukowe Scholar.

Independent Evaluation Group – Overview (2008). Washington: World Bank.

Koncepcja good governance – refleksje do dyskusji (2008). Warsaw: Ministerstwo Rozwoju Regionalnego.

Kupiec, T. (2014). Użyteczność ewaluacji jako narzędzia zarządzania regionalnymi programami operacyjnymi. *Studia Regionalne i Lokalne*, 2(56). bit.ly/37X48rN [accessed: 9 January 2020].

Obarymska-Dzierzgwa, J. & Boratyńska-Karpiej, E. (2017). *Ewaluacja wpływu programów transportowych – teoria i praktyka*. Warsaw: Centrum Unijnych Projektów Transportowych.

Oxford English Dictionary (2002). Oxford: Oxford University Press.

Petraś, Z.J. (1990). *Teoria decyzji politycznych*. Lublin: Uniwersytet Marii Curie-Skłodowskiej.

Project Cycle Management Manual (2001). Brussels: European Comission EuropeAid Office for Cooperation.

Project Cycle Management guidelines(2004). Vol. 1. Brussels: European Commission EuropeAid Office for Cooperation, Luxembourg: Office for Official Publications of the European Communities. ec.europa.eu/europeaid/evaluation/methodology/index_en.htm [accessed: 9 January 2020].

Scriven, M. (1967). *The Methodology of Evaluation*. Chicago: Rand McNally.

Sekutowicz, K. (2006), Monitorowanie i ewaluacja. In: D. Głażewska (ed.), *Zarządzanie, planowanie i ocena własnych działań*. Warsaw: Akademia Rozwoju Filantropii w Polsce.

Suwaj, P. & Szczepankowski, R. (2009). Wokół pojęcia polityki i polityki publicznej. In: B. Kudrycka et al. (eds.), *Nauka administracji*. Warsaw: Wolters Kluwer.

Svara, J.H. (2006). Introduction: Politicians and Administrators in the Political Process – A Review of Themes and Issues in the Literature, *International Journal of Public Administration*.

Szarfenberg, R. (2016a). *Polityka publiczna – krótkie wprowadzenie*. Warsaw: Instytut Polityki Społecznej Uniwersytetu Warszawskiego.

Szarfenberg, R. (2016b). Polityka publiczna – zagadnienia i nurty teoretyczne, *Studia z Polityki Publicznej, 1*(9).

Szarfenberg, R. (2017). *Teoria i praktyka polityki społecznej*. Warsaw: Instytut Polityki Społecznej Uniwersytetu Warszawskiego.

Sześciło, D. et al. (2014). *Administracja i zarządzanie publiczne. Nauka o współczesnej administracji*. Warsaw: Stowarzyszenie Absolwentów Wydziału Prawa i Administracji Uniwersytetu Warszawskiego.

White, H. (2009). *Theory-Based Impact Evaluation: principles and Practice. 3ie Working Paper 3*. 3ie – International Initiative for Impact Evaluation.

White Paper of the European Union on Good Governance (2001). Brussels: European Commission.

Working Document No. 2 – Indicative Guidelines on Evaluation Methods: Monitoring and Evaluation Indicators, DG-REGIO, The New Programming Period 2007–2013 Working Documents (2006). Brussels: European Commission, Luxembourg: Office for Official Publications of the European Communities. ec.europa.eu/regional_policy/sources/docoffic/working/sf2000_en.htm [accessed: 9 January 2020].

Wytyczne w zakresie ewaluacji polityki spójności na lata 2014–2020 (2015). Warsaw: Ministerstwo Infrastruktury i Rozwoju.

Zybała, A. (2012). *O lepszą jakość polityk publicznych*. Warsaw: Institute of Public Affairs.

Zybała, A. (2013). Ewaluacja i jej "uczenie się". Wiedza w działaniach publicznych, *Animacja Życia Publicznego, Zeszyty Centrum Badań Społeczności i Polityk Lokalnych, 4*(11).

Unnecessary burden, quality assurance, or opportunity for development? A study of coordinators' attitudes towards project evaluation

DOI: 10.47050/66515246.20-41

Monika Bartosiewicz-Niziołek

Abstract

This chapter presents the findings of qualitative research which aimed to explore the attitudes of Polish coordinators of EU-funded projects towards evaluation. The data is presented with reference to the findings of similar research conducted among civil servants, teachers and social services staff, as well as NGO workers. Additionally, this chapter may contribute to an analysis of different ways the above professional groups use evaluations as they gain an increasing understanding of them through "learning by doing". It may also serve as a basis for reflection on the need to educate adults about evaluation, as those two factors influence their attitude towards it.

Introduction

Evaluation practices first appeared in Poland in the mid-1990s, mainly in academic circles and among non-governmental organisations, in connection with the implementation of projects financed by foreign grants. It became widespread only after Poland's accession to the European Union in 2004, especially during the 2007–2013 programming period, and as a result of the 2009 inclusion of evaluation activities in the education system under the pedagogical supervision regime. Although today evaluation is inherent in any public administration activity related to the implementation of European funds and the functioning of educational institutions, examination of **attitudes towards evaluation** is still rare. The occupational groups which have been subject of this type of research (usually within a single study) include public institution staff, primarily civil servants, as well as those working in the education sector, the social services system, and for NGOs.

Monika Bartosiewicz-Niziołek
Graduate of doctoral studies at the School for Social Research of the Institute of Philosophy and Sociology of the Polish Academy of Sciences, and has completed Postgraduate Studies in Project Evaluation. She has evaluated around 100 projects and programmes and is the author of publications and reviews on evaluation. She is Vice President of the Polish Evaluation Society, as well as author and host of several dozen training sessions on evaluation. She is a member of the Evaluation Steering Group of the Operational Programme Knowledge Education Development, a certified reviewee and a reviewer of the European Evaluation Society, and Member of the Advisory Board of the Network for Evaluation Societies in Europe.

KEYWORDS:
EVALUATION,
ATTITUDE TOWARDS EVALUATION,
ADULT LEARNING

Attitudes towards evaluation manifested by representatives of various public sector entities, mainly civil servants[1], were examined as part of broader quantitative research in 2010 (*Katalog Polskiego Towarzystwa Ewaluacyjnego*, 2011). The results of that survey show that attitudes towards evaluation are largely a derivative of its **usefulness**. This is indicated by the fact that the respondents said one of the most important characteristics of evaluators was their ability to translate research results into practical solutions. This ability is a key competency, especially in utilisation-focussed evaluation, which is based on participation and active involvement of its target users in the entire evaluation process (Patton, 2008; Patton, 2013).

The attitude of contracting parties towards evaluation companies and the research they offer is, unfortunately, not very encouraging. The vast majority of respondents in the survey agreed with the statement that many contractors are only interested in "easy money", do not care about the quality of research and resort to various legal loopholes. Less than one in five contracting parties said they **trusted** external evaluators and the quality of their services. Although the majority of respondents declared that evaluation brings measurable benefits to public administration, less than half considered evaluation studies to be highly useful. This may result, among others, from the fact that those in charge of institutions which order such studies, and are responsible for implementing their recommendations, show moderate interest in their findings. One of the reasons for this state of affairs could be that public institutions often have evaluations conducted solely to meet external requirements and be compliant with applicable laws, and not to obtain information that would help them improve their operations. At the same time, the lack of trust in evaluators and the quality of their research may be caused not only by the negative experiences of the contracting parties, but also by the general low level of social trust in Poland, which may constitute a significant **barrier to the development** of an evaluation culture.

However, while staff in central and regional government generally seem to have a high level of knowledge about evaluation (due to the implementation of EU programmes by their home institutions), these competencies are very low among local government officials and in smaller urban centres. Low awareness of evaluation practices at the **local level** translates into their limited popularity and underestimation – due not only to the lack of education in this area among staff, but also to financial deficits and organisational difficulties. Another barrier to the development of evaluation at the local government level is a very low degree of its **institutionalisation**. The vast majority of units have no procedures or ordinances regarding evaluation, although how advanced individual communes are in this respect **varies** (Grzywa, 2008). Nevertheless, even officials who are familiar with evaluations frequently equate them with a way to **monitor** and **account for** the results achieved, while attaching less importance to their usefulness for the beneficiaries of activities or to their relevance to the needs of the local community (Worek, 2013).

Education staff are another professional group which, similarly to civil servants, is required to conduct evaluations. Research carried out in this population lead to the conclusion that the attitudes of teachers and evaluation inspectors[2] were **diverse** and **ambivalent**. Some inspectors, who previously often held the position of "super headmasters", after becoming external evaluators felt they had lost influence over the school and had a sense of lesser responsibility for its functioning. Others, however, felt more helpful in the new situation and were seen as "good spirits" by the teaching staff (Kołodziejczyk

1 The study also included representatives of school education institutions, Higher Education Institutions and employment services. The sample comprised 106 respondents and was not representative.

2 These are persons responsible for conducting external evaluation in schools and educational institutions.

& Kołodziejczyk, 2015). In addition, teachers seemed to slip into certain **dissociation**, separating the notion of evaluation from its practical application at school. The very idea of evaluation was often seen as valuable and useful, yet the evaluation activities that the respondents encountered at work (and which they themselves undertook) triggered feelings such as aversion and fear due to various difficulties, as well as a sense of dissatisfaction, helplessness and submission (Sury, 2017).

Teachers' attitudes are also worth considering from a **dynamic** perspective – by analysing the way in which they can change under the influence of newly acquired competencies and experience, following the trajectory of development of evaluation practices in various societies (Guba & Lincoln, 1989; Korporowicz, 2011; Mizerek, 2016; Jaskuła, 2018). The initial response of the teaching staff to the introduction of evaluation in the supervision system of educational institutions was generally filled with fear and negative feelings, as they associated evaluation activities with control and judgement. However, after completing the entire evaluation cycle, many – although only **after some time**, which needs to be emphasised – did see the benefits of evaluation (Wasilewska et al., 2014; Walczak, 2015). One of the reasons why the change in attitude towards evaluation practices is gradual and progressive in nature is that they are part of a **learning process** which happens mainly in practice ("learning by doing"). It should be noted that the attitude to evaluation is significantly influenced by **how its findings are used** (Wasilewska et al., 2014), although this relationship works both ways: implementation of recommendations legitimises evaluation activities and contributes to their recognition as useful.

In principle, however, in educational settings, like among civil servants, evaluation practices are often perceived in terms of control, criticism and judgement, instead of support (Sury, 2016). Interestingly, the surveyed teachers were **more critical** of evaluations than the management, because they were more burdened with their conduct. In other words, the headmaster's positive attitude towards evaluation did not necessarily translate into similar attitude among teachers, who were also, in most cases, responsible for addressing the findings of evaluation activities, i.e. for implementing the recommendations. Evaluation competencies also play an important role in shaping attitudes, because the teaching staff's approach to evaluation depends largely on how much they **know** about it. Individuals who participated in relevant training had a more positive attitude towards evaluation practices compared to those who relied solely on the opinions of headmasters and other teachers (Wasilewska et al., 2014).

The issue of trust mentioned earlier in relation to civil servants seems to be equally important in the case of social services workers, as highlighted by a 2010 study conducted in the region of Małopolska by the Regional Centre for Social Policy in Krakow. In this case, however, the respondents pointed to a lack of trust not so much in evaluators or the quality of evaluation studies, but in the way their findings were used. It is worth noting that these fears were mainly found among the "rank and file" employees of the surveyed entities, rather than those in charge of them. Social services workers were also characterised by a more **complex** attitude to evaluation in comparison with that exhibited by the management. In addition to listing the advantages of evaluation, social services workers would also expressed certain concerns and fears which were not shared by their superiors, such as being burdened with additional tasks or suffering the consequences of instrumental use of evaluation findings. This **ambivalent** attitude to evaluation stemmed from the belief that it is a tool that may be used in a variety of ways, depending on the circumstances. That is why respondents perceived evaluation activities as a source of the risk of demotivation of social services workers, as well as of an exacerbation of their relations with clients. In addition, they emphasised that competence-related barriers and deficient good practices in the field of evaluation, as well as staffing and financial shortages, limit not

only the possibility of conducting evaluation studies, but also that of **implementing** the resulting recommendations.

Although the vast majority of managerial staff felt that evaluation of social services activities is useful, those who saw evaluation practices as a way of **controlling** and **judging** their institution showed a negative attitude towards evaluations and underestimated their benefits. It is also worth noting that some of the fears of social services workers were due to a lack of evaluation competencies, and the conviction that evaluation plays only an instrumental role was the result of individual beliefs and experiences. On the other hand, the more knowledge the respondents had about evaluation research, the more aware they were of the need to carry out evaluations and the stronger they believed in their usefulness for social services (*Model ewaluacji w pomocy społecznej*, 2010; Hryniewicka, 2008).

A similar state of affairs can be observed in non-governmental organisations. The majority of NGOs not only do not evaluate, but also neglect monitoring, meaning that they do not systematically assess the effectiveness of their activities or analyse their results. Some organisations equate evaluation with **satisfaction** surveys (constructed in a rather intuitive and schematic way) carried out among their beneficiaries. The main reasons why NGOs fail to adopt professional evaluation practices are similar to those indicated in other professional groups: lack of the need to do so related to a lack of awareness of the benefits of evaluation, negative connotations with it, such as judgement and control, as well as competence deficits (Stowarzyszenie Klon/Jawor, 2012).

The findings of the selected studies on the attitudes of different circles towards evaluation discussed above paint a fairly coherent, if internally diverse, picture, which is a derivative of the current stage of development of the **evaluation culture** in Poland. This background is precisely why evaluation is perceived through the prism of the familiar mechanisms and tools of control or judgement, deeply rooted in our society, which trigger negative emotions and discourage from adopting good evaluation practices. And when evaluation is seen in such light, it tends to be dismissed as an unnecessary burden, not only because of the shortage of resources required to conduct it, but also because of the perceived lack of benefits. Nevertheless, any positive experiences gained in the course of commissioning, conducting and participating in evaluation pay off in the form of a change in attitude towards it. Those circumstances create a mosaic of **complex** and **heterogeneous** attitudes, which may be the result both of individual experiences gained in the evaluation process and of the attitudes of superiors towards using evaluation findings – voices that evaluations are pointless and, in effect, disappointment in the whole practice are often the result of **failure** to implement the recommendations. It is worth noting that evaluators are not only to a high degree responsible for the course and atmosphere of this process, but can also, through an interactive and dialogical approach to evaluation, contribute to increasing its usefulness (Korporowicz, 2008).

Ignoring or marginalising the role of evaluation in the management process significantly hinders, or even prevents, rational decision-making aimed at improving implemented projects. The **organisational culture** of the entity which commissions evaluation activities or conducts them using its in-house resources is of considerable importance here. The development of an evaluation culture depends not only on the degree of its institutionalisation in a given sector, but also on the readiness of the organisation to self-reflect and learn from its mistakes, as well as on its desire to self-improve and its willingness to be open to changes. Another important aspect of evaluation practice is **cooperation** – both with the external evaluator and within the institution. If evaluation activity

is restricted to a specialised unit, while other organisational units[3] do not participate in e.g. conceptualisation and data collection processes, they are later not interested in the findings or perceive them as criticism or a personal attack (*Ocena zdolności...*, 2010). Conducting non-participatory evaluations results in lack of involvement of important actors in, or even their alienation from, this process. In effect, they lose the motivation to develop and change, and thus end up undermining the importance of evaluation as a whole, which, in turn, may weaken the system of implementing recommendations.

Furthermore, practice shows that misconceiving evaluation, for example **reducing** it to assessment of the effectiveness of reaching certain indicators (that is equating it with monitoring), may cause various distortions. It is also worth noting that in Poland there are still areas where evaluation practices are absent for various reasons (e.g. at the local level, in social services). This increases the risk of introducing inadequate, inefficient or even feigned actions (Trawkowska, 2007; Worek, 2013). A similar effect may also be brought about by the suspension of EU project evaluation[4], which, according to decision-makers, was supposed to be a remedy for its low quality, yet we cannot help but exclaim: *Don't throw the baby out with the bathwater!* Concern for a high level of evaluation should be manifested in providing opportunities to acquire and develop relevant competencies, and in continuous work both on raising awareness of the importance and usefulness of evaluation and on improving the system of selecting contractors and implementing recommendations.

Study of the attitudes of EU project coordinators – methodological issues

The question of attitudes towards evaluation was again raised by the Polish Evaluation Society (PTE) in 2018 as part of a qualitative study based on in-depth individual interviews (IDI). In the period from February to April 2018, PTE members[5] conducted 14 interviews with coordinators of EU-funded projects. The sample included persons who coordinated and evaluated projects in the years 2007–2013, implemented mainly under the Operational Programme Human Capital (OP HC)[6]. IDI respondents represented all three sectors, that is central and regional public institutions, non-governmental organisations, and enterprises (mainly small and medium-sized). Most of the projects implemented by them had an **educational** or **social** profile, with single cases of investment projects and initiatives related to the development of entrepreneurship (it is worth noting that many were innovative or pilot in nature). The value of the projects varied widely and ranged from several hundred thousand to several hundred million PLN[7].

3 For example, those responsible for the implementation of projects being evaluated, as well as bodies supervising these processes, such as the Evaluation Steering Group or the Monitoring Committee.

4 According to the guidelines of the Minister of Infrastructure and Development concerning cohesion policy in the years 2014–2020, project evaluation is only possible in justified cases as a departure from the principle of evaluation at the level of a particular Programme, Priority or Action, rather than the individual project level.

5 The study was conducted by PTE members: Beata Ciężka, Sławomir Kozieł, Małgorzata Niemkiewicz, Jakub Wróblewski, Agnieszka Szczurek, Agnieszka Śnieżek, and the author of this chapter.

6 In addition to OP HC, they included the following EU initiatives: Regional Operational Programme, the Operational Programme Innovative Economy (OP IE), the Operational Programme Infrastructure and Environment (OP I&E), and Erasmus.

7 For example, in the case of infrastructure projects.

Evaluation experiences

IDI respondents had very **diverse experiences** in the field of evaluation. Among the interviewees were persons with no experience or knowledge about it who carried out evaluation activities only because they were **obliged** to do so: *I commissioned the evaluation because it was a requirement*[8]. On the other hand, some coordinators had a background in social sciences and/or had experience in conducting or commissioning evaluations, often considering it an **inherent part of a project**: *I can't imagine a project without evaluation; I have been involved in writing projects for a long time, that's why I think of evaluation as a basic tool that is inherent to EU projects.*

Some respondents had previously encountered evaluation in **areas other** than project implementation, e.g. in relation to university activities and educational programmes (e.g. SMART) or training. Some obtained information on evaluation independently, sourcing it from online publications, while others gained this knowledge during training sessions and courses.

The sample also included people who learned to evaluate "**in action**", i.e. during the implementation of EU projects. If the coordinators' experience in this area happened to be **positive**, they tended to show a **positive attitude** towards evaluation practices.

Evaluation resources

Although the majority of respondents believed that the financial resources allocated to evaluation and the duration of evaluation were **sufficient**, the scope of evaluation was generally quite limited, and they had to commit their own resources: *0.5% of the budget was allocated to evaluation. That amount was sufficient because it was basic research, with significant use of our own resources.* Some coordinators, however, emphasised that the small financial support and very short duration of evaluation could have had a **negative impact** on the usefulness of its findings, and, furthermore, that those two factors contributed to **limiting the scope of evaluation**: *You could say that there's never enough funds, and the final evaluation plan is a matter of negotiations with the contractor; The evaluation was conducted to a fairly limited degree (e.g. only surveys), because the funds for evaluation are usually small; The problem is that the budgets of some projects don't foresee any money for evaluation and it ends up being done voluntarily by team members. The time allocated for evaluation is too short. Evaluations are carried out for the sake of audits, so only the bare minimum is done.*

Underlying assumptions and conduct of evaluation

The respondents performed both internal and external evaluation, as well as, although much less often, self-evaluation. In relation to **innovative projects**, internal evaluation was often conducted in-house and combined with monitoring, while external evaluation was carried out by experienced evaluators and research or consulting firms. In the case of **internal evaluation**, emphasis was placed on the effectiveness of completed activities and opportunities for improvement, as well as on the current and future needs of the beneficiaries and their satisfaction level. Information was collected

8 Italics indicate quotes from respondents.

through surveys, individual and group interviews, observations, as well as analysis of monitoring data, i.e. values of relevant indicators.

External evaluation, on the other hand, largely involved an analysis of the usability and potential uses (sustainability) of the findings[9]. This type of evaluation aimed to summarise the effects of a completed initiative, but also to assess the possibilities of its continuation. The evaluation process itself included (as in the case of internal evaluations) documentation analysis, interviews (individual and group), surveys and observations, however it relied to a greater degree on the opinions of experts and panels involving the beneficiaries themselves[10]. As part of external evaluations, case studies were also sporadically carried out.

For **systemic** projects, similarly to innovative ones, evaluation was the norm. Where a professional evaluator was not hired, evaluation was carried out by the project coordinator: *Every project was evaluated. It was often included in the budget, and if not, the coordinator would see the need for it and conduct it in the form of supervision, through conversations and meetings with persons involved in project activities, and with beneficiaries. Local coordinators also talked to beneficiaries. External evaluations were also commissioned* [...]. *There were no projects without any form of evaluation, information was always collected, although in different ways. Most coordinators appreciated project feedback*.

In systemic projects, **ongoing** and **ex-post** (summative) **evaluation** were most often carried out. The former focussed on questions, such as how work is organised, how project activities are implemented and how to increase their efficiency, as well as on obtaining feedback from project beneficiaries: *The goal was to obtain information on the activities carried out – how they are conducted, what problems are encountered by both those who run the project and those who benefit from it, how they can be helped, what cooperation in the group is like, what helps the group and what disturbs it. The focus depended on what information was needed in the particular case and on the nature of cooperation with people; The idea was to achieve better results, a greater efficiency of activities, to learn how to support colleagues so that they are more effective in their work*.

Where the initiative was aimed at **developing competencies**, the evaluation also served to measure their growth, which would be one of the project's outcomes. It was emphasised that acquisition of competencies and change of attitudes is a lengthy process. Therefore, it is difficult to research and demonstrate such effects immediately after the project is completed, although grant cycles often force such practices. Ex-post evaluation was carried out at the end of the project, so the time devoted to it was often **too short**.

According to most respondents, **cooperation with the evaluators** was good and no problems were encountered in this respect: *I can't think of any problems I've had with evaluation. We were lucky with our specialists. Or maybe it was because of our carefully prepared contractor selection procedures.* In one case, the contractor quit, which gave rise to the necessity of quickly replacing them. Good relations with evaluation contractors were often due to **long-term**, as well as **previous** cooperation. Many emphasised the importance of the evaluators' knowledge and experience in a particular **evaluation area** or specific nature of the project, which translated into the quality of their evaluation. Some coordinators attributed the quality of external evaluations to their contractor **selection procedure**. When **price** was

9 For example, in relation to the final product, which the innovative project aimed to develop and test.

10 That is users of the final product developed in the course of the project.

the leading selection criterion, they would end up with unreliable companies which would not meet the agreed deadlines and provided poor reports: *We have had poor external evaluations, especially when price was the only selection criterion. The evaluation was simply badly done. Those companies failed to meet their obligations, were late, and the evaluation was sloppy. One time, reports kept being rejected for three months because they were of such poor quality. It was the effect of Public Procurement Law – it was impossible to choose the companies we trusted. Personal contact, meetings, conversations, explanations, openness on the part of the company and its willingness to work together would all contribute to successful cooperation, but if the company persisted and kept doing things wrong, then often nothing could be done.*

For those who had not previously carried out evaluations and did not know what they involved, cooperation with the evaluator was a **burden**, because they had to provide the necessary documents and contact details of project participants, and then analyse the content of the report to add their comments: There were no difficulties, only a certain burden, because we had to provide the evaluator with aggregated data and then read the report. More experienced coordinators were aware of the fact that their **engagement** in the evaluation translates into its **quality**: *Some contracting parties think that they'll sign a contract and that's it – take care of it and don't bother me. But this is the first step towards failure. The contracting party must be involved at every stage of the evaluation, be vocal about what it needs and what has been done badly, and monitor the quality of evaluation.*

Assessment of the usefulness of completed evaluations

Evaluation findings were used by the coordinators in various ways. In the case of innovative projects, they served to i**mprove the quality** of the final product and to meet the **formal requirement**: *Internal evaluation allowed us to improve project deliverables on an ongoing basis, especially since both projects were challenging and did not go as smoothly as we expected. External evaluation is more of a formal requirement, necessary for validation, it was difficult to use its findings to improve the deliverables; At first, I did not understand what evaluation was and why it was done. I came to appreciate its importance during project implementation. Both those of us who implemented the project and those in charge from the IB and the MA* [Intermediate Body and Managing Authority – editor's note] *were more careful and accurate. Now this is missing* [...]. *Where there was no such requirement, we would not do an evaluation. We just didn't think about it when writing projects.*

Evaluations of non-innovative projects were also aimed at **improving** the activities conducted and **avoiding mistakes** in subsequent projects, as well as **disseminating best practices**: *The findings of evaluation mainly contributed to drawing conclusions for the future and adjusting the support provided; The main focus was on those aspects which performed poorly so they could be changed or improved. But we also found it useful to know which areas were successful, so that we knew what worked and was worth disseminating.* Evaluation served not only as "quality control", but also helped to **implement** the project, **manage** it, and develop **corrective measures** and **directions for further work**: *I have always been very serious about evaluation findings. They gave me an objective insight into the quality, relevance and effectiveness of our activities. They were a trigger to introduce changes and improve forms of assistance and project management. The findings of intermediate reports and recommendations helped improve the project itself. On the other hand, the findings of final reports provided information on how to plan subsequent activities and projects. The findings of evaluations* [...]

helped us implement the project, were discussed by the project team. They were also the basis for implementing changes.

Evaluation findings were used to assess the project in terms of **various criteria**, e.g. relevance, effectiveness, efficiency or utility: *The purpose of the evaluation was to determine the degree of achievement of project goals and effects; The evaluation was used to answer questions about the quality of project implementation and appropriateness of funds use; The evaluation was used to analyse the degree to which the expected results had been achieved, and to determine the effects of the methods and activities used.* The respondents emphasised that evaluation helped them get **feedback** from the project beneficiaries, and also supported **team collaboration**: *The reason we did the evaluation was actually to implement its recommendations. Comments were often analysed on a regular basis within the project team and decisions were made. Immediately after each meeting and whenever help was needed, action was taken. We contemplated what could and couldn't be done.*

Factors which translated into good use of evaluation findings included not only successful cooperation with the evaluation contractor and its high quality, but also the way the whole process was conducted, especially the **participatory approach**: *The evaluation was carried out very professionally and very thoroughly. As project implementers, we were involved in its every stage. Because of this, we acquired a lot of knowledge about evaluations – their rationale and advantages. We have benefited from it.*

Evaluation findings were most useful to the **project team**, followed by **project stakeholders** and the **management of the implementing entity**: *The findings of an evaluation are usually analysed by the project team or by teams of employees who work in the relevant area; The evaluation was most useful to the team that implemented the project. Evaluation contractors also delivered a presentation to the stakeholders* [...]. *It was a form of feedback for them. To a lesser extent the findings of the evaluation were used by the management, because there were frequent changes of managers. If managers change frequently, they're not always interested in evaluation, they're not always aware of it. In such a case, evaluation reports are not discussed or shown any interest in.* In addition to frequent changes of management, another **difficulty** with using evaluation was **not taking account of its findings** and their **insufficient dissemination**: *The problem was insufficient use of evaluation findings. During project implementation, the evaluation was done very well, but its findings remained within the institution, and no conclusions were drawn from it.*

Those aware of the **benefits** of evaluation emphasised that it should be conducted **on the initiative of** every coordinator who **cares about project quality**: *Any project implementer who is interested in whether the project is going well will perform an evaluation or collect feedback anyway, for example through meetings with employees, carers and beneficiaries. They are interested in how the project is going, what's working, what isn't and what can be improved, because it's a lesson for the future. If you don't feel the need for it, evaluation will never help you.*

Assessment of the **usefulness of evaluation** depended on the respondents' **previous experiences**. Some respondents who had experience with evaluation emphasised the greater value of **internal evaluations**: *The findings of internal evaluation were more useful, they were more aligned with the needs than those of external evaluation; The evaluator kept us updated so that we could use the findings to improve the deliverable as quickly as possible; There should only be ongoing evaluation,*

conducted by an internal evaluator or coordinator; We have received reports commissioned by the IB/MA at the programme level, but they were terrible. Low quality, no knowledge of the evaluated field, the companies did not understand what they were evaluating. Often, the draft evaluation was so poor that it should not have been approved, but no one said "stop" at any point. Some respondents indicated that external evaluations were characterised by greater superficiality.

Other coordinators valued **external evaluations** more, mainly due to their **objectivity** resulting from lack of involvement in project implementation, as well as **expert knowledge** and the possibility of confrontation with a **broader context**: *Evaluations should be outsourced and conducted by a third party. When project implementers were responsible for conducting the evaluation, we ended up wanting to show the project in the best light. Some criticisms we heard from the evaluator were later not included in the report because we had agreed not to mention them.*

Among the respondents were those who saw the usefulness of **both types** of evaluation, saying that their value depended on cooperation with a particular contractor: *Both forms of evaluation were performed – external and internal. The evaluator was chosen through official tendering procedures, while internal evaluation was often carried out by coordinators themselves, as long as they knew what they were doing, or with the support of someone from another project. I wouldn't say that one of those types of evaluation was better, it depended on the cooperation with a particular evaluator. External evaluation has its advantages, because it's done by someone who's not emotionally involved in the project.*

As regards innovative projects that aimed to test new solutions, the opinions of the respondents were **varied**. Some coordinators pointed to the high usability of evaluations of such projects and appreciated the value of the information obtained, related to both the project **implementation process** and the **results achieved**: *Without evaluation, it would be practically impossible to run an innovative project, because how else would we get information on whether the idea is being developed in the right direction and is bringing the expected effects.* A respondent who commissioned the evaluation of innovative projects regretted that it was not carried out for other projects, especially investment ones: *I feel that without evaluation the project was implemented "in the dark". Had there been evaluation, perhaps we could have avoided many of the mistakes we made in that project. I believe that the project in question should have been evaluated in terms of sustainability of the business that was created as a result.*

Evaluation was described as highly useful in particular by those respondents who implemented **pilot projects**: *Evaluation was particularly useful in a pilot project, where new system solutions were introduced and current information on ongoing activities was needed; There must be evaluation in every pilot and systemic project, because it is a form of control of how large public funds are being spent. Although, in retrospect, I think that the evaluation could have been done with less money if its quantitative part was smaller* [...]. *And fewer evaluations in general could have been carried out, because a final evaluation was also done in that project, which largely repeated the findings of the ongoing one. Initially, the project was to have slightly different goals, so the needs related to evaluation were also different. Then the project had to be changed during its course, but the evaluation assumptions remained the same, so the evaluation component ended up being too large.*

Those who had not previously participated in an evaluation and did not have any knowledge of this area generally equated evaluation to **monitoring**, **reporting** and **control**, and focussed mainly on the degree of implementation of the **assumed outcomes**. In those cases, evaluations were **not rated as highly useful**, because the coordinators felt they had not learnt anything new. Those persons said that they would not have evaluated their projects if there was no such requirement: *I was mostly interested*

in proving the achievement of indicators. I had analysed the outcomes myself before, so nothing in the report surprised me. Those respondents who performed evaluation only because of an **external requirement** believed its main purpose was to demonstrate that the project was **carried out correctly**. This was due to the **fear** of losing the allocated funds: *Our evaluations were conducted with a certain "ready conclusion" in mind because there was a risk that if something went wrong, we would have to give the funds back or they would be withdrawn.* In such cases, the evaluation was only an **additional obligation** and a **burden** for the project coordinator: *It was a redundant obligation, completely unnecessary.*

If I didn't have to, I wouldn't have done it because of the workload. If management procedures are in place, goals are kept in sight, and individual elements are implemented, then evaluation is not needed, because indicators are achieved. The design of the project was simple, but in order for it to receive funding, various additional elements were needed, such as gender and competence growth indicators. There were also too many indicators for evaluation, which greatly complicated the process.

The usefulness of evaluation carried out only to meet **formal requirements** was **rated low**. The fact that such evaluation was considered less useful was also due to its **superficial nature** and lack of sufficient **engagement of the contractor**, who did not bother to get to the bottom of the specific nature of the project: *The evaluator was not able to delve into the specifics of the project. The conclusions were general, no new lessons were learnt.* Lack of **expert knowledge** on the part of contractors also reduced the usefulness of external evaluation carried out in innovative projects: *In our innovative project, the partner* [i.e. internal evaluator – editor's note] *had the time and many opportunities to learn, but this was missing in the external evaluation. This problem was difficult to solve because the principle of competitiveness does not allow us to define the selection criteria for an evaluator in too much detail. As a result, the offer with the lowest price usually wins, because even the criterion of three or five years of experience is not helpful here.*

Some respondents emphasised that obtaining useful, quality evaluation requires considerable **commitment** not only from the contractor, but also from the **contracting party**: *Contracting parties are often unaware that this requires a lot of work on their part too. You have to contribute to the concept, tools, report... you just need to know what you're doing. The contracting party often lacks those features or provides sloppy input, so poor quality is not only the fault of the company. Preparing a description of the subject of the contract*[11] *and all that follows is very labour-intensive and time-consuming, and if you don't do it, you end up with a poor-quality report. Both sides are responsible for the quality of evaluation.*

Current interest in evaluation

Another area of the research concerned the **change of approach** to project evaluation, which was reflected in the 2014–2020 Guidelines of the Minister of Infrastructure and Development for Cohesion Policy Evaluations. Provisions relating to the eligibility of expenditure under the European Regional Development Fund, the European Social Fund and the Cohesion Fund have significantly cut down on evaluations for individual projects in favour of evaluations in relation to a given Programme, Priority or Action, and therefore at a more general level. Although project evaluation is still possible, for it to be recognised as an eligible cost (i.e. one that can be covered by EU funds) the approval of the

11 It is a part of a Terms of Reference (ToR) – in this case, regarding an evaluation.

Managing Authority of the relevant Operational Programme is required. In this situation, evaluation at the project level is possible only in particularly justified cases – as an exception to the rule.

The respondents were to a **varying degree** interested in conducting project evaluation in the current financial perspective (i.e. in the years 2014–2020) and had differing opinions about the new approach of the Ministry of Infrastructure and Development[12]. Respondents who had **positive experiences** with evaluation and appreciated its benefits believed that project evaluation was necessary to ensure a **high quality** of their activities, especially in relation to educational and social projects: *This was a very useful requirement, because not only did it help project implementers to control whether the project was going well, but also the Ministry was able to assess the entire intervention based on all evaluations; We care about quality, so we would definitely evaluate projects, if possible* [...]. *However, we wouldn't be able to hire additional people or external contractors. I regret that there's no longer such a requirement.*

As a result of the new project evaluation regulations, the respondents have been deprived not only of knowledge about the **quality and effectiveness** of their actions, but also of information about improvement opportunities, prevention of **difficulties and risks** and **in-depth analysis**: *I'm so glad that the previous perspective included a requirement to evaluate projects. I believe that evaluation is necessary to properly implement projects or activities. It is a source of knowledge about quality and effectiveness. It makes it possible to introduce improvements and risk prevention; I think it* [universal project evaluation – editor's note] *was valid and justified. It provided a basis for reliable, evidence-based, constructive conclusions. It also allowed for an analysis of problems from different perspectives; I think this* [introducing restrictions on project evaluation – editor's note] *is a bad move. Not everyone feels the need for or is aware of the importance of evaluation. Without specific guidelines regarding the necessity and value of evaluations, many implementers will not evaluate their actions, which may translate into lower quality, misalignment between actions and goals, and exposure to project risks.*

From this perspective, the current approach to project evaluation (that is requiring MA approval, among others) is seen as **limiting**: *I consider the current solution to be a limitation. We were thinking about including evaluation as a task to be commissioned to an external contractor within a cooperation project, but due to those limitations we abandoned the idea.* Respondents who **disapproved** of the new guidelines on evaluation pointed out that previously they were "forced" to care about the **quality** of projects, e.g. by collecting **feedback** from the beneficiaries and **improving** not only the current activities, but also subsequent undertakings: *If someone was not interested in the quality of what they did and did not want to improve it, this requirement made them adopt a different approach and showed them this needed to be done. What happened with the findings of such evaluation is another matter. Awareness of such a need is not widespread, sometimes people have to be made to do it; After participating in the SMART programme, I came to understand the need and became aware that you have to listen to what people want to say about the project; I consider this to be a mistake, because it lowers the quality of projects; I disapprove of it, because we've now lost the positive impact of evaluation findings on subsequent projects.* Although some projects may be evaluated **in-house**, i.e. using internal resources, for others external evaluation would prove more useful: *In some projects, it's sufficient to perform an in-house evaluation, but sometimes it makes more sense to use an external expert, who has a fresh perspective and is not involved in the execution process.*

12 In the previous programming period, i.e. in the years 2007–2013, evaluation of projects financed from EU funds was obligatory.

Coordinators who performed evaluations only because of the external requirement and who did not feel they benefited from them **approve** of the current limitations on project evaluation. Nevertheless, even some of those who consider this solution to be right conceded that whenever coordinators see such a **need** they should be able to conduct or commission an evaluation at their own discretion – possibly in consultation with the **project implementer** – rather than based on the approval of the MA: *This requirement was unnecessary in innovative projects but if someone wants to do it, they should be able to do so without unnecessary formalities.* Respondents are worried about the **quality** of current projects due to **lack of evaluation** and feedback: *It isn't right* [that no evaluation is carried out – editor's note] *mainly because IB project supervisors are not interested in the quality of projects, they're only interested in the indicators. Nobody cares if the right people are part of the project and how they benefit from participating in it.*

Yet another group of respondents shows an **ambivalent attitude** towards the introduced changes – on the one hand, they appreciate the benefits of evaluation, but, on the other, they perceive it as an additional burden or unnecessary expense when it is conducted without a **clear need**: *Clear-cut evaluation does not exist. I appreciate the value of evaluation, but I believe that it should not be a task of the project implementer. I'm glad that this duty was taken away from project implementers. If I were to implement a pilot or innovative systemic project, I would definitely carry out an evaluation, but it's only useful if I have the expert knowledge and experience of evaluators; For some – who don't see the need for it and think it's of no use to them – it may be a relief that they don't have to do the evaluation, they feel they've stopped spending money on nonsense. In my opinion, evaluation is a sign of respect for your colleagues and beneficiaries, a question about the competencies of our audience, how they behave in difficult situations, what they're unable to do; I'm glad that the requirement is gone, but it's bad that in practice it's also no longer a possibility.*

According to the respondents, the requirement of conducting an evaluation **does not guarantee** that its findings will be used – this is due not only to the experience and quality of the evaluators' work, but also to the level of **awareness of decision-makers**: *A lot depends on the awareness of project and institution managers. If they're not interested in and convinced about evaluation, its findings won't be used.* When asked if they would have evaluated projects if there had not been such a requirement in the previous financial perspective, respondents generally **answered in the affirmative**: *The requirement of evaluation was justified; In each project, regardless of the requirement, I conducted an evaluation; Yes, because I have very good experiences from previous evaluations. They were done by high-class experts and professionals who contributed a new quality, knowledge and viewpoint to the project and the project team, and this couldn't be replaced with anything else.* However, others stated that they would not have carried out the evaluation without an external requirement because they did not have sufficient **awareness or the need** to conduct it: *I probably wouldn't have conducted it, because awareness of the essence and usefulness of evaluation was at that time generally very low. Moreover, evaluation is expensive; An important problem is the continuing low awareness of what evaluation is and what it can offer. The level of this awareness is definitely higher among NGOs operating in the business environment, e.g. those that obtain funds and grants from companies. In those circles the expectation of effectiveness, and the ability to demonstrate it, is much greater. The level of knowledge among small organisations operating locally and financed from public funds is definitely lower.* The requirement to carry out evaluations was also linked to the efficiency of spending **public money**: *Evaluations should be done and should be a requirement. Both for pragmatic reasons and out of respect for public money.*

If project evaluations were not currently required to be approved by the MA, many respondents would perform them, especially in relation to **difficult**, **innovative** or **pilot** projects, as well as in the case of planning **new activities** or working for **other target groups**: *We would ask for an evaluation of difficult and demanding projects where, for example, the target group is different than usual, or if we were planning any new solutions or activities. Evaluation would be desirable for pilot projects, where there is a high degree of uncertainty.* Others indicated that evaluation would provide them with **feedback** that would **improve their activities**, **eliminate errors**, and better tailor projects to the **needs of beneficiaries**: *I would conduct evaluations to introduce improved activities and innovative solutions, and eliminate errors in the future. Thanks to evaluation, subsequent projects better match the needs of their beneficiaries; It's mostly about project participant feedback that is as independent as possible. It provides an insight into the effectiveness of the actions taken and the accuracy of selection of the working methods.*

Some respondents declare that they would prefer to evaluate only **some projects**, e.g. of an educational nature, while others indicate the need to evaluate **all their projects**: *I would evaluate educational projects* [...], *other projects – no; I think that evaluation is a very valuable tool, it should be done for every project; I would conduct evaluation for every project. The purpose of such evaluation would be to determine the level of achievement of objectives and outcomes, determine the effects of the actions taken, the relevance of proposed activities, selected methods and target groups, and whether funds are used adequately; All educational activities must be evaluated, how else would you obtain information on how effective the training or other support has been or what people need and expect? Evaluation in education (including adult education) is now a standard, I can't imagine not collecting feedback from participants or not giving them the opportunity to comment on what they are taking part in.* Some coordinators would **limit the scope** of evaluation in the absence of such a requirement: *We would evaluate even without being required to do so, but probably to a lesser extent; We have some projects in the pipeline, to be financed with Norwegian funds, and the 2014–2020 POLAND–RUSSIA Cross-Border Cooperation Programme. We could really use evaluations in these projects, outsourced to entities with a serious approach, rather than done as an in-house effort.*

Some respondents also pointed out the risk associated with the fact that evaluations are an **imposed requirement**, which may lead to **inefficient spending**. In the previous financial perspective, there were sporadic fabricated evaluations – without foundation in the **actual need** of the project originator: *There were certain very superficial, even fictitious evaluations; I wonder if the MA read those reports with comprehension.* If all this boils down to is ticking off a requirement and spending money, then such evaluation is pointless. If the previous regime is restored, i.e. widespread evaluation, in some cases this would lead to a **waste of funds**, especially for low value activities.

Some coordinators believe that evaluations should focus on in-depth analysis of the **usefulness and sustainability** of projects, rather than merely on their **ongoing implementation**: *Usefulness and, above all, sustainability of the project should be examined. A business model is needed to maintain the outcomes.*

Evaluations should focus on whether the proposed products or services will sell, whether there is a market for them, whether it will be feasible to maintain and develop them. Potential should be evaluated, and user satisfaction surveys should be conducted. According to many respondents, evaluation should **obligatorily** accompany pilot projects which test new solutions.

Reasons for not applying for project evaluation

It should be emphasised that some coordinators **did not know** that project evaluations are still **possible**:

If it were possible, we would include evaluation in the application for funding; Evaluation is now unwelcome, because it's seen as a cost. Nobody told us that we could still ask the MA for evaluation, I've not heard of such an option; If it were possible to include evaluation as an external cost, we would use it in the 2014–2020 perspective by outsourcing it [...]. *Due to the fact that according to the current rules evaluation would have to be financed from non-EU funds, it could be considered non-frugal and so the idea has been abandoned.*

Others pointed to the **difficulties** associated with the need to **apply** for evaluation: *The application would have to be submitted together with the project application, and no one will do it at this stage, because we don't know if we'll get funding. And then, when the project is approved, it's too late to change anything. The whole thing is designed to get rid of project evaluations so that nobody scrutinises the officials and discovers their incompetence; Currently, the company is implementing 17 social projects under the ROP*[13]*. There's not the slightest evaluation in any of them, no participant surveys, no feedback is collected from them at all.*

Although the respondents' opinions on the principles of financing evaluations in the current financial perspective were varied, they may indicate certain **ignorance of these issues**. Some expressed the willingness to do evaluations as part of **indirect costs**, while others intended to **conceal the cost of evaluation** under other eligible costs: *We now have a part-time evaluator under indirect costs (in an ROP project), they're an employee of the institution implementing the project; I would hide the cost of evaluation in some other costs.* Some declared that, should they feel the need to conduct evaluation, they would disguise it as some other type of research (e.g. analyses, measurement of competence growth).

The vast majority of respondents **disapproved** the fact that project evaluation was subject to the decision of the MA. According to them, the **project coordinator** and/or the **governing bodies** of the entity which is implementing the project should **make the call** on evaluation: *The project coordinator should be the only person to decide whether there's such a need and whether feedback is required. This shouldn't be a top-down obligation, it should depend on the type of project, for example it's not needed for simple projects. This should be at the discretion of the beneficiary; The MA could recommend the use of evaluation. But the ultimate decision should be made by the project implementer; The decision to conduct an evaluation should belong not only to coordinators, but also to the management of the relevant institution; It's a pity that the possibility of evaluating a project depends solely on the decision of the MA. The opinion of an external evaluator could be valuable. To my mind, the MA should make it possible for all project implementers to carry out evaluation if they signal such a need.* It was emphasised that the decision on project evaluation should be based on **specific criteria**, which should be the same for all beneficiaries: *The possibility of conducting an evaluation should be based on clear criteria and premises which are the same for all. And not on an arbitrary decision of an official/expert.*

One respondent was of the opinion that project evaluation should be carried out by an **independent institution** specially appointed for this purpose: *There should be an external institution responsible*

13 Regional Operational Programme.

for project evaluation. Every project asking for, say, PLN 200,000 should be automatically given e.g. an additional PLN 10,000 for evaluation, not from the project budget, but independently. And the evaluation would be carried out by an independent agency. Its role would be to support project implementation by providing evaluation data in the course of the project. Another person suggested that the decision whether to evaluate a project should belong to the **Intermediate Body**, which has more knowledge about the **particular project** than the MA: *Possible negotiations and decisions should be made by the IB rather than the MA* [that is at the level of the institution organising the competition – editor's note]. *This should be an element of the application, a part of negotiations and determining the scope of the project, which takes place at the application stage anyway. Then the beneficiaries would probably be more willing to use evaluation, without fear that the decision is made by an institution which doesn't really know their project.*

The respondents emphasised that evaluation carried out at the programme level **cannot replace** evaluation of projects and does not provide their implementers with **adequate support**: *The MA should also know what's happening in the projects, what facilitates the implementers work and what disturbs it. Evaluation at the programme level will not ensure this. It'd be good to evaluate at least certain types of projects, not only at the end of activities, but also during, so that the MA knows how it can help the implementers.*

Needs in the area of evaluation-related competencies

Some coordinators **were not interested** in developing their own evaluation competencies, which they could use during the implementation of projects, because they preferred to **outsource** evaluation to **specialised contractors**: *I'm in charge of project management, it's just not something I want to explore; If necessary, I would hire an external evaluator, because such evaluation is more objective, the person is less biased and offers a fresh look.* Several persons said that they would like to take part in **training on evaluation**: *I haven't conducted any evaluations because I don't have the qualifications.* One of them said that they could use a course that would provide them with knowledge enabling them to **partner up** with other evaluators – for example they would like to learn how to design evaluations and what their benefits are. In addition, the importance of **levelling up evaluation competencies** within the **project team** was emphasised: *It is important to improve the skills and knowledge of research methods and evaluation usability, striving to ensure that awareness of the needs and legitimacy of evaluation are at a similar level throughout the project team.*

According to one of the respondents, evaluation knowledge is **universal** and can be **useful** in other work, not directly related to the implementation of EU projects: *Knowledge about evaluation is universal, it's useful in any work, not just project work. Especially as the degree of knowledge and skills among project teams and other employees is low. There is often no knowledge of the basics, such as methods, tools, goal-setting, outcomes, deliverables, indicators, inference, and making recommendations.*

In the respondents' opinion, other members of the project team have **diverse needs** when it comes to developing their evaluation competencies: *Some would like to learn something extra, but this is a very diverse need. Some want to know the findings of an evaluation so they can change something, and there are also those who think it's nonsense, even though they are teachers, because there's this widespread conviction in schools that evaluations are of no use.*

Some people pointed to the need for **external consultations** regarding evaluation, as well as the lack of a coherent and complete **source of knowledge** on this subject. This information is available on the Internet, but it is very scattered: *Such a source of knowledge is needed, not only for me, but for a wide range of people operating in the third sector. For many people, participation in away workshops is not an option. Mobile consultations via Skype would also be valuable. Such solutions could contribute to popularising evaluations.*

Respondents' recommendations

When asked which elements or solutions they would recommend, and which practices are better avoided in relation to conducting evaluations, the respondents gave answers that referred to **various stages** of the evaluation process.

Some people pointed to the need to take into account different procedures for awarding a public contract, as well as to extend the criteria for **selecting contractors**, so that price is not the only, or main, condition for commissioning evaluations: *There should be a wider range of possibilities in developing criteria for selecting evaluators; Price cannot be the only criterion of evaluator selection. You need to take into account experience and references, although it's probably always a bit of a lottery. We have experience in negotiating and this has worked very well; The evaluator should have in-depth knowledge of the area they are investigating* [...]. *Without such expert knowledge, the evaluation is less thorough and therefore less useful; Evaluation quality should be ensured by requiring evaluators to have the appropriate competencies.*

It would be advisable for the **description of the subject of the contract** to be developed by **experts**: *The description of the subject of the contract should be prepared by someone who knows about evaluation or at least research in general.* **Knowledge about evaluation** should be **promoted** to a greater extent so that it is not reduced to surveys, awareness of how **findings can be used** should be raised, the process of **implementing recommendations** should be monitored, and the effect of the attitude and commitment of the **contracting party** on the quality of evaluation should be highlighted: *Evaluations cannot be equated with surveys. You have to strongly promote awareness of what to do with the finished report and evaluation findings, and how to prepare documents. It would be worth looking at what happens with the reports, whether they have contributed to changing anything. And remember that the responsibility for evaluation lies on both sides – the contractor and the contracting party.*

In addition, attention was drawn to the need to allocate **more resources** to evaluation, which would make it possible to broaden its **scope** and increase its **usefulness**: *There should also be larger amounts earmarked for evaluation, so that something extra could be expected from it and it is really useful to the contracting party; Don't cap the cost of evaluation below a certain threshold, because this definitely lowers its quality and makes it pointless. Choose research methods in a thoughtful and purposeful way.*

People who equated evaluation with the effectiveness criterion pointed to the burden of **monitoring project indicators**: *It would be better without these complicated indicators. It turns out that it's important what you say in the application, because then you have to prove it and account for it. There were stupid assumptions about indicators in the programme, but they had to be included in the application and then evaluated.*

It was also emphasised that evaluation should not be carried out solely because of **formal requirements** – the most **valuable** and **useful** evaluations respond to the **information needs** of

the contracting party, so the evaluator should be able to identify them: *Evaluation should not be done just to meet formal requirements. The evaluator should be able to elicit the needs and listen to the contracting party, which has knowledge of the project; When carrying out an evaluation, it's worth paying attention not only to its cognitive, but also, and above all, to its utilitarian nature. Rely on practitioners and researchers, rather than theoreticians. Work closely with the contractor in the case of external evaluations, in particular in terms of their objectives and scope. In the case of internal evaluations, ensure objectivity and lack of bias potentially resulting from the evaluator's place along the chain of command.*

Conclusions and recommendations

The findings of this qualitative research have led to the following conclusions:

- → The surveyed coordinators had **diverse experiences** with evaluation, mainly related to the implementation of educational/social projects;
- → How they assessed the **usability** of evaluations, and how strongly they felt about the **need** to conduct them in the current financial perspective (2014–2020), depended on their prior experiences – primarily the **benefits** of **evaluations** carried out in respect of previously implemented projects;
- → Respondents who had not had any prior evaluation experiences or conducted it because of an external requirement equated evaluation with **monitoring** and **control**; this approach meant that no new knowledge was acquired – instead, evaluation served merely the purpose of demonstrating the **degree of achievement** of predetermined outcomes and verifying the **correctness** of project implementation, which was to prevent the possible loss of awarded funds; these persons perceive evaluation as an **additional burden** and an **unnecessary expense**, and they **approve** the current limitations on project evaluation; because of the **low usefulness** of such a conception of evaluation, they claim they would not normally perform it, unless it was required; some respondents drew attention to the risk of **compulsory evaluation**, which could lead to various **irregularities**, **inefficient spending** and **waste of funds**;
- → Coordinators with a **favourable attitude** towards evaluation, which stems from their positive experiences in this area, perceive it as a tool to **support** the process of project implementation, or even an inherent element of project management, mainly serving the following purposes: **obtaining feedback**, **ensuring quality**, **improving** the tasks carried out by developing **corrective actions**, **dissemination** of best practices, and **avoidance of errors** in subsequent projects, as well as signposting the **direction for further work** and strengthening **team spirit**; the need for evaluation was also linked to the efficiency of spending **public money**;
- → Some respondents were more appreciative of **internal** evaluations, which they associated with ongoing feedback, as well as better knowledge of the project and the needs of its implementer, while others preferred **external** evaluation due to its objectivity, expert knowledge and the possibility of placing it in a wider context; according to the respondents, the **quality of evaluation** depends on the knowledge and experience of the evaluators in a given **research area** or **specific project theme**, as well as the procedure and **criteria for selecting** contractors; some persons emphasised that **their own involvement** in evaluation has an impact on its quality;

- → Evaluation findings are most useful to the **project team**, followed by project **stakeholders** and the **management** of the implementing entity; **frequent changes** in management were one of the key difficulties in using evaluation, along with disregarding its findings in subsequent activities and their insufficient dissemination; low usefulness of evaluation was also due to its **superficial** nature and **insufficient involvement** of the contractor, who failed to sufficiently delve into the specifics of the project or lacked expert knowledge; on the other hand, the factors which contributed to rising the degree of evaluation usability included its **high quality**, **good cooperation** with the contractor, as well as the manner of conducting this process, especially adopting a **participatory approach**;
- → Coordinators aware of the **benefits** of evaluation believe that it should be motivated by their **own needs**, rather than an imposed requirement, and are genuinely interested in evaluating currently implemented projects; the new approach to project evaluation, which requires MA approval, is often **disapproved** of and seen as a **limitation**; according to the respondents, the decision should belong to project **implementers/coordinators** or the management of the implementing entity, or, alternatively, to Intermediate Bodies, which are more **familiar** with a given project than the MA;
- → Some respondents **did not know** that it was possible to evaluate projects in the current financial perspective, while those who were more versed in the current guidelines said that in practice **evaluation has been eliminated** at project level; this state of affairs results in lack of **feedback** from beneficiaries, as well as absence of knowledge about the **quality** and **utility** of implemented actions, how they can be **improved** and how to **prevent** various **difficulties** and risks, as well as about the **sustainability** of the outcomes; according to the respondents, evaluation at the programme level **cannot replace** the evaluation of projects because it does not provide their implementers with **adequate support**;
- → Some coordinators say that in the previous financial perspective they would not have performed evaluations **on their own initiative** (i.e. without an external requirement) because they lacked sufficient **awareness** and felt no **need** in this regard; many respondents declared that they would now happily evaluate projects – especially difficult, innovative or pilot ones – if this **did not require approval** by the MA; coordinators pointed out the **difficulties** associated with the need to apply for approval of an evaluation and the **reduction** of the scope of evaluations conducted in-house; some intended to conceal evaluation expenditure under other eligible costs e.g. by conducting evaluations under the pretence of other studies;
- → The needs of coordinators and other members of project teams to develop evaluation competencies are **varied**; some respondents prefer to commission evaluations to specialised contractors, while others would like to have the knowledge to be able to **partner up** with evaluators; in addition, the need to **level up competencies** across the **project team** was emphasised.

Based on these findings, the following recommendations can be proposed:

- → Feeling the need for evaluation stems from **positive experiences** in this regard; current solutions related to project evaluation largely prevent learning from **evaluation in practice**, therefore they should be more **project originator-friendly** and **encourage** those project teams which are **motivated** to do so to perform evaluations;

- → A requirement/obligation to perform evaluations may lead to a decrease in their quality and make them essentially fictitious, therefore evaluation at the project level should be **voluntary**; at the same time, it is worth encouraging the evaluation of projects which **test** new solutions (pilot and innovative projects) or which are addressed to a new target group;
- → Project coordinators/implementers should educate themselves about the current principles of **financing evaluations** (e.g. from the summary prepared by the Ministry of Infrastructure and Development); decisions regarding project evaluation issued by the MA should be based on **uniform criteria**;
- → It is worth promoting **various procedures** for awarding public contracts concerning evaluation (e.g. competitive negotiations), as well as various **criteria for selecting** contractors, so that price is not the leading condition; tender documents (e.g. description of the subject of the contract) should be developed by **experts** with in-depth knowledge of evaluation; contracting parties should be made aware that the **resources** allocated for evaluation (financial resources, time for its implementation) should be adjusted to the **scope** of the proposed evaluation;
- → In some cases, it is worth empowering project teams to carry out **deferred** evaluations, which will focus, for example, on changes in attitudes, increase in competencies, and the **usefulness and sustainability** of outcomes;
- → There should be greater efforts to **promote knowledge** about evaluation so that evaluations carried out at the project level are not based solely on quantitative methods; furthermore, appropriate **use** of evaluation findings and monitoring of the **implementation of recommendations** should also be fostered; it is worth making project teams aware that the attitude and commitment of **contracting parties** has an impact on the quality of evaluation;
- → Due to the fact that evaluation competencies and experiences translate into a higher opinion of its usefulness and greater use of its findings, coordinators and other project team members should be **educated**, including through the use of **e-learning/blended-learning** (such as the course "Weź kurs na ewaluację" developed by the Polish Evaluation Society), as well as **webinars** and **remote consultations**; raising awareness of the benefits of project evaluation and of the differences between evaluation, monitoring and control, may contribute to shaping **positive attitudes** towards evaluation;
- → This qualitative research should be followed up by a **quantitative** study, i.e. computer-assisted web interviewing (CAWI) with the participation of project coordinators in order to determine the **scale** of the opinions/phenomena identified through individual interviews.

Bibliography

Grzywa, M. (2008). Ewaluacja w małopolskich jednostkach samorządu terytorialnego. In: M. Grzywa et al. (eds.), *Ewaluacja w samorządzie lokalnym – perspektywy i możliwości.* Krakow: Fundacja Rozwoju Demokracji Lokalnej, Małopolski Instytut Samorządu Terytorialnego i Administracji.

Guba, E.G. & Lincoln, Y.S. (1989). *Fourth Generation Evaluation*. London: Sage.

Hryniewicka, A. (2008). Ewaluacja w pomocy społecznej – badania własne Instytutu Rozwoju Służb Społecznych. In: J. Staręga-Piasek & A. Hryniewicka (eds.), *O potrzebie ewaluacji w pomocy społecznej*. Warsaw: Instytut Rozwoju Służb Społecznych.

Jaskuła, S. (2018). *Ewaluacja społeczna w edukacji*. Krakow: Impuls.

Katalog Polskiego Towarzystwa Ewaluacyjnego (2011). 1st ed. Warsaw: Polskie Towarzystwo Ewaluacyjne.

Kołodziejczyk, J. & Kołodziejczyk, J. (2015). Wizytatorzy o swojej pracy w kontekście ich nowej roli zawodowej. In: G. Mazurkiewicz (ed.), *Jak budować dobrą szkołę. Potencjał i bariery ewaluacji w oświacie*. Krakow: Wydawnictwo Uniwersytetu Jagiellońskiego.

Korporowicz, L. (2008). Interakcyjne aspekty procesu ewaluacyjnego: pomiędzy analizą a animowaniem zmian społecznych. In: A. Haber & M. Szałaj (eds.), *Środowisko i warsztat ewaluacji*. Warsaw: Polska Agencja Rozwoju Przedsiębiorczości.

Korporowicz, L. (2011). Zmienne losy polskiej ewaluacji. Pomiędzy nadzieją, animacją i konfuzją. In: B. Niemierko & M.K. Szmigiel (eds.), *Ewaluacja w edukacji: koncepcje, metody, perspektywy*. Krakow: Polskie Towarzystwo Diagnostyki Edukacyjnej.

Mizerek, H. (2016). Ewaluacja edukacyjna w Polsce. Trajektorie, perspektywy i dylematy rozwoju, *Zarządzanie Publiczne, 1*(33), pp. 1–16.

Model ewaluacji w pomocy społecznej (2010). Krakow: Regionalny Ośrodek Polityki Społecznej.

Ocena zdolności monitoringu i ewaluacji we wdrażaniu RPO WZ na lata 2007–2013 oraz systemu wdrażania rekomendacji wykonanych badań ewaluacyjnych. Final study report (2010). Warsaw–Szczecin: CASE – Doradcy.

Patton, M.Q. (2008). *Utilization-Focused Evaluation*. 4th ed. Thousand Oaks, CA: Sage.

Patton, M.Q. (2013). Ewaluacja skoncentrowana na wykorzystaniu. In: A. Haber & Z. Popis (eds.), *(R)ewaluacja. Poszukiwanie nowych metod oceny efektów*. Warsaw: Polska Agencja Rozwoju Przedsiębiorczości.

Polski rynek ewaluacji i badań dla sektora publicznego (2011). In: *Katalog Polskiego Towarzystwa Ewaluacyjnego*. 1st ed. Warsaw: Polskie Towarzystwo Ewaluacyjne.

Programming Period 2014–2020. Monitoring and Evaluation of European Cohesion Policy, European Social Fund. Guidance Document (2015). Brussels: European Commission.

Rozporządzenie Ministra Edukacji Narodowej z 7 października 2009 r. w sprawie nadzoru pedagogicznego. Journal of Laws of 2009, No. 168, item 1324.

Rozporządzenie Ministra Edukacji Narodowej z 27 sierpnia 2015 r. w sprawie nadzoru pedagogicznego. Journal of Laws of 2015, item 1270.

Society at a glance 2016. A spotlight on youth, how does Poland compare? (2016). Paris: Organisation for Economic Co-operation Development. www.oecd.org/poland/sag2016-poland.pdf [accessed: 5 March 2019].

Sury, Z. (2016). Wybrane przejawy postaw wobec ewaluacji w szkole, czyli Baśń o Kopciuszku w murach polskiej szkoły, *Edukacja Elementarna w Teorii i Praktyce*, Vol. 11, *2*(40).

Sury, Z. (2017). Ewaluacja w szkole – w kierunku kontroli czy wspomagania?, *Szkoła – – Zawód – Praca*, *14*.

Trawkowska, D. (2007). Działania pozorne w pomocy społecznej. Przypadki (i przypadłość) pracy socjalnej, *Problemy Polityki Społecznej*, *10*.

Walczak, B. (2015). Ewaluacja oczami szkoły. In: G. Mazurkiewicz (ed.), *Jak budować dobrą szkołę. Potencjał i bariery ewaluacji w oświacie*, Krakow: Wydawnictwo Uniwersytetu Jagiellońskiego.

Wasilewska, O. et al. (2014). *Wykorzystanie ewaluacji zewnętrznej i wewnętrznej przez szkoły*, Warsaw: Educational Research Institute.

Worek, B. (2013). Uspołecznienie ewaluacji czy rytualizacja badań ewaluacyjnych? Refleksje o wykorzystaniu ewaluacji w społecznościach lokalnych, *Animacja życia publicznego. Zeszyty Centrum Badań Społeczności i Polityk Lokalnych*, *4*(11).

Wytyczne w zakresie ewaluacji polityki spójności na lata 2014–2020. Warsaw: Ministerstwo Infrastruktury i Rozwoju.

Życie codzienne organizacji pozarządowych w Polsce. Study report (2012). Warsaw: Stowarzyszenie Klon/Jawor, fakty.ngo.pl/raporty/organizacje-pozarzadowe-2012-artykuly-badawcze [accessed: 13 February 2020].

Evaluation as a tool supporting the management and development of seniors' organisations on the example of the third age university movement

DOI: 10.47050/66515246.42-49

Jakub Wróblewski

Abstract

The demographic structure of Polish society is changing. A growing share of the country's inhabitants are people of non-working age. As any other new phenomenon, the process of society ageing involves challenges, some of which can only be answered by top-level authorities, and some by grass-roots social activity. But there also exists an area which requires cooperation between central and local authorities and non-governmental organisations: it is support provided to the development of seniors' organisations. This chapter deals with the part of senior citizens' movement that has been developing particularly dynamically over the past dozen or so years, in response to the needs of social, cultural and educational activity of senior citizens. Here we mean third age universities. These institutions face an extremely important task of confronting the problem of an ageing society and counteracting the negative effects of this demographic process. It is, therefore, necessary to look for solutions that will make it possible to build stable foundations for the development of third age universities in all their dimensions and organisational forms, while not limiting the grass-roots character of their activities. One such solution is including evaluation activities in creating development strategies as well as in the current management of seniors' organisations.

It is not possible to discuss current and future developmental challenges in isolation from the demographic context. In Poland, like in other European countries, the population of people of non-working age is growing. Decision-makers at both the national and transnational level are commencing to take this challenge into consideration. What best reflects the fact that public attention has been drawn to this process was the decision to declare 2012 the European Year for Active Ageing and Solidarity between Generations (*Decision...*, 2011). This had a symbolic meaning and was aimed at stressing the significance of this challenge.

Jakub Wróblewski
Sociologist, evaluator, social researcher and analyst. Expert of the Centre for Evaluation and Analysis of Public Policies of the Jagiellonian University. Co-creator of one of the first Polish Internet panels ResearchOnline.pl. Member of the Polish Society of Market and Opinion Researchers (*Polskiego Towarzystwa Badaczy Rynku i Opinii*), as well as President of the Polish Evaluation Society and Member of its Management Board. He specialises in conducting research and providing assistance for the public market, central administration and local government entities.

KEYWORDS:
EVALUATION,
THIRD AGE,
THIRD AGE UNIVERSITIES,
SOCIETY AGEING,
SENIOR ACTIVATION

According to Edward Rosset (1959), population ageing is a combination of two different demographic processes. One is a direct result of the undisputed success of civilisation, the outcome of which is not only a significant increase in average life expectancy, but also a remarkable enhancement in the quality of life at all its stages. For the first time in history so many people have an opportunity to reach the age of 70 or 80 and remain in good physical and mental health. However, in parallel with this positive and optimistic process, the number of births is steadily decreasing. Similarly to the inhabitants of many other developed countries, Polish men and women decide to start families at a much older age and on average have fewer children than their parents and grandparents. In 2018, the total fertility rate in Poland was 1.4[1]. It is worth noting that a fertility rate below ca. 2.1 means that the population replacement rate is lower than replacement level fertility.

In Poland, the negative impact of a decreasing number of births is additionally augmented by labour migration (Holzer 2003, p. 139), especially to Western Europe. The emigrants are mostly young people in their prime who initially might assume that they will return to the country but who, having acclimatised and started professional careers abroad, change their decision as to where they want to live and start a family. This process does not only apply to disadvantaged areas. All regions of Poland are depopulating – villages as well as small and medium-sized towns. Large cities and urbanised areas are in a slightly better position, as they are a natural destination of domestic migration, although depopulation processes are also observed there, and the average age of their population is gradually increasing (Śleszyński, 2018, p. 226).

The question should be asked if the growth of the share of people aged 60+ in the population is accompanied by a reflection on how modern Polish society can prepare for this process, not only in order to minimise negative effects of the trend, but also to use its potential. Taking note of the observed demographic trends is crucial for understanding the importance for the country's sustainable development of the creation of mechanisms which not only prevent the exclusion of senior citizens, but also foster their activity, development and integration into the mainstream of social life, which is a challenge due to the culturally conditioned low level of activity of Polish society (Halicki & Halicka 2003, pp. 200–201).

Reflection on how to build infrastructure supporting senior citizens' activity is important for yet another reason. It is worth noting that the demographic processes described above are accompanied by a cultural change manifesting itself in the reformulation of the model of life of non-working age people. A growing number of contemporary seniors want to find fulfilment, make their dreams come true, pursue their interests, and discover new, previously unknown areas of activity. It is not surprising therefore that there are more and more public and grass-roots organisations and initiatives that meet these needs. The development of senior citizens' movement observed in Poland is a response to the demographic trends described above, which started to be prominent in social sciences in the second half of the 20th century. But it is also a result of the change in the way of thinking about the third age. While in the 19th century much more attention than previously was paid to children, which resulted in the emergence of modern models of education focussed on their opportunities and needs, current trends in pedagogy and andragogy indicate that in the 21st century the focus will shift from the youngest to the oldest members of society (Tomczyk, 2015, p. 7).

These changes have gradually moved from the sphere of scientific considerations to public awareness. The modern way of thinking about the third age is based on the conviction that it is simply

1 Statistics Poland, Local Data Bank.

one of the phases of life and therefore individuals should be supported in dealing with the problems that may arise at this stage. It is also important to ensure an equal standing of old age with other phases of life, to promote its positive image, and to prepare younger generations for it (Szatur-Jaworska, 2012, pp. 9–10).

A change in the way of thinking about older age also forms part of the **lifelong learning** concept, which makes it possible to keep up with and understand social, cultural or technical changes. The Decision of the European Parliament and of the Council[2] establishing an action programme in the field of lifelong learning is a document highlighting its multifaceted nature. Learning, broadening horizons, improving knowledge and acquiring new competences are no longer the domain of youth. Lifelong learning is a new lifestyle characterised by the inclusion of continuing development at all stages of human life.

Thanks to the changes described above, old age ceases to be a taboo subject, and the third age is no longer considered a phase in which people are excluded from social life. On the contrary, it may be a time when, with adequate conditions provided, it is possible to achieve self-actualisation and make use of leisure time that was lacking at earlier stages. The awareness of the existence of this resource has resulted in the need for institutional support allowing for creative and productive use of free time. The establishment of Third Age Universities (TAUs) has been an answer to these needs. The first TAU was established in 1972 at the University of Toulouse in France on the initiative of Prof. Pierre Vellas. In the following years, more institutions of this type were founded in France and other countries. Already in 1975, the International Association of Universities of the Third Age (*Association Internationale des Universités du Troisième Age* – AIUTA) was established. At that time, the idea of educating seniors also reached Poland, where, as early as in 1975, on the initiative of Prof. Halina Szwarc of the Medical Centre for Postgraduate Education in Warsaw, the Third Age School was established. Subsequent Polish TAUs were opened in Wrocław, Opole, Szczecin, Poznań and other academic hubs. Currently, the TAU movement in Poland encompasses more than 600 organisations (*Universities...*, 2019) operating throughout the country. They work both within the structure of universities and as independent institutions, the latter being founded by local authorities and via grass-roots initiatives of local social leaders. It is important to note that these organisations offer diversified activities as there is no single and universal model of senior citizen activation. The key to effectiveness of such activities is to take into account individual seniors' desires, needs and competences and to enable them to make independent decisions on taking up activity and its nature (Schonbrodt & Veil, 2012, p. 74).

Significantly, the development of the TAU movement supported by, among others, the Polish-American Freedom Foundation[3], as well as by activities undertaken on the local level, such as the Programme for Social Participation of Senior Citizens[4], cannot be analysed only from the perspective of the positive value of providing room for creative leisure in a relaxed atmosphere and the pursuits of people who are no longer professionally active. These are indisputable advantages of participation in senior organisations, but they certainly do not exhaust their mission.

Activation of seniors also brings other benefits, which are important from the point of view both of individuals and of the pragmatics of the government's operations. The growing share of seniors

2 Decision No. 1720/2006/EC of the European Parliament and of the Council of 15 November 2006 establishing an action programme in the field of lifelong learning.

3 PAFF launched a programme to support TAUs in 2005.

4 Government programme implemented in the years 2012–2013 and 2014–2020, which aims at providing conditions for the development of social activity of the elderly by means of actions in four priority areas: education of older people, social activity promoting intra- and inter-generational integration, social participation of seniors, and social services for the elderly.

in the population, as well as growing costs of medical procedures, make it necessary to place much more emphasis than in the past on preventing diseases and limiting their negative effects. The senior movement responds to this need. Active seniors are not only healthier, but thanks to their overall better physical and mental condition they cope much better with diseases both typical of any age and specific to old age. The efforts of TAUs aimed at the activation of seniors perfectly fit this paradigm shift and emphasising the promotion of healthy lifestyle and prevention (Golinowska, 2008, p. 45), which should determine the direction of public healthcare development.

We should strive to make activity in third age organisations, such as rural women's clubs, folk universities, interest clubs, rural and housing estate clubs or third age universities, as natural as professional activity, which not only serves the purpose of making money, but is also an important element of life, especially in its intellectual and social dimension. However, in order to promote participation in pursuits offered by such organisations, it is important to understand what senior organisations are, what makes people use their offer and how their operations affect the daily lives of seniors, their families and friends. The dynamic development of TAUs in recent years and positive changes brought about by these institutions, for both seniors and local communities, may overshadow the problems faced by these organisations and their leaders in their daily activities, which may have a significant impact on the future of the movement. If we agree that even a spectacular success hides a seed of future failure, we can put forward a tentative thesis that the dynamic development of the Polish TAU movement also carries such a risk. A large part of Polish TAUs – especially those founded in the last 10 years – have been grass-roots initiatives of people promoting social activity. These institutions are brainchildren of individuals who, with their passion and commitment, as well as frequently hard and almost always unpaid social work, set up organisations which have given seniors in many towns and municipalities the opportunity to spend free time in a constructive manner. It is local leaders in towns and localities distant from urban hubs who created the structures, worked out agreements with local authorities, reached out to trainers and lecturers, promoted the organisations among seniors in their regions, and who every day organise and coordinate the operation of these organisations. It is thanks to their work that tens of thousands of people can regularly take part in sports activities, go to opera houses, theatres and museums. Without the personal commitment of those individuals the success of their organisations would not have been possible. However, stable and durable structures cannot rely on those leaders alone. Regardless of the documents developed in the course of operations (statutes, regulations, etc.), daily practice and the ability to achieve specific objectives build knowledge resources that are closely related to the leader or a narrow group of people who manage the organisation on a day-to-day basis. This is tacit knowledge. It is individualised, intuitive, and its transferability is limited. The existence of such knowledge and having it at one's disposal can be ignored and underestimated until the person, e.g. a volunteer or an employee, leaves the organisation (Mielus, 2011, p. 107).

Such a model, although it can initially function effectively, is unsustainable in the long run. A leader leaving the organisation means the end of its operations or their significant limitation. Even in an optimistic variant, assuming the appearance of determined successors, the operations will be limited until these persons acquire indispensable knowledge and experience. A solution to this problem can be the standardisation of activities and the creation of stable organisational structures, which will function regardless of the turnover of members of key statutory bodies.

In this context, the need to support the senior citizens' movement, not only in terms of finance and infrastructure (which is addressed, among others, by the Senior+ multi-annual programme[5]), but also in terms of organisation and advice, is particularly important. This challenge is all the more important, because an effective senior policy must impact many different areas. This was noted in the latest strategy *Social Policy for Older People 2030. Safety. Participation. Solidarity.* Each of the three pillars has been comprehensively defined in a way that takes into account different levels and dimensions of fundamental values indicated above. **Safety** is a state and situation in which people have a firm belief that their material, living, mental and spiritual needs are and can be met thanks to adapting material conditions, institutions and social environment organisations to the changing needs of the elderly. **Participation** assumes the presence and involvement of older people in various forms of family, social, sports, professional and civic activities. **Solidarity** provides for society taking into account in its actions the opinions and interests of various generations, which reflects the strength of intergenerational bonds (i.e. a sense of connection with people of different ages) and an awareness of responsibility for those representing other generations.

In view of the above, as well as of the challenges related to building a stable basis for the functioning of senior organisations, the question arises how to establish sustainable and stable organisations at the regional level which will be flexible enough to respond to the changing needs of local communities. This question is all the more important, as the operations of senior organisations are influenced by a number of factors – from legal through cultural to technological, linked to new forms of learning, including distance learning. Therefore, proper identification and diagnosis require a holistic view of an institution's activities.

While programmes implemented at the national level are subject to monitoring and evaluation, the results of which are published[6] so that they contribute to improving subsequent editions of projects, the lack of comprehensive support in the area of evaluation of hundreds of individual TAUs and other seniors' organisations is noteworthy. It should be emphasised, however, that this does not mean that there are no studies focussing on select organisations or policies addressed to seniors adopted by cities and municipalities. However, they constitute a set of independent initiatives of local government officials or third sector representatives or are undertakings of a strictly scientific nature, in which given organisations or programmes form a case study in order to analyse a phenomenon or to diagnose the analytical capabilities of a selected research method or approach. Regardless of the high scientific value of such research projects, they are not intended to provide comprehensive management support to a wide range of leaders of senior organisations on a national scale, as they simply do not serve this purpose.

Also, a number of high-quality studies and guides to support leaders in establishing NGOs in general, including seniors' organisations, is published. As an example, we can quote Third Age Forum (*Forum III Wieku*) publications, such as *Standardy działania uniwersytetów trzeciego wieku w Polsce* (Borczyk et al., 2019). This publication is a structured set of proposals for standards suitable for third age universities in three basic areas of their activity, i.e. formal and legal aspects, content-related and organisational factors, and cooperation with other entities. Another valuable source of information

5 Since 2015, the Ministry of Family, Labour and Social Policy has been implementing the Senior+ multi-annual programme for the years 2015–2020 addressed to local government units, which is a continuation of the Senior-WIGOR programme.

6 Examples include published reports on successive editions of the government's Programme for Social Participation of Senior Citizens or evaluations of particular axes of operational programmes, some of which (e.g. these related to digitalisation or improvement of accessibility of medical services) also support people aged 60+.

are the publications of the Zaczyn Foundation, which boasts the *Polityka Senioralna* magazine devoted to the issues of seniors' movement and activation.

It is of note, however, that the majority of publications and guides omit evaluation or treat it very superficially. Evaluation is an element of organisation management which forms an integral part of the management process and can support its development in the long term. Evaluation of the overall functioning of an organisation and of specific projects can serve as a source of knowledge, which is a foundation for making decisions that allow for modifying actions so that they more efficiently respond to the changing needs of recipients. Particular emphasis should be placed on the fact that reflection on evaluation combines pragmatic and operational threads coupled with axiological ones to describe actions through the prism of their values.

Despite these advantages of including evaluation in the process of managing a third age university or other seniors' organisation, management practice is here often the result of a development strategy and activity planning based mainly on leaders' intuition and seniors' interest in specific forms of activity, reflected e.g. in attendance.

Such an approach, although it may prove successful in some cases, may not necessarily be a good planning method. In the process of evaluation, not only the outcomes of intervention are of interest, but also recommendations on how to implement the developed solutions and how to formulate assumptions concerning the method of implementation of specific undertakings, so that they can translate into achieving the assumed objectives (Górniak & Keler, 2008). It should be stressed that the aim of seniors' movements is not to achieve high participation rates in various forms of activity, but rather to activate and integrate the elderly at various levels. Including evaluation in the process of creating a strategy for the development of seniors' organisations, programme planning or specific project evaluation can therefore provide assistance to leaders in planning the activities so that they not only are positively perceived by the participants, but also correspond to the organisation's values and objectives.

A risk specific to the dissemination of evaluation in seniors' organisations, although it is of a general nature and can also be seen in other fields, is the promotion of evaluation in a limited, even residual form, reduced e.g. to assessing the satisfaction of participants in various forms of activity included in a given TAU's offer. Such a misunderstood and misplaced para-evaluation results in non-realisation of the potential of such reflections and, in the long run, in a gradual decline in leaders' interest in educating themselves about evaluation, which is caused by a limited range of benefits it brings to the organisation. Proper evaluation is not just monitoring, and its role is not limited to recording the number of participants and measuring their satisfaction. Evaluation requires, above all, opening up to all actors of processes subject to analysis and not treating them as the subject of the study, but empowering them to act in a participatory process, in which researchers also focus on the context of the organisation's operations, the recipients of services, and the impact of various external factors that change over time.

Participatory and socialised evaluation can be a step that allows for a transition from spontaneous activity to one based on local potential, capable of long-term development, and responding to evolving needs of subsequent generations of senior citizens.

In the light of data available today, it must be assumed that the process of population ageing will intensify and the social role of the senior citizens' movement will grow. There will also be a growing need for the creation of conditions for active ageing, which, according to a definition proposed by OECD in 1998, is *the ability of people to lead productive lives in the social and economic sphere, despite the passage of time. This means that people can make flexible choices related to managing the free time*

in their life – time spent studying, working, resting and caring for others (Perek-Białas & Worek, 2005, p. 13). It is impossible to ignore these processes or leave the seniors' movement at the mercy of chance, hoping that without support it will efficiently meet the needs of this growing and increasingly self-aware part of society. Taking a responsible approach to supporting seniors' organisations at both the national and regional level, we should strive to create mechanisms that will ensure their autonomous operations, but at the same time provide them with a knowledge capital enabling their development in the long term. Evaluation does not provide ready-made strategies but it is a source of information necessary in the process of their creation (*Ewaluacja w strategicznym zarządzaniu...*, 2010, p. 79) and decision--making. Properly designed and conducted evaluation, thanks to its pro-developmental role, can support leaders and decision-makers in taking actions that satisfy real needs of seniors' organisations, in terms both of meeting the needs of their members, and of their contribution to the overall social mission of the seniors' movement at the national level.

For evaluators, the development of seniors' movement forms a challenge and an opportunity. It is an opportunity to educate about evaluation and disseminate it. Yet evaluation must be understood as a socialisation instrument supporting quality and development, and not as a tool for assessment. Such an instrument allows for changes to be introduced not only in the programme, but also in the organisational structure. What is more, it supports the development of a widespread and diversified seniors' movement.

Bibliography

Borczyk, W. et al. (2019). *Standardy działania uniwersytetów trzeciego wieku w Polsce*. Nowy Sącz: Ogólnopolska Federacja Stowarzyszeń Uniwersytetów Trzeciego Wieku, Fundacja Rozwoju Ziem Górskich.

Decision No. 1720/2006/EC of the European Parliament and of the Council of 15 November 2006 establishing an action programme in the field of lifelong learning. Brussels: European Parliament, Council of the European Union.

Halber, A. & Szałaj, M. (eds.) (2010). *Ewaluacja w strategiczne zarządzania publicznym*. Warsaw: Polska Agencja Rozwoju Przedsiębiorczości.

Golinowska, S. (2008). Społeczno-ekonomiczne konsekwencje starzenia się populacji. In: *Konsekwencje ekonomiczne i społeczne starzenia się społeczeństwa*. Warsaw: Polska Akademia Nauk, Komitet Prognoz Polska 2000 Plus.

Górniak, J. & Keler, K. (2008). *Rola systemów wskaźników w ewaluacji. Teoria i praktyka ewaluacji interwencji publicznych*. Warsaw: Akademia Leona Koźmińskiego.

Halicki, J. & Halicka, M. (2003). Integracja społeczna i aktywność ludzi starszych. In: B. Synak (ed.), *Polska starość*. Gdańsk: Wydawnictwo Uniwersytetu Gdańskiego.

Holzer, J.Z. (2003). *Demografia*. 4th ed., revised. Warsaw: Polskie Wydawnictwo Ekonomiczne.

Mielus, M. (2011). Organizacja oparta na wiedzy jako podstawa przedsiębiorczego zarządzania. In: *Przedsiębiorcze aspekty rozwoju organizacji i biznesu*. Krakow: Oficyna Wydawnicza AFM.

Perek-Białas, J. & Worek, B. (2005). Aktywne starzenie się. Aktywna starość – określenie pojęć. In: *Aktywne starzenie się. Aktywna starość*. Krakow: Aureus.

Social Policy for Older People 2030 (2018). Ministry of Family, Labour and Social Policy, Appendix to Resolution No. 161 of the Council of Ministers of 26 October 2018, item 1169.

Decision No. 940/2011/EU of 14 September 2011 on the European Year for Active Ageing and Solidarity between Generations. OJL 303, 14 December 2011.

Rosset, E. (1959). *Proces starzenia się ludności. Studium demograficzne*. Warsaw: Polskie Wydawnictwo Gospodarcze.

Schonbrodt, B. & Veil, K. (2012). Zjawisko wycofania społecznego w kontekście "aktywnego starzenia się". Potrzeba działania i przykłady dobrych praktyk w Niemczech, *Problemy Polityki Społecznej, 18*(63).

Szatur-Jaworska, B. (2012). Zasady polityk publicznych w starzejących się społeczeństwach. In: *Strategie działania w starzejącym się społeczeństwie*. Warsaw: Commissioner for Human Rights Bureau. www.rpo.gov.pl/pliki/13541772380.pdf [accessed: 13 February 2020].

Śleszyński, P. (2018). *Demograficzne wyzwania rozwoju regionalnego polski*. Warsaw: Instytut Geografii i Przestrzennego Zagospodarowania im. S. Leszczyckiego PAN.

Tomczyk, Ł. (2015). *Edukacja osób starszych. Seniorzy w przestrzeni nowych mediów*. Warsaw: Difin.

Uniwersytety trzeciego wieku w Polsce w 2018 r. Warsaw: Statistics Poland.

Implementing an evaluation culture among stakeholders of the Integrated Qualifications System

DOI: 10.47050/66515246.50-60

Tomasz Kasprzak
Paweł Szymborski

Abstract

The paper[1] presents the practice of implementing an evaluation culture among stakeholders of the Integrated Qualifications System (IQS), mainly among external quality assurance entities and ministerial officers. The statutory duty of such bodies is to monitor and conduct external evaluation of the qualifications awarded by certifying institutions. The role of external evaluation is to ensure that qualifications are awarded correctly and in the spirit of improving the process. The lack of an evaluation culture or its unusual understanding may be a serious problem that impacts on the outcomes of independent evaluation carried out by external quality assurance bodies even post-support. The analyses undertaken by the authors of this paper were based on experience in promoting an evaluation culture and on data they had collected in the field of implementing the IQS[2]. Subsequently, the analyses were examined in reference to the external evaluation culture which exists in formal education.

Tomasz Kasprzak, Paweł Szymborski
Sociologists and staff members of the Educational Research Institute. They support external quality assurance bodies in their preparations for and in fulfilling their role. The Institute conducts interdisciplinary studies on the functioning and effectiveness of the education system in Poland. It supports the implementation of the Act on the Integrated Qualifications System, including its management and development.

Introduction: main assumptions of the Integrated Qualifications System

One of the goals of the Integrated Qualifications System (IQS)[3] is to shorten the educational path that an individual must complete to find employment, whilst providing consistent quality assurance principles for different systems of attesting staff competence. The IQS was introduced

KEYWORDS:
EVALUATION CULTURE, LIFELONG LEARNING, INTEGRATED QUALIFICATIONS SYSTEM, QUALITY ASSURANCE, PDCA/PDSA CYCLE, STAKEHOLDER SUPPORT

1 The text is based on internal analyses conducted by staff members of the Educational Research Institute who are responsible for supporting external quality assurance institutions. The authors of this paper present the findings of the analyses.

2 The research material, which illustrates the evaluation potential and experience of institutions that act as external quality assurance bodies, comes from, among others, surveys commissioned by the Educational Research Institute and working materials summarising individual support activities.

3 For more information about the IQS, its assumptions and definitions, see: Sławiński, 2017.

by the Act of 2015 on the Integrated Qualifications System (Act on the IQS)[4]. The intervention combines existing solutions concerning qualifications earned in the course of formal (school or higher) and non-formal education (e.g. training courses) with the possibility of validating competences acquired on an informal path (e.g. on-the-job training, i.e. so-called "market qualifications")[5]. The most important element of the IQS, which makes it possible to verify competences regardless of how they were acquired, is the validation process[6]. If the conclusion of this process is positive, the person in question receives a certificate. Both the validation and certification processes lead to the award of a qualification. However, while formalised methods of verifying acquired competences exist in formal and non-formal education[7], as far as market qualifications are concerned it is necessary to think about and develop such a process from scratch.

The method of integrating a market qualification into the IQS is strictly determined by the provisions of the Act on the IQS[8]. In the process of integrating a qualification proposed by a market entity, the competent minister[9] determines and verifies the learning outcomes that an applicant should present to obtain a specific qualification. The minister also specifies the staffing and organisational settings required from an entity interested in validating and certifying the qualification, i.e. interested in becoming an awarding body (AB[10]). Apart from awarding qualifications, an important responsibility of a AB is quality assurance and its improvement, for example, by monitoring and conducting internal evaluation[11].

Once a qualification is integrated into the IQS, it becomes a public asset and its contents become accessible to the public. The qualification is supervised by a competent minister who examines applications for the authorisation to certify qualifications. The whole process, which ultimately concerns the person who acquires a qualification, ends with the signing of an agreement between the competent minister and an external quality assurance entity (EQAE)[12]. An EQAE is an external institution which monitors if an AB meets the staffing and organisational requirements for awarding a specific qualification. It also evaluates the validation and certification processes, as well as monitors and evaluates the AB's internal quality assurance system (see: Figure 1).

The IQS provides for using different methods of attesting the same qualifications, hence quality assurance processes may differ depending on the institution that verifies the holding of specific competences. The requirements towards the institutions in the system presuppose that they employ

4 Act of 22 December 2015 on the Integrated Qualifications System, Journal of Laws of 2018, item 2153.

5 A qualification is *a set of learning outcomes in the scope of knowledge, skills and social competences acquired in the course of formal education, non-formal education or informal learning, which complies with the requirements of the particular qualification, the achievement of which has been validated and formally certified by a qualified certification body* (Act on the IQS, par. 2(8)). The learning outcomes and requirements are part of the "qualification's description". The term "market qualification" refers to qualifications which have been described and introduced into the system by entities representing the market in its broad sense.

6 Validation is *he examination of whether an applicant for a particular qualification, regardless of their method of learning, has achieved a specified part or all of the learning outcomes required for the qualification* (Act on the IQS, par. 2(22)).

7 These include standards, legal regulations and international guidelines.

8 For more information on integrating market qualifications, see: Sławiński et al., 2017; Ziewiec-Skokowska et al., 2016.

9 A competent minister is a minister in charge of a public administration department responsible for the developing or functioning of a specific qualification.

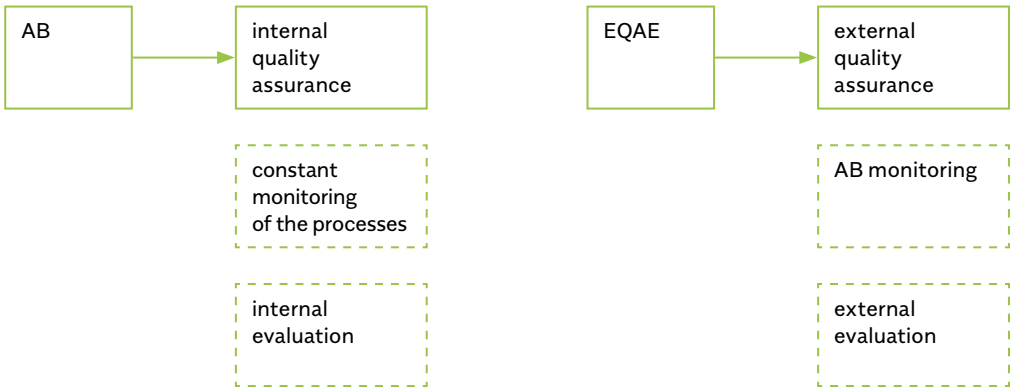
10 An AB is an entity that is authorised to award specific qualifications integrated in the IQS. A business entity can become an AB if it meets the staffing and organisational requirements to conduct a validation process, as defined in a qualification's description included in a notice on integrating a qualification into the IQS (Act on the IQS, Art. 41).

11 Pursuant to Art. 63 of the Act on the IQS, Certifying Institutions are required to develop and use an internal quality assurance system. The system comprises solutions to ensure proper validation and certification, as well as their improvement. In particular, this includes separating education and training processes from validation, constant monitoring and ongoing evaluation of validation and certification, and internal evaluation of validation and certification.

12 An institution which aspires to become an EQAE must meet specific prerequisites (Act on the IQS, Art. 53) which are verified during a selection procedure for entry on the EQAE list (Act on the IQS, Art. 52).

experts in methodology and in analysing data concerning the complex processes of validating competences. Based on existing experience in implementing the system, we should accept that these requirements constitute the target model. The preparedness of ABs and EQAEs for determining the scope of evaluation (i.e. defining the evaluation criteria and asking appropriate questions), for conducting evaluation, and for drawing conclusions that serve to improve the process of awarding qualifications requires support from the Educational Research Institute (IBE)[13]. We should also note that this does not only include support in the scope of planning and conducting the process, but also in building an evaluation culture, as opposed to a culture of control and supervision that is common in many institutions.

Figure 1. Internal and external quality assurance of validation and certification in the IQS



AB – awarding body.

EQAE – external quality assurance entity.

Source: authors' own work based on the Act of 2015 on IQS, par. 50(1–2).

Evaluation in other systems. Point of reference for the IQS

From the perspective of building an evaluation system for quality assurance, of particular interest are considerations on implementing a system of external and internal evaluation in education presented in a series of texts published by the Jagiellonian University (Mazurkiewicz, 2012; Mazurkiewicz & Gocłowska, 2014a; Mazurkiewicz & Gocłowska, 2014b)[14]. Moreover, between 2009 and 2015 the IBE commissioned a study on the implementation of a system for evaluating education. The authors of the study emphasised that external evaluation is conducive for reflection. This, however, requires proper management of the process and the preparation of its participants[15]. At the same time, it is necessary

13 Educational Research Institute (Instytut Badań Edukacyjnych – IBE) is responsible for supporting the implementation of the IQS. For more information on the role of the Institute and on implementation stages, go to: kwalifikacje.edu.pl/o-projekcie [accessed: 9 December 2019].

14 Attempts have also been made to systematise the discourse on evaluating public interventions in Poland. A good example of this is the academic textbook *Teoria i praktyka ewaluacji interwencji publicznych. Podręcznik akademicki*, which focusses on the evaluation of support programmes (cf. Haber, 2007; Haber & Szałaj, 2010; Haber & Olejniczak, 2014).

15 The preparation of EQAEs to conduct evaluation is discussed later in this chapter.

to constantly monitor the adopted solutions and consult major changes with as many different groups of stakeholders as possible. The evaluation of various solutions requires time, and introducing changes without extensive research can pose a significant risk of disrupting the system (Walczak et al., 2017, pp. 40–41). The constant monitoring of the adopted solutions is consistent with the PDCA/PDSA cycle (Plan-Do-Check-Act/Plan-Do-Study-Act), which is currently being integrated into quality assurance in the IQS (Moen & Norman, 2010)[16].

Based on the example of evaluation in pedagogical supervision, the authors of the cited study present an important challenge, which is the different level of acceptance and understanding of change (in this case the introduction of evaluation and its subsequent implementation) due to a varying level of involvement in such change. Thus, the authors recommend involving as many stakeholders as possible and facilitating their participation in designing the change, in order to give them a stronger understating, sense of ownership and acceptance. At the same time, the authors advise to extend the period of preparing and supporting the stakeholders (Walczak et al., 2017, p. 19).

Evaluation in the IQS and the PDCA/PDSA cycle

When an AB or an EQAE decides to enter the IQS, it accepts the necessity to adapt its institution to a new role. Constructing a quality assurance system which will serve entities that have so far held a different position on the market and which will ensure the proper playing of their roles in the IQS is a challenge that is both unprecedented and difficult. The way the quality assurance system works will, in fact, determine the success of the implementation and subsequent functioning of the IQS as far as market qualifications are concerned.

Evaluation plays and important part in the IQS: as a tool for collecting and analysing data to check if various entities are working towards a desired goal, and as a concept that combines different approaches to quality management (iteration and continuous improvement). The Act on the IQS does not specify how an AB should monitor its processes and conduct internal (self-) evaluation. Following internal evaluation, the AB is required to produce a report, which should include an analysis of documentation on the validation and certification processes, an assessment of the suitability of validation methods in terms of qualification descriptions, and information about the steps taken to improve these processes. The Act on the IQS outlines the direction the AB should follow and the frameworks which it should subsequently fill with good practices. A similar solution is used in formal education, where internal evaluation in schools and other education institutions is required by law, however its scope and form are not precisely defined. It is acknowledged that upon entering the IQS an entity accepts the abovementioned responsibilities, which means, among others, that it accepts the possibility of conducting monitoring activities according to a self-evaluation plan.

The Act on the IQS also formulates certain guidelines for EQAEs, assuming that they have the resources and competences needed to conduct external evaluations[17]. Figure 2 shows the actions that ABs and EQAEs should perform in one cycle of external quality assurance. The Act on the IQS does

16 The concept entails an organised and cyclic development of a service, activity or product. In this paper the letter "S", which stands for "study", appears in the PDCA/PDSA cycle. It is indicative of improving an action through evaluation studies and their effect on subsequent iterations.

17 In comparison to ABs and the terms of their integration into the IQS, EQAEs must meet more specific requirements, e.g. employ staff members with competences needed for monitoring and evaluation, as well as with knowledge of the IQS, of validating procedures and of principles of assuring the quality of various processes. The actual process of integrating an entity into the system is preceded by a selection process. For applicable legal provisions, see: the Act on the IQS, Art. 53–60.

not specify the cyclic approach by name[18], but this method becomes apparent when the individual provisions are organised into a schedule that considers the time limits for particular activities.

Figure 2. Framework schedule for internal and external quality assurance of validation and certification in the IQS

Activity	Year 1	Year 2	Year 3	Year 4	Year 5
EQAE monitors the internal quality assurance system, including:					
quarterly information (provided by the AB)					
AB operations reports					
AB internal evaluation report					
EQAE monitors how the AB meets requirements					
External evaluation and report on external quality assurance (conducted/prepared by EQAE)					

AB – awarding body.

EQAE – external quality assurance entity.

Source: authors' own work based on the Act of 2015 on IQS.

Figure 2 shows that quality assurance is a continuous process. The concept of an endless chain of process improvement is consistent with the Act on the IQS and is based, among others, on the Deming cycle (Moen & Norman, 2010; Dybaś et al., 2014). The PDCA/PDSA cycle is treated as a tool for managing the quality of the process. In literature, this cycle is exemplified by public health service programmes (e.g. Coury et al., 2017).

Within the scope of supporting institutions responsible for assuring the quality of awarded qualifications (ABs and EQAEs directly, and ministries in a supervisory role), the IBE works with the institutions to define in greater detail the methodology of evaluation research and of other analyses which they perform in their role within the IQS.

Evaluation culture in the IQS

In Poland, the first examples of implementing evaluation were found in the education sector (cf. Korporowicz, 1997; Mizerek, 1997), while, according to Karol Olejniczak (2008, p. 80), evaluation culture in public administration began developing in the mid-1990s. It was mainly brought about by the requirements of structural funds, which began flowing into Poland at that time. Olejniczak

18 For example, an AB is required do conduct an internal evaluation "at least every 3 years". Such provisions concern most of the products that both ABs and EQAEs are required to deliver.

stresses that this was an element of a broader, demanding process of implementing output-driven management in the public sector. In 2008, when Olejniczak reached this conclusion, he also hypothesised that in the near future the process of integrating evaluation into national programmes would probably intensify. In other words, through the use of structural funds evaluation would cover national policies and programmes (ibidem, p. 81).

Over a decade later, we may wonder to what extent evaluation culture has been integrated into public administration. The subject of this analysis, the IQS, or the system of education evaluation mentioned earlier, are good examples of programmes and projects which are supported by structural funds and develop (or can develop) an evaluation culture.

In literature, we can come across the concept of organisational culture, where evaluation is used, rather than a description of evaluation culture as such (Dorczak, 2012, pp. 91–113). In this sense, evaluation can be treated as an outcome or component of a given organisational culture. As far as supporting IQS stakeholders in the process of implementing a qualification system is concerned, evaluation culture is tentatively defined in terms of a process. This entails the creation of an environment for analysing activities, drawing conclusions and formulating recommendations, which serve as a basis for decisions made by the actors of the system. Finally, bearing in mind that evaluation culture is strictly linked to organisational culture, it seems reasonable to speak of not one, but many evaluation cultures in the IQS. Because there are many entities in the system that differ in terms of industry, sector, size and business model, as well as objectives and roles within the system, the IQS embraces many organisational cultures and, consequently, many potential evaluation cultures.

EQAE potential and preparedness for evaluation

During their selection procedure, institutions working in the system as EQAEs prove that they meet the requirements specified in the Act on the IQS[19]. However, these institutions differ greatly in terms of type (consulting firms, universities, training institutions, non-governmental organisations, etc.) and size (number of employees and departments). An external survey[20] also revealed other differentiating factors, including how the EQAEs perceive their role in the IQS, their motivation to work in the system and their openness to cooperation. These are soft competences, which are not measurable at the early stage of integration. Having collected opinions on self-evaluation of competences in the scope of detailed aspects of evaluation surveys, the authors of this paper were able to measure the competence-related and organisational potential of staff members delegated to managing external quality assurance. This has helped the authors to improve the support programme addressed to EQAEs. The authors also observed that despite their differences, the EQAEs were open to working with ABs on a partnership basis.

By structuring the internal and external quality assurance system in compliance with the provisions of the Act on the IQS, the system prioritises the quality of awarding qualifications. Based on an analysis of relevant data and the experience of various systems in validating competences, the IBE developed solutions to promote evaluation culture rooted in cooperation between various institutions, in particular

19 For a description of the procedure and additional information about EQAEs, go to: kwalifikacje.edu.pl/podmioty-zewnetrznego-zapewniania-jakosci-pzzj [accessed: 2 December 2019].

20 Centrum Ewaluacji i Analiz Polityk Publicznych Uniwersytetu Jagiellońskiego – CEiAPP UJ (2018–2019). *PZZJ jako podmioty przygotowujące się do pełnienia funkcji – zbiór raportów.*

between ABs and EQAEs. These solutions include information materials, tools to support the work of the institutions, as well as a training and workshop programme.

Strengthening the surveying and organisational potential of EQAEs

As part of building a common evaluation culture, the IBE offers training and workshops for both new EQAEs and those already operating within the IQS. This approach is consistent with the recommendations resulting from the analysis of the previously mentioned system of education evaluation. The authors of the analysis (who were also co-authors of the system) emphasise that *it is necessary to ensure that all reports or other forms of communicating evaluation results serve reflection, the development of knowledge about a given organisation, working together on applications, and planning changes. To make this possible, the teams taking part in evaluation should undergo an appropriate preparation process concentrating on methods of using evaluation data* (Walczak et al., 2017, p. 11).

The offered support targets staff members responsible for EQAE activities in their respective institutions[21]. The series of workshops is modular, which means that they can be organised in an order that suits the needs of a given entity. The areas covered by the workshops include:

- → system-related topics (process of describing and integrating qualifications into the system, validation process);
- → the role and operations of an EQAE (statutory requirements, day-to-day operations and reporting);
- → cooperation (between EQAEs and ABs, and between EQAEs and competent ministries);
- → evaluation and monitoring (survey models and survey tools).

The workshops are run by IBE staff. In addition, EQAEs take part in external workshops organised by commercial companies selected via a tendering procedure. Currently, the following workshops are organised:

1. Managing change[22];
2. SCRUM[23];
3. Design Thinking[24].

A model of external evaluation of a given qualification is developed by IBE and an EQAE. Currently, the first qualification integrated into the IQS, "carpentry assembly in construction", is being evaluated[25]. An IBE team and the EQAE for this qualification (Łukasiewicz Research Network – Institute for Sustainable Technologies) are operationalising the process in a series of workshops, meetings and consultations. The team is also developing specific evaluation criteria (accuracy, reliability and adequacy), as well as evaluation tools and report templates. The goal is to design an evaluation process that matches the qualification, taking into account its actual context, the AB and the recipients of the qualification. Being the first external evaluation in the IQS, it will be used to examine the approach to testing

21 Since the turn of 2018, when the support became available, the majority of entities operating within the IQS have declared their willingness to use it.

22 The training is designed to help teams to adapt to the new conditions and deal with their institution's new role as an EQAE.

23 Learning agile planning and controlling the design work process is intended to help EQAEs effectively manage the process.

24 Design Thinking may be helpful in designing external evaluation tailored to the specific features of a certifying institution.

25 The qualification was integrated into the system and is validated by the VCC Foundation.

the functioning of the internal quality assurance system and to treating external quality assurance as an ongoing cycle. Furthermore, the actual survey questions and methods will be tested.

Activities to date and expanding the field of dialogue

In line with the premises of the PDCA/PDSA cycle, support provided to IQS stakeholders is monitored and evaluated on an on-going basis. An example of this is the evaluation of the selection of potential EQAEs, which will be discussed later in this section.

Two selection processes have been conducted since the Act on the IQS came into effect. The first was run between November 2016 and September 2017, and EQAE status was awarded to 14 institutions in 37 qualification groups. In late 2017, the process was evaluated[26]. Based on conclusions and recommendations, representatives of IBE were able to prepare the second selection run, which began in April 2018. Modifications in relation to the first selection process included:

- → a more precise scope and definition of qualification groups,
- → operationalisation of criteria,
- → communication with applicants,
- → IT system management,
- → organisation of the committee's work.

The selection was conducted for 11 qualification groups and ended with the entry of 8 new entities on the EQAE list[27]. As in the previous case, also this selection process was evaluated[28].

Subsequently, many conclusions and recommendations were made concerning information and promotion activities, the formal and content-related evaluation stage, the work of the committee and voting procedures, as well as cooperation between entities during the selection process. The recommendations are currently analysed as part of the preparation for the next selection procedure. The evaluation shows a higher level of satisfaction among stakeholders thanks to the improvements. Representatives of the ministries involved compared the two selection processes, emphasising the improvement of the procedures, new functionalities and tools to enhance work quality.

Evaluation culture is developed as part of a process of ongoing individual and group reflection, as exemplified by recent workshops attended by representatives of various ministries and EQAEs (October 2019). The workshops focussed on sharing experience in the scope of concluding agreements, defining key challenges and presenting modification proposals. The process also included surveys among the ministries and EQAEs which had already singed such agreements. The conclusions were used as a starting point for in-depth reflection during a workshop on the challenges concerning the actual signing of the agreements and, in a broader sense, on cooperation from the point of view

26 The goal of the qualitative survey was to collect opinions from individual actors of the process, including about the formal and content-related evaluation, the committee's work, voting and entry on the EQAE list by the Ministry of National Education. The following methods were used: desk research, group and individual interviews, including individual in-depth interviews (IDI) and telephone in-depth interviews (TDI). The survey was conducted by P. Stronkowski, M. Szostakowska and A. Szczurek of Ośrodek Ewaluacji and the Idea for Development Foundation (an external entity).

27 In addition, one of the entities entered on the list during the previous selection process reapplied and was added as an EQAE to another qualification group.

28 The goal of the survey was to collect opinions from the individual actors of the process regarding its subsequent stages and, in general, the selection of EQAEs. The survey also focussed on the current situation and options for implementing the recommendations formulated during the evaluation of the first selection procedure (2017). From a methodological perspective, this was also a qualitative survey. In late 2018, an external research agency conducted an evaluation of the process, which included desk research, as well as group and individual interviews with the stakeholders involved.

of both sides. Based on these analyses, the participants proposed changes and formulated recommendations regarding solutions to problems they considered crucial.

Figure 3. Experience in concluding agreements from the point of view of EQAEs

Source: material developed during workshops, based on a presentation and discussion of the survey findings.

The workshops provided material to be used in an in-depth analysis of the process of signing agreements from the participants' point of view. Furthermore, the participants formulated recommendations for changes in the actual process and for cooperation between the EQAEs and the ministries involved. In addition, the workshops demonstrated that involving all parties in the discussion on the current and future situation gave them a stronger stimulus to participate in the system. This corresponds to the analysis of the education evaluation system mentioned earlier, according to which *various forms of involving stakeholders in the process of change determine their level of understanding and approval of the process* (Walczak et al., 2017). Ultimately, according to the findings of the surveys and workshops, the next step is to translate the recommendations into practice.

Conclusions

Similarly to the IQS, which is currently under development, the tasks and roles of the entities involved in the system, outlined in the Act on the IQS, are being clarified and tested in practice. All the stakeholders of the IQS are learning and adjusting their organisational cultures to the system. The quality assurance system, which the IQS aims to provide, is intended as an effective tool for awarding high quality, credible qualifications. As demonstrated in this paper, creating a quality assurance system on such a wide scale is an unprecedented and challenging project, where external and internal evaluation are of key importance. Ultimately, the quality which the system is to assure depends on

the evolving evaluation cultures of individual stakeholders: private individuals and public institutions, as well as non-governmental organisations or businesses, which play different roles in the system.

Bibliography

Act of 22 December 2015 on the Integrated Qualifications System. Journal of Laws of 2018.

Coury, J. et al. (2017). *Applying the Plan-Do-Study-Act (PDSA) approach to a large pragmatic study involving safety net clinics*. BMC Health Serv 2017. DOI: 10.1186/s12913-017-2364-3

Dorczak, R. (2012). Znaczenie kultury organizacyjnej szkoły w procesie wprowadzania systemu ewaluacji oświaty. In: G. Mazurkiewicz (ed.), *Jak być jeszcze lepszym? Ewaluacja w edukacji*. Krakow: Wydawnictwo Uniwersytetu Jagiellońskiego.

Dybaś, M. (2014). *Zapewnianie jakości kwalifikacji w zintegrowanym systemie kwalifikacji*. Warsaw: Educational Research Institute.

Haber, A. (2007). *Ewaluacja ex-post. Teoria i praktyka Badawcza*. Warsaw: Polska Agencja Rozwoju Przedsiębiorczości.

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

Haber, A. & Olejniczak, K. (2014). *R(ewaluacja) 2. Wiedza w działaniu*. Warsaw: Polska Agencja Rozwoju Przedsiębiorczości.

Korporowicz, L. (1997). *Ewaluacja w edukacji*. Warsaw: Wydawnictwo Oficyna Naukowa.

Mazurkiewicz, G. (2012). *Jakość edukacji. Różnorodne perspektywy*. Krakow: Wydawnictwo Uniwersytetu Jagiellońskiego.

Mazurkiewicz, G. & Gocłowska, A. (2014a). *Jakość edukacji – dane i wnioski z ewaluacji zewnętrznych przeprowadzonych w latach 2013–2014*. Krakow: Wydawnictwo Uniwersytetu Jagiellońskiego.

Mazurkiewicz, G. & Gocłowska, A. (2014b). *Jakość edukacji. Dane i wnioski z ewaluacji zewnętrznych przeprowadzonych w roku 2014*. Krakow: Wydawnictwo Uniwersytetu Jagiellońskiego.

Mizerek, H. (1997). *Ewaluacja w szkole. Wybór tekstów*. Olsztyn: Wydawnictwo MG.

Moen, R.D. & Norman, C.L. (2010). Clearing up myths about the Deming cycle and seeing how it keeps evolving, *Quality Progress*, November.

Niedzielska, E. & Szulżycka, A. (2018). *Ewaluacja ex-post dotycząca naboru PZZJ w systemie ZSK zakończonego w czerwcu 2018*. Warsaw: Infondo, commissioned by Educational Research Institute.

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

Olejniczak, K. (2012). *Organizacje uczące się. Model dla administracji publicznej*. Warsaw: Wydawnictwo Naukowe Scholar.

Olejniczak, K. et al. (2008). *Teoria i praktyka ewaluacji interwencji publicznych. Podręcznik akademicki*. Warsaw: Wydawnictwa Akademickie i Profesjonalne.

Rozporządzenie Ministra Edukacji Narodowej z dnia 10 października 2016 r. w sprawie regulaminu pracy komisji do oceny wniosków o wpis na listę podmiotów uprawnionych do pełnienia funkcji zewnętrznego zapewniania jakości, wzoru umowy z podmiotem, któremu powierzono funkcję zewnętrznego zapewniania jakości, oraz sposobu ustalania wysokości wynagrodzenia z tytułu tej umowy, Journal of Laws of 2016, item 1687.

Sławiński, S. (2017). *Mała Encyklopedia ZSK*. Warsaw: Educational Research Institute.

Sławiński, S. et al. (2017). *Włączanie kwalifikacji do zintegrowanego systemu kwalifikacji*. Warsaw: Educational Research Institute.

Stronkowski, P. & Szostakowska, M. (2017). *Ewaluacja ex-post dotycząca naboru Podmiotów Zewnętrznego Zapewniania Jakości w systemie ZSK zakończonego w czerwcu 2017. Raport końcowy*. Warsaw: Ośrodek Ewaluacji Idea, commissioned by Educational Research Institute.

Walczak, B. et al. (2017). *Analiza doświadczeń we wdrażaniu systemu ewaluacji w ramach nadzoru pedagogicznego*. Warsaw: Educational Research Institute.

Ziewiec-Skokowska, G. et al. (2016). *Opisywanie kwalifikacji nadawanych poza systemami oświaty i szkolnictwa wyższego. Poradnik*. Warsaw: Educational Research Institute.

Developmental capacities of an educational institution

DOI: 10.47050/66515246.62-86

Mirosław Warowicki

Abstract

The first part of the chapter[1] presents a summary of the discussion on the theoretical basis of the changing role of educational and training institutions. The author discusses the traditional and proactive role of an educational institution in developing individual, organisational and institutional capacities and, consequently, in improving the quality of education services. Playing the role of an active agent of change by an educational institution requires adequate preparation, from capacity assessment in several areas to planning a development scenario. The second part of the chapter is a step-by-step presentation of the application of the assessment methodology and capacity development. The author proves that the presented methodology allows for diagnosing and developing the potential of an educational institution. This approach also gives a perspective for the discussion on the application of qualitative and quantitative methods in social research, including in education, and in the assessment and evaluation of the processes related to education and development presented in the article.

Developmental trends

The topic of introducing changes in the process of education[2] in its broadest sense is presented here on the basis of the discussion regarding trends and the effectiveness of international development aid held over the past 20 years. It is not only about popularising the idea of an educational institution as a learning organisation (Sange et al., 2002), but also about research and analyses on sustainable development. The essence of the latter concept is to take into account three key dimensions – social,

1 This article is a revised version of Mirosław Warowicki's unpublished report for UNHABITAT, Nairobi 2012.

2 Following the adoption of the Millennium Development Goals by the UN General Assembly, the publication of the Paris Declaration, the Accra Agenda for Action, the reports of the OECD/DAC (Pearson, 2011) and other documents, it is a common view that the limited effectiveness of international aid that has been observed so far calls for a shift towards greater use of local developmental resources and putting greater emphasis on the role of teaching/training and learning in cooperation with less developed countries.

Mirosław Warowicki
PhD in social sciences (social psychology) of the Warsaw University. He has many years of academic experience (University of Warsaw, University of Warmia and Mazury, and Olsztyn School of Management) and in carrying out training, research, implementation and evaluation projects for the public, private and non-governmental sector in Poland and abroad (the Balkans, CEE, Eurasia and Africa). He was a Member of the Management Board of the European Network of Training Organizations operating alongside the Council of Europe. Since 2018, he has been a Member of the Management Board of the Polish Evaluation Society. He is the author and co-author of publications on social sciences.

KEYWORDS:
DEVELOPMENTAL RESOURCES, METHODOLOGY FOR ASSESSING AND DEVELOPING RESOURCES, INTERACTIVE RELATIONSHIPS, QUALITATIVE AND QUANTITATIVE METHODS, EVALUATION FROM TEACHING TO LEARNING

economic and environmental – from the very beginning. The reference model for this reflection is a triangle-based pyramid which includes:

- → the social dimension, i.e. broadly understood culture, social participation, and distribution of goods and services;
- → the economic dimension related to the mechanisms of growth, effectiveness and economic stability; and
- → the environmental dimension understood as the preservation of biodiversity of ecosystems, the availability of natural sources and resources, and the quality of life in a given environment (Mondini & Vale, 2007; Goodland, 1994).

The concept of social capital is based on similar principles. It places emphasis on interpersonal relations combined with the socio-economic sphere (Edwards et al., 2001; Gajowiak, 2012; Sztompka, 2016).

Many studies and observations confirm the importance of the relationship between education systems and the multidimensional complexity of socio-economic life, on both the global and local scale. This complexity cannot be reduced to a closed list of causative mechanisms and indicators, due to, among others, the interactive nature of interdependencies (Czapiński, 2004; Harré, 1980; Warowicki, 1985)[3]. On the other hand – and this is important for the success of educational projects – more and more arguments speak for the fact that the development of local resources and potential forms the basis for social capital and problem-solving (Lewenstein & Thesis, 2008; Pearson, 2011b; Kazimierczak, 2014; Gliński, 2001; Czyżewska, 2001, 2012; Warowicki, 2001, 2013a, 2013b).
The resources which are the basis for, among others, the development of social capital, are broadly understood and relate to several spheres: social, economic, educational, functional, spatial, technical and infrastructural. In line with the abovementioned reflection trends, the success of the introduction of various education models depends on the local resource's stage of development (understood as capacity development), which determines the scale of initiatives and the effectiveness of actions taken to solve problems within educational institutions[4].

Two roles of educational institutions

The discussion on the traditional role of an educational institution and the growing demand for its new, proactive role, leads to two conclusions (Pearson, 2011b). Firstly, capacity development helps to provide high quality training services. And secondly, it makes it easier for an educational institution to play the role of a facilitator of the learning process and to introduce organisational, institutional and systemic changes[5]. The two roles are interlinked and equally important for the effective strengthening of an educational institution.

3 Transformation in Poland is a good example of this. After the 1989 political transformation and Poland's accession to the European Union in 2004, the country's development has visibly changed. Several economic development indicators and these pertaining to the quality of life – including the quality of life indicators adopted in UNDP reports and in *Diagnoza Społeczna* (Czapiński & Panek, 2008, 2011, 2013, 2015) – confirmed consistent growth in Poland and the country's reaching higher positions in international comparisons. However, much of the transition period is characterised by an increase in income inequality and poverty. There are also differing opinions about the education reform, which raises many doubts. This means that the assessment of whether the situation of Poles improved over this period is ambivalent. In late 2019, Poland, with its good economic situation and low unemployment, became one of the countries in Europe which stood out in international comparisons due to great income inequality (according to World Inequality Lab calculations, Poland has the greatest income inequality in Europe and the highest change dynamic in this area over the last 15 years, cf. Blanchet et al., 2019).

4 It is assumed in this article that an educational institution is an organisation or company founded to transfer or develop knowledge and skills. In this respect, schools, higher education institutions (public and private ones) and training centres are considered educational institutions.

5 The terms "system" and "systemic" refer to a given whole (e.g. an institution) composed of elements that interact with one another and are influenced by external factors. The systemic approach to human behaviour in a given organisation consists in the inclusion of such elements as structure, communication (with emphasis on the importance of feedback), interaction profiles between employees, objectives, values, organisational culture, and a certain level of autonomy and adaptability of individual units, departments and the organisation as a whole (cf. Anderson & Carter, 1984.)

An educational institution plays a major role not only in the transfer of knowledge or competencies, but also in starting new undertakings leading to the development of resources by both the recipients of educational services and the educational institution itself. It can therefore be described as an active agent of change. The effectiveness of initiating and implementing new solutions depends on the status of a given institution, on its interactions with other institutional actors and on its amount of agency in the discussion on new ideas and changes with representatives of other institutions. Change agents generally look for opportunities to create new, strong links with other organisations and to modify their relations with key actors in order to make them more useful for their operations.

Facilitation starts with reaching an understanding by the change agents. The quality of relations and subjectivity are preconditions for exerting influence on others, introducing new ideas to the discussion and considering them with a wider circle of players. Agency is crucial in any process of introducing changes, but it is especially important when considering it at a local level. Change agents should look at themselves as important components of resource development, hence strengthening an educational institution also forms part of this process.

Capacity development – from theory to practice

This notion is understood in subject literature as developing the capacity to perform a task or a specific function. Therefore, capacity development is defined as:

- → the ability to carry out activities (in a better way);
- → the process and form of introducing changes that focus on raising organisational standards, performance, level of service provision and/or introducing new learning methods;
- → an attribute of individuals, organisations and groups of organisations/entire institutions and the outside world;
- → a combination of individual competencies, team skills, resources, and internal and external relations (organisational culture and quality of relations with stakeholders) that will make it possible to carry out the learning process and provide better quality education services.

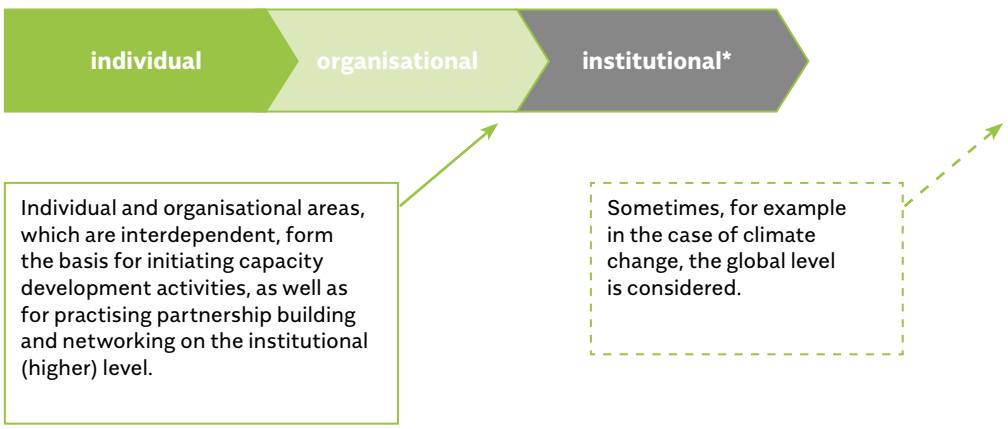
As can be seen from the above considerations, capacity development covers a wider range of concepts than growth of individuals' knowledge and skills. Although the term capacity building is often used interchangeably with capacity development, it is worth pointing out the difference between the two. Capacity development results from a given organisational structure and external support. In other words, it consists in an entity's growth which cannot be attributed to external influence alone. On the other hand, capacity building is a result of external support (e.g. the involvement of experts from outside the institution in order to introduce a new teaching method) and concerns mainly the initial stage – assuming the institution did not have any resources.

A more effective use of resources can consist in searching for an adequate methodology (in order to adapt it, among others, to available training materials) or in promoting cooperation between educational institutions. Therefore, apart from broadening knowledge, capacity development concerns the relations between stakeholders, decision-making hubs and the environment, including the institutional environment, in which a given educational institution operates.

Capacity development in several areas

Capacity development refers to several areas, including social, economic, legal, functional and spatial, environmental and technical. The figure below presents the sequence of analysing capacity development.

Figure 1. Capacity development analysis

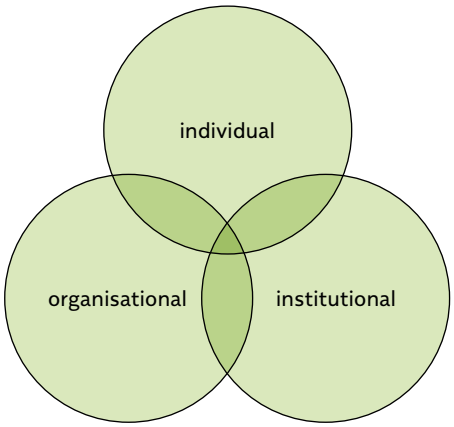


* The institutional level concerns legal regulations, public policies, as well as the institutional and, more broadly, social environment. As a whole, this level can create an enabling environment.

Source: author's own work after Pearson, 2011a, p. 2.

To the same extent, capacity development concerns the introduction of creative curricula, unconventional relationships between the teacher and the learner, new organisational solutions, social and economic activity in the field of education, as well as the development of relations between stakeholders, local authorities and the broader environment. Capacity development naturally concerns broadening knowledge and raising competencies at the individual, organisational and institutional level. The figure below shows the three dimensions of capacities that underpin development. These dimensions are, to a large extent, interconnected. The development of knowledge, competencies and relations at each level affects the other two levels.

Figure 2. Capacity development – three dimensions



Source: author's own work based on Taylor & Clarke, 2008[6].

In accordance with the concept of capacity development, the strategy for taking action within the education sector goes beyond the linear interdependencies of causative mechanisms and focusses on systemic actions aimed at strengthening and developing local resources. An educational institution's capacities are summarised in the table below. A more extensive description can include additional categories, such as internal and external communication, management (including financial and human resource management – HRM), adaptation of the organisation's structure to adopted objectives and functions, adaptation of the organisation to market trends, relations with the external environment, etc.[7]

Table 1. Examples of individual, organisational and institutional capacities

Human resources of an educational institution
→ Teachers and trainers: they have the right level of knowledge in many areas, such as planning, conducting and evaluating training/training programmes in many subject-areas; → Administrative staff: they have professional skills required to perform administrative tasks, and use modern information and communication technologies (ICT) with the support of IT services; → Managerial staff: they professionally manage an educational institution, set strategic directions, introduce elements of coaching, support employees and co-workers, encourage and foster open communication, skilfully engage employees, face difficult problems and analyse them from several perspectives, and encourage innovativeness and new ideas.

6 Peter Taylor and Peter Clarke (2008) adopted three dimensions of capacity development: individual, organisational and social. This article assumes after Jenny Pearson (2011) that the social dimension is an extension of the institutional one.

7 In *Designing Effective Organizations*, Michael Goold and Andrew Campbell (2002) proposed a conceptual framework consisting of nine categories, which makes it possible to verify an organisation's structure adaptation to, among others, the external market, adopted goals and strategies, and the use of internal resources. This framework also makes it possible to verify the extent to which an organisation and its structure comply with modern management standards and to which the organisation itself is capable of adapting and innovating.

Capacities of an educational institution
→ It runs various education and training programmes; → It is able to simultaneously and quickly introduce several new training programmes or educational services; → It has an organisational structure that provides good conditions for teamwork and cooperation between departments; → It boasts a flexible organisational structure, which facilitates learning how the organisation or its departments operate(s) and enables an adequate response to changes and trends observed in the external environment.

Relationships as resources forming an enabling environment
→ Legal regulations – policy instruments enabling the development of formal and informal links, cooperation and partnerships between various public and private entities, institutions and organisations operating in the education and training sector; → Networking between institutions and organisations (cooperation with other organisations consisting in information sharing, provision of complementary services, carrying out joint projects, interaction and joint exertion of influence on the environment); → Experience in involving stakeholders in addressing difficult problems and challenges related to internal resources and relations with the external environment, including new trends and changes; → Cooperation at the local and regional level – local government and other public institutions and stakeholders actively support educational institutions through their policy actions.

Source: author's own work.

An enabling institutional environment strengthens educational institutions and provides them with systemic incentives (*The Challenge of Capacity Development*, 2006). The term systemic in this context refers not only to interactions between the three dimensions of capacities, but also to the role and importance of knowledge in relation to development, learning and change. This is crucial for developing resources and strengthening training institutions (Taylor & Clarke, 2008).

The more you get into it, the more complicated it becomes

In each of the three capacity areas, dimensions are defined to describe the level of their development, e.g. on a scale or descriptively (*Measuring Capacity*, 2010; Otoo et al., 2009; Wignaraja, 2010). Capacities can be found in the following areas:

→ participants in education and training programmes: their individual knowledge;
→ teams: capacities allowing them to perform typical tasks;
→ educational institutions: capacities enabling them to provide education and training programmes.

Moreover, training institutions and organisations form part of relationship resources, which create an enabling environment.

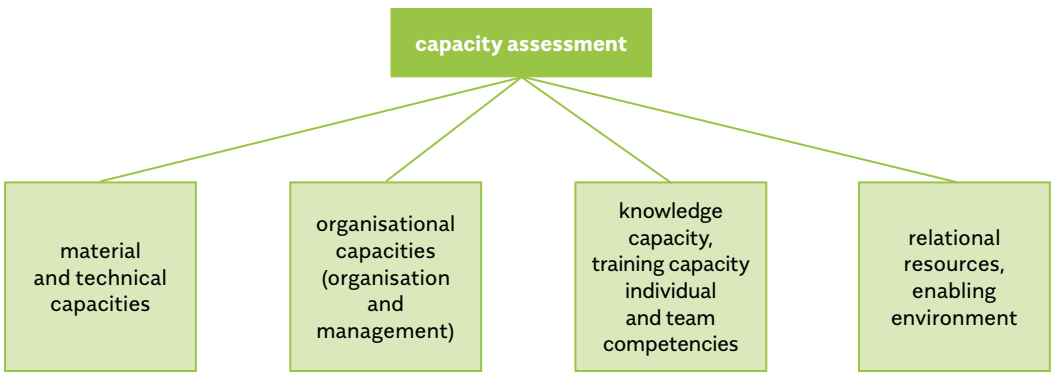
> *The data collected in many countries, which form country case studies, show that capacity development is still tantamount to human resource management. In recent years, experts and practitioners have gradually become more aware of the fact that sustainable growth of individual knowledge and competencies to a large extent also depends on the quality and level of educational institutions' organisation and the extent to which the institutional environment and the wider context in which an educational institution operates are enabling* (Supporting Capacity Development..., 2011, p. 13).

Increased performance and effectiveness of an educational institution or organisation depends on the development of human and financial resources, on the availability of modern technology,

and on management standards, organisational culture and relations with the external environment. What also matters is the extent to which the environment is conducive to the development of a given educational institution.

When defining the resources of an educational institution, the question arises whether the capacities described in the three dimensions have been accurately identified and diagnosed and whether they enable the institution to provide educational services at a sufficiently high level (Walters, 2008). In practice, the assessment of capacities includes, in addition to these three dimensions, the assessment of material and technical capacities. Thus we are dealing with a description of four areas on scales[8] (the description uses interchangeable dimensions). They are presented in Figure 3.

Figure 3. Assessment of capacities – four categories



Source: Warowicki, 2012.

The four categories of capacities to a large extent overlap with the already mentioned three dimensions and include: material (buildings, infrastructure), organisational (organisational structure, management), substantive (quality of programmes and training) and relational (enabling environment) capacities.

Competencies are the outcome of knowledge, skills, abilities, motivation and attitudes, and – in a broader sense – of human resource management. The competencies of the staff include, in addition to knowledge of the training's or education programme's subject-matter, knowledge of the specific character of public (and/or private) educational institutions. These capacities also comprise knowledge of and experience in planning, conducting and evaluating educational undertakings, as well as carrying out follow-up after the training.

The strategic and operational level of educational institution management is, in turn, a consequence of how employees and co-workers are focussed on tasks and to what extent they can work together to provide educational services of the highest quality. The organisational capacity of an institution is its readiness to pro-actively react to internal and external changes.

8 The dimensions are used interchangeably in the description, because the description on scales is possible when we take the dimensions of capacities (in the description of individual areas) and not the types of resources/capacities. The division into types of resources/capacities is relevant as long as it is not a matter of indicating the level of development of particular capacities.

The organisational level affects the behaviour and attitudes of employees (including teachers, lecturers and trainers) and training participants. This concerns the way in which knowledge, skills and competencies are acquired. Information and knowledge transfer is only part of the teaching and learning process. Participants in training programmes must be motivated to acquire, test and modify their knowledge and skills. This requires constant readiness to expand them. In turn, if it is enabling, the external environment provides incentives and reinforcements for continuing education and competency development.

In many respects, capacity development in an educational institution is the same as in other organisations. The same applies to the assessment of the institution's capacities. The difference is reflected in:

- → the specific character of competencies and skills of staff members as teachers, trainers, and facilitators (individual dimension);
- → specific capacities of an educational institution as an organisation (organisational dimension); and
- → relations with the external and social environment (systemic dimension, enabling environment dimension)[9].

A key success factor in capacity development is the systemic integration of these dimensions, and the integration of the political, social, economic and cultural context. Measures aimed at capacity development should take into account, on the one hand, the importance of creating individual capacities, and, on the other, that of developing favourable relations with internal and external stakeholders. Strengthening relational resources can be achieved by improving communication, which facilitates dialogue and partnership building (Ubels et al., 2010).

Apart from typical operations of an educational institution, it is worth paying attention to networking, facilitation and building strategic partnerships, also with local authorities. By taking up new challenges related to internal and external stakeholders, a new kind of capacity is created within the training institution. An enabling environment affects the behaviour of the organisation and its members (employees) through incentives and positive reinforcement (*The Challenge of Capacity Development*, 2006).

9 Peter Morgan (2006) proposed an interpretation of the link between competencies, capabilities and capacities, treating the former two as components of the latter. From this perspective, competencies are specific abilities of an individual, and capabilities are specific abilities of an organisation in a systemic (or sub-systemic) sense.

The table below outlines the development of individual, organisational and institutional capacities taking into consideration the three stages of institution-building.

Table 2. Capacity development of an institution or organisation

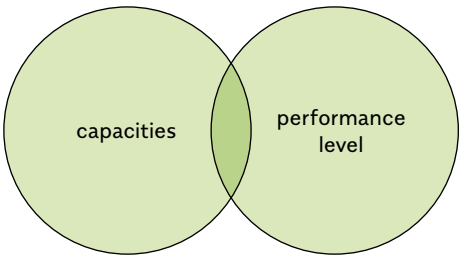
Capacity development	Creation	Use	Maintenance
Individual	development of appropriate skills, knowledge, competencies and attitudes	application and use of acquired skills, knowledge and competencies in the workplace	reduction of staff turnover, dissemination of knowledge and skills within the institution
Organisational	establishment of effective structures, processes and procedures	integration of structures, processes and procedures in the daily work of the institution	consistent adaptation of structures, processes and procedures
Institutional and environmental	establishment of adequate institutions and relevant laws and regulations	enforcement of laws and regulations concerning good governance	consistent adaptation of institutions, laws and regulations

Source: author's own work based on Pearson, 2011a, p. 3.

Capacities and performance levels

Educational capacity development in the traditional sense refers to technical and content-related aspects of linking competences to the level of performance or, in other words, to the level of service provision by an educational institution.

Figure 4. Linking capacities to performance levels



Source: author's own work.

The link between resources and performance levels can be illustrated by the following example. In general, trainers strive to:

1. develop training and educational programmes that respond to the needs of prospective participants;
2. prepare training materials containing up-to-date knowledge, supplemented by good practices, case studies and handouts;
3. plan the use of specific training methods to facilitate the exchange of opinions and experiences by participants;

4. ensure that training results are transferred by the participants to specific workplaces and applied;
5. use ready-made evaluation sheets upon the completion of the training.

These elements constitute educational capacities, each is an important part of the training, and together they determine its quality and adopted quality standards. In this respect, the availability and usability of these capacities is crucial for quality, which is also determined by an appropriate needs analysis, proper design, preparation and realisation of the training, monitoring and evaluation, as well as follow-up, i.e. complementary support activities after the training. Quality depends on the relevance and usefulness of training, monitoring and evaluation, and transfer of results.

Stakeholder involvement

Involving stakeholders in projects is a widely accepted element of project implementation and evaluation. Stakeholders are individuals, groups or organisations which are or can be affected by the activities and impact of an educational institution. They are also the entities to which a given educational institution addresses its products or services. Yet, at the same time, they are entities that can themselves influence the educational institution. Before making a resource assessment, it is worth identifying individuals, groups, organisational structures and entire institutions that can support the activities of an educational institution through involvement in the entire process. The degree of stakeholder involvement can vary and change over time. It can start with simple two-way communication and take more advanced forms, such as consultations on relevant issues, joint decisions and support actions. Communication and consultations facilitate the identification and analysis of common problems. They form the basis for understanding and taking action in the area of individual stakeholders' responsibility (e.g. supervisory authorities). As research proves (Pearson, 2011b, 2012), the involvement of stakeholders makes it possible to provide services at a higher level if, as a result of stakeholders' actions, project implementers modify their intentions. This also applies to projects aimed at capacity building, especially these focussing on the enabling environment.

To sum up, involving stakeholders in the assessment of educational institution capacities is important for two reasons. Firstly, obtaining feedback significantly increases the chances of a project's or programme's success. Secondly, it builds accountability and confidence in the results of capacity assessment.

Below are some examples of varying degrees of stakeholder involvement:

- → involvement in addressing new challenges related to internal resources and the external environment from the perspective of the latest trends in management and legal and organisational solutions;
- → taking action to maintain and use strategic partnerships and links from the perspective of dialogue with local or central government;
- → building strong stakeholder coalitions that contribute to effective implementation of planned changes; creating a bottom-up impact on assuming institutional responsibility by external actors, creating and developing advocacy, creating lobbies and offering training to media and government representatives;
- → offering support to educational institutions through policies and actions (local government, public institutions and stakeholders share new information and challenges related to reforms and new legal regulations).

The conditions for the success of institutional change – based on learning – include an accurate assessment of capacities and cooperation of stakeholders with educational institutions.

Participatory self-assessment

Self-assessment means that the employees of an educational institution carry out an assessment of resources for, and to a large extent by, themselves (a facilitator helps them) using special diagnostic methods, e.g. a question grid, answer scales, unfinished statements or – in the drama approach – not fully defined role descriptions. Such a process encourages employees to become involved, gives them the feeling that they have developed the results of the assessment themselves, and inspires them to take up challenges faced by the educational institution.

The assumption of the participatory approach is the creation of a common internal perspective for the analysis of the state of affairs, behaviours, attitudes and competencies. The experience of the participants and their knowledge are very valuable in capacity assessment. Therefore, it is worth focussing on encouraging them to comment and present opinions after each attempt to describe and assess a selected aspect. It will also be an added value to share examples, stories, events and anecdotes.

Assessment in the presented approach consists in individual and team self-assessment. In addition, where possible with the involvement of all parties, stakeholders collect information and evaluate the current state of a selected area of education, as well as the degree of implementation and results of existing projects. The aim is to create a precise picture of the capacity level in a given context. The cyclic diagnosis created in this way helps process participants to appraise the initial state and describe the target state, and the path to achieve the assumed results can be translated into an action plan.

The process of self-assessment, which is discussed in more detail below, is guided by an external or internal facilitator. During the workshop, a series of interviews or other forms of cooperation they provide guidance and participate in the selection of process participants and in defining areas of operation of an educational institution.

Capacity assessment / capacity development methodology

The presented capacity assessment / capacity development methodology (CA/CD) is an adaptation based on a combination of three approaches: the Capacity Assessment Framework (*Measuring Capacity*, 2010), the Mc Kinsey Capacity Framework (2001) and the Analytical Framework (EPCDM, 2008). The model, developed for the education sector, assumes that capacity assessment must take into account objectives, functions, training and teaching methodology, organisational culture and the language of communication. The starting point is the availability of material, human and financial resources, educational infrastructure, technology and communication. The analysis additionally covers organisational resources and management of an educational institution, as well as the sphere of educational competencies, including the thematic scope of training and education and the teaching methodology used. In turn, the category related to the environment (how enabling it is) refers to legal regulations, policy actions of educational institutions and, more broadly, public policies, networking and relational resources (in the form of strengthening or weakening links with other institutions and organisations, adaptation to change and learning of the organisation, and stakeholder involvement).

It is worth noting here the difference between models applied in social sciences (including education) and these used in science. In traditional physics models and these describing technical devices, it is possible to precisely reproduce real relationships. In social sciences, education, economics and finance, where human behaviour plays an important role, it is very difficult to achieve this effect. The behaviour of people, including decision-makers, should be taken into account in the developed models (cf. Cichocki, 2016; Koopmans, 1957).

After certain modifications, the CA/CD methodology can be used to describe and diagnose educational, production and service institutions. In the process of making the diagnosis and planning the activities, participants' knowledge and their accounts are used[10]. The active involvement of participants allows for taking a closer look at various aspects of an entity's functioning at the individual, organisational and institutional level, taking into account its enabling environment. The CA/CD methodology makes it possible to verify the state of progress of the existing strategy of an educational organisation/institution and its implementation. The key stage is taking action – from diagnosis to change – to increase effectiveness and efficiency, expand the service market and modify the organisational culture. In a selected, narrow area, these activities can concern e.g. individual dispositions, learning management, adopted learning models, internal and external communication and/or relations with the environment.

CA Assumptions

The CA (capacity assessment) methodology facilitates the acquisition of new knowledge about the organisation thanks to involving representatives of different levels of the educational organisation/institution in the diagnostic process. At a lower level, i.e. management and functioning of the institution, it makes it possible to see how self-diagnosis changes the perception of the institution/organisation (its mission, goals, and internal and external communication), as well as relations with the environment.

In order to better understand the current and potential state of development of an educational institution, representatives of various departments, including managerial staff, and representatives of stakeholders need to take part in all stages of the process. Depending on the size of the educational organisation/institution, participants in e.g. a workshop (it is one of several possible solutions) can be all the employees and managerial staff, lecturers, teachers and trainers. For larger educational organisations/institutions the number of participants can be limited to key persons.

An important element of the capacity assessment methodology is to involve representatives of the organisation in the entire evaluation process, starting from the diagnosis and ending with the implementation of the action plan. The aim is to bring about change based on assessment and ongoing evaluation, and the essence of the process is to create convergence through jointly agreed and shared opinions of the participants in the self-diagnosis process initially carried out individually and, at a later stage, in groups. This sequence of actions facilitates the identification of relevant and critical issues and encourages in-depth reflection on priorities before deciding what should be included in the action plan.

The application of the described methodology makes it possible to verify organisational priorities and change the perception and treatment of the educational institution's priorities by employees and stakeholders. It can also result in changes in internal communication within teams, departments or units

10 Accounts concern remembered events and anecdotes illustrating a certain state of affairs in terms of individual behaviour, organisational solutions, interactions and the extent to which the external (institutional, legal, functional and spatial) environment is (or is not) enabling vis-à-vis effective teaching and learning.

of the organisation and in building better relations with the environment. Prior to conducting capacity development activities, expert practitioners familiarise themselves with the skills and competencies of the employees, the organisational structure as well as institutional relations with the environment in order to take this information into account in their subsequent work. This knowledge is a starting point for analysing and working on resources in all their dimensions: individual, organisational and institutional.

An accurate assessment of the initial state of capacities is an important step preceding the development of an action plan. If it provides for participation of all interested parties, its very course builds up organisational resources and individual competences.

Capacity and opportunity analysis – diagnosis

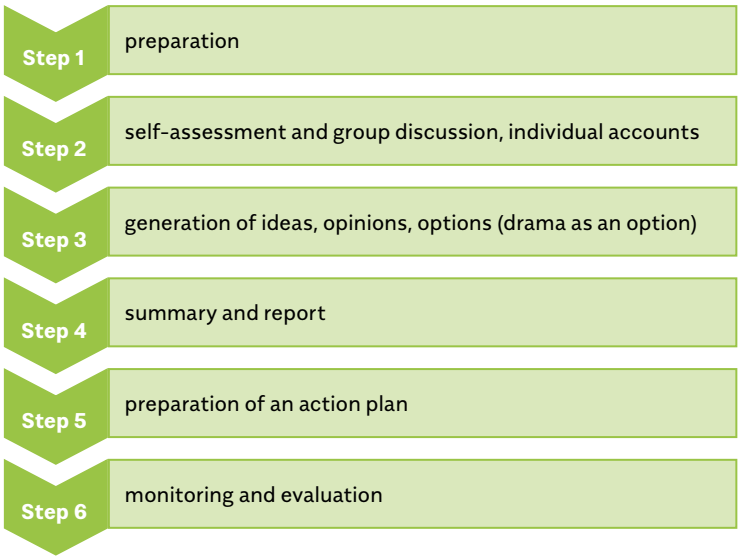
Analysing and assessing an institution's capacities is primarily a diagnostic tool. It is used to verify selected components of activity and results of all the institution's operations – the quality of services and/or the level of competencies of service recipients. A key stage is the integration of evaluation into the planned improvement, growth and change process, taking into account external and internal challenges and opportunities for the institution. In accordance with the presented approach, capacity assessment quite often consists in self-assessment in order to better understand the conditions of activity, in the case of teachers especially the specific quality and way of teaching and the possibility of raising this performance level in relation to the teaching process in the future.

Diagnosis based, among others, on self-assessment helps the process participants to:

- → gain an agreed, fuller understanding of the challenges faced by the educational institution and relating to development opportunities;
- → agree on selected development opportunities;
- → identify factors that can facilitate or hinder the introduction of changes.

When carrying out the diagnosis, several organisational solutions can be used, including written diagnoses, interviews, small group work and/or a workshop. In the case of the latter, which is the most recommended form (in practice it is usually a series of workshops), a facilitator conducts sessions during which they assist participants in describing the organisation in selected areas. The participants analyse the status of the institution in terms of the availability of capacities, capabilities and competencies of employees, and they review selected strategic documents (such as the mission statement, vision, objectives, challenges), management and communication practices, established processes of creating products/services or – if the evaluation is to have a narrower scope – only selected areas that are critical for the functioning of the organisation/institution.

Figure 5. A workshop plan



Source: author's own work.

During a workshop devoted to capacity assessment and development, it is necessary to create a climate of teamwork that fosters deeper reflection and understanding of the specificities of the achieved results. The self-assessment grid divided into selected categories of capacity assessment is designed to guide the discussion and encourage dialogue, exchange of opinions, sharing anecdotes and giving individual accounts. Questions are asked to initiate discussion on key areas of capacity assessment. Four descriptive stages in each category of capacity assessment serve the purpose of better reading the organisation and managing it more effectively. A group discussion in turn creates an opportunity to share hidden knowledge about the educational institution as an organisation.

One methodological solution is to use the self-assessment grid designed in such a way as to direct the participants' attention and, at a later stage, discussion towards capacity assessment and key challenges related to development opportunities. The capacity assessment grid (CAG) is based on a four-stage scale and it is a description rather than an assessment tool. It describes four areas: (1) capacities, (2) organisation and management, (3) teaching methods[11] and (4) relations with the environment.

Results

Having conducted the self-assessment, the appointed team with the help of a facilitator summarises and interprets the results obtained. They are presented in a report, which should contain:

→ a description of the context in which the self-assessment was carried out;
→ a summary of the analyses carried out in relation to current and planned capacities;
→ a review of comments and suggestions presented by stakeholders;
→ recommendations for the development of an action plan.

11 The area of teaching and training methods can be replaced by another, depending on the needs of the given organisation/institution.

The results of the self-assessment can indicate the need to strengthen a specific component or area of the educational institution's activity and to focus on it. They can also be grouped in order to set priorities, taking into account the team's needs, real capacity and time available for the implementation of the action plan.

Thanks to the analysis of the self-assessment results, the representatives of an organisation can broaden their knowledge of current and prospective opportunities for its development at various levels (individual, organisational and institutional) and monitor their growth. The assessment is designed to help:

- → better understand the strengths and weaknesses of leadership in the organisation;
- → assess the institution's/organisation's adaptability;
- → assess the effectiveness of capacity use;
- → better understand the extent to which the institution is capable of making organisational and operational changes;
- → know the extent to which the institution/organisation is capable of promoting learning and change;
- → know how power relations affect development opportunities;
- → better understand how to strengthen relationships between institutional actors through networking, partnerships and creating an environment that enables the development of organisations, institutions and companies;
- → better understand individual needs concerning employees' competency development;
- → learn more about opportunities for changing the organisation's culture.

A well-planned self-assessment should result in setting a clear direction for change. The results of discussions on progress towards achieving the agreed indicators, as well as perceived opportunities for development (e.g. in terms of increasing the effectiveness of the teaching process or the quality of educational services) should be included in the action plan. The organisation's reflection on selected components, the results of self-assessment and lines of action planned serve as a basis for ongoing and ex-post evaluation, both at the time of and after introducing changes.

Preparing an action plan

Preparation of the action plan is the stage following capacity diagnosis. Thanks to the diagnosis, strengths and weaknesses of an educational institution are known. An important part of the plan is setting priorities and establishing who is responsible for the implementation of its individual elements. The plan should describe the starting point (i.e. identified needs concerning capacities), feature a list of planned activities, their justification and a preliminary schedule (who?, what?, when?, etc.). In addition, it should take into account evaluation and monitoring methods, expected results and initially envisaged stakeholder participation.

The preparation of a baseline action plan is important for several reasons. Firstly, it allows for assessing whether the planned actions are realistic. Secondly, it makes it easier to identify risks, drivers and obstacles, which is often difficult at the planning stage. And thirdly, it helps to diagnose the state of current capacities and to justify why some of them are limited and relatively difficult to develop. A model action plan format is presented in the table below.

Table 4. Model format of an action plan

Capacity area /factor	Planned actions	Expected results /behaviour	Schedule	Evaluation and monitoring methods	Who will be involved in the implementation

Proposed stakeholder participation

Source: Warowicki, 2012.

Capacity development evaluation

Capacity development is a complex process and its results are often difficult to predict. This is so because there are interactions (and cycles of interactions) between factors at the individual, organisational and institutional level, the results of which are unknown. Therefore, as Douglas Horton (2011) puts it, evaluation of the capacity development process serves to identify relevant factors and conditions to improve and enable the process (as well as to obtain external support) rather than to make a traditional assessment using results based on a logical framework. The team responsible for capacity assessment carries out a learning-oriented evaluation in the form of one or several sessions and, where appropriate, involves process participants, including evaluation stakeholders. Using lessons learned from past experience in the diagnosis leads to a better, more in-depth capacity self-assessment.

The evaluation of development capacities self-assessment and the action plan give raise to the following questions:

→ what have we learned from the process?
→ how can capacity self-assessment be improved?
→ what changes can be made to the future assessment of resources and the next action plan?
→ what are the results of external support?
→ how can support for capacity development be increased?

It seems that, from the perspective of the planned change and due to the local context and the educational institution's environment, the participatory evaluation process is more adequate. It offers more than an evaluation carried out from outside without local participation. Moreover,

as already mentioned, participatory evaluation can be considered the next step in capacity development. In this context, it is a learning-oriented evaluation.

Interactive complexity of change models

A key strength of the application of the capacity development concept in educational institutions is the greater effectiveness of solving problems by means of activities aimed at strengthening local resources than by means of activities focussed on linear interdependencies. It is about the relationship between the diagnosed causes of problems and the expected results of actions aimed at reducing or eliminating negative phenomena.

This has important implications for the evaluation model as well as for the research profile of the education process. In accordance with the presented perspective, conditions for capacity development are analysed. Strengthening local resources consists in, among others, creating conditions which foster raising individual and team competencies and building a network of constructive relations with stakeholders. The traditional intervention model based on the popular logframe does not take into account the functional relationships between independent variables and factors, with which they are linked in real life. Learning and teaching implies interaction at least between the teacher and the learner. It also involves the reaction of the environment and the impact on this environment during the learning and teaching process (cf. Fontana, 1998; Feldman, 1990).

The learning and teaching process can be described as the transfer of knowledge and acquisition of competencies (although this does not exhaust its full complexity), as well as the elaborate sequence of individual and social behaviours occurring in a specific context. Social behaviours, including those relating to education, cannot be written down in the form of an additive function[12]. This limits the possibility of using quantitative methods (in a way that fully reproduces the complexity of links between variables) in social research, including research on education. Most phenomena and behaviours, not only social ones, are the result of interactions between variables that are either very difficult to identify or cannot be detected at all (cf. Wieczorkowska & Wierzbiński, 2012)[13].

Educational behaviour, just as, broadly speaking, social behaviour, always has cultural and situational significance. For example, depending on the context, shaking hands can signify: saying thank you, congratulating someone, reaching an agreement, saying hello or saying goodbye. Moreover, human behaviour has a functional meaning in a broader context and therefore eludes parametrisation. Hence the recurring interest of researchers and practitioners in theory and research focussing on the combination of ethnomethodology and the concept of interactionism initiated in the late 1970s by, among others, Rom Harré (1979, 1980)[14]. Unlike in typical research into social roles and rules governing interactions in education, the problem is analysed from the perspective of individual accounts

12 Using the additive function we can, for example, record the behaviour of billiard balls on the table, because the parameters determining the path of the balls' movement are not internally linked (there is no interaction between them). The situation is different in the case of social behaviour, where more often we deal with a contextual or functional definition of the nature of factors than the one used in natural sciences to define parameters. Among others, this is the difference between the cognitive dissonance phenomenon and the law of gravity. In science – in physics and chemistry, as well as in psychology – there are two basic methods of conducting research: quantitative (parametric) and structural. In psychology, an example of the application of quantitative and structural methods is the study of discourse in psycholinguistics (number of defined acts while taking into account separate structures divided into interaction units).

13 A *trivial example concerns sleeping pills and alcohol, which also puts us to sleep. If the influence of both variables was additive, taking sleeping pills and drinking alcohol would make us fall sleep even faster. However, the impact is not additive (the effects of the two variables do not add up), but interactive. We are at risk of dying, and this effect could not have been predicted knowing the impact of the individual variables* (Wieczorkowska & Wierzbiński, 2012, pp. 36–37). Similarly, the effect of combining oxygen and hydrogen resulting in H2O would not be predictable if we did not have the knowledge of the result of this combination and based it on adding the properties of oxygen and hydrogen.

14 Rom Harré proposed a reconstruction of the dramaturgy of events by attempting to recreate their context.

of event participants in their natural language (account analysis). Workshop participants or interviewees are asked to present their own account of events and incidents. This makes it possible to reconstruct the social "scene" based on the experience of its participants.

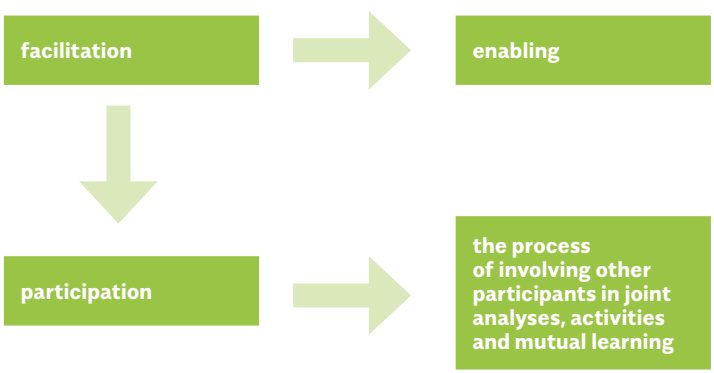
Drama can serve a similar function (cf. Somers, 1994, 1996). The analysis of the same incidents from the perspective of individual participants creates an opportunity to confront social understanding of social roles and the resulting rules of behaviour with their individual understanding (role-rule analysis). Thanks to the use of such qualitative methods, the process participants and researchers can better grasp the essence of the differences between the observer and the actor in a specific sequence of events and, as it seems, come closer to the nature of the studied phenomena. This also helps to smooth away differences between the "myself"/"others" perspectives (differences in access to information, the effect of focussing attention, the aims of myself/others perception, cf. Wieczorkowska-Wierzbińska, 2011). The analysis of participants' accounts in the capacity analysis process, which involves several dimensions, combined with quoting anecdotes, brings the capacity self-assessment method closer to Harré's (1980a) dramaturgical approach and to discourse analysis in psycholinguistics.

The role of facilitator

As already indicated, the participatory workshop format requires the presence of an experienced facilitator[15] who will lead the session, foster group dynamics, evenly distribute the activity of individual participants, and prepare and conduct exercises. It is the facilitator who, at the beginning of the workshop, provides participants with information about the objectives of the training, its programme and expected results. The facilitator is also responsible for logistics and preparation of a joint report. In practice, this means they play several roles: they are an exercise instructor and guide, active listener, supporting moderator, and process accelerator (facilitator). They also oversee the course of individual parts of the workshop and plan subsequent actions.

A facilitator – just like an experienced trainer – does not provide ready-made solutions, but asks the right questions, which stimulate participant reflection. During the workshop, facilitation and participation are linked.

Figure 6. Relationship between facilitation and participation



Source: Schwarz, 1994, pp. 269–270.

15 Facilitation is more focussed on enabling and supporting the group process, which goes in a not entirely known direction. Moderation consists mainly in convening a session or meeting and setting a direction for it.

The basic skills of a facilitator, particularly useful during workshops devoted to capacity assessment, are:

→ encouraging and sustaining participants' active participation and engagement;
→ introducing important new elements to the discussion;
→ supporting the participatory process;
→ summarising key points resulting from exercises and discussions;
→ taking decisions by consensus;
→ focussing the discussion on the subject of the workshop;
→ stimulating reflection and learning;
→ understanding cultural and contextual conditions;
→ promoting a sense of ownership and authorship of workshop results.

The role of facilitator seems essential in the process of capacity evaluation and development. As already mentioned, a facilitator helps create a climate of teamwork that fosters deeper reflection and understanding of the specifics of the results achieved earlier.

Application of the capacity development method in an educational institution

Every educational institution has certain capacities[16]. The key question is whether their current state allows for providing services of a sufficiently high quality.

Change-oriented process is the underlying assumption of the concept of capacity development. From this broad perspective it can be assumed that education is also change-oriented (e.g. new knowledge leads to changes in behaviour, attitudes and/or organisational structure). Discussions held over the last decade on the conceptualisation of the learning process have gone beyond the process itself and directed the attention of researchers and practitioners to a process that supports learning. This has certain implications, which include:

1. Going beyond the traditional way of defining learning resources (such as educational needs analysis, training design and planning), the education process (taking into consideration different forms and methods of training/learning and assuming an evaluation of the results), and moving from an analysis of "educational events" to follow-up, which forms an important part of education[17];
2. Transition from teaching to learning and focussing on an enabling environment, which requires facilitation, and the involvement of stakeholders and important institutional actors; these social and relational resources were not included in the traditional understanding of education;
3. Monitoring and evaluation, which is more focussed on learning outcomes (change of participants' behaviour and/or the procedures in an educational institution) than on recording judgements and observations of participants of a training or education process; therefore, there is a growing need to use participatory methods in evaluation, which provide for the involvement of instructors, teachers, trainers, IT specialists, training participants and stakeholders' representatives.

16 The same is true of the state of social capital resulting from the state of and access to capacities, as well as from social networks.

17 Taking support measures to complement the standard duration of teaching or training.

The new, emerging role of an educational institution as a facilitator of learning and developmental change[18] requires a non-traditional approach to evaluation[19]. In practice, this means that in addition to offering curricula and training courses aimed at participants' achieving a certain level of technical or social competencies (or in addition to them), an educational institution is tasked with strengthening relationships and involving stakeholders at the institutional level, in order to support learning.

This approach shifts the emphasis from the transfer of knowledge and skills at the individual level to consideration of mutual relations between organisations, institutions, networks and the system in which they are embedded. Therefore, the assessment of resources should be designed so that, with the involvement of selected process participants, new solutions can be developed, and a wider range of learning and change factors can be identified and articulated.

It is worth noting that both the aforementioned roles of an educational institution are equally important. From a broader perspective on learning and on the assessment and evaluation of the learning process, it seems that it is easier to diagnose relational resources at a higher, more complex level and as part of a more lengthy process (e.g. in relation to a two- or three-year programme, where training is only one form of intervention), while it is easier to evaluate a one-off short training course in a traditional way.

Conclusions

The development of educational institutions' capacity is in many respects similar to capacity development in other entities. The same goes for capacity evaluation. The difference lies in the specific character of competencies at the individual level, the nature of capacities resulting from the organisational structure of the institution (the organisational level), and its relationships with the institutional and social surroundings (its environment).

An approach based on capacity strengthening and development, which is more systemic than the study of linear dependencies, takes into account the functional relationships affecting the strengthening of capacities in a broad sense – from social, economic and cultural to environmental ones. Focussing on the needs of education service users, their deficits and problems (need based approach) and/or on the needs of the educational institution does not offer full analytical capability when searching for potential and strengthening local capacities (asset based approach). Such an approach often leads to the fragmentation of remedial actions and "hides" the fact that, in reality, problems are interconnected and result from a weakening or even collapse of potentials and capacities in a broader sense. Therefore, it seems that in education and in the process of capacity development it is also worth analysing the enabling environment.

18 To learn more on this subject, see: Grolnick & Ryan, 1987.

19 Meaning assessment.

Bibliography

Anderson, R.E. & Carter, I. (1984). *Human Behavior in the Social Environment. A Social Systems Approach*. New York: Aldine Publishing Company.

Blanchet, T. et al. (2019). *How Unequal Is Europe? Evidence from Distributional National Accounts. WID.world WORKING PAPER N° 2019/06*. Paris: World Inequality Lab.

Capacity Change and Performance: Insights and Implications for Development Cooperation (2008), *Policy Management Brief*, 21 December. Maastricht: European Centre for Development Policy Management.

Capacity Change and Performance. Capacity development: between planned interventions and emergent processes Implications for development cooperation (2009). *Policy Management Brief*, 22 March. Maastricht: European Centre for Development Policy Management.

Czapiński, J. & Panek, T. (2008, 2011, 2013, 2015). *Diagnoza społeczna. Warunki i jakość życia Polaków*. Warsaw: Rada Monitoringu Społecznego.

Cichocki, K.S. (2016). Wybrane modele wspomagające wieloletnie planowanie finansowe w jednostkach samorządu terytorialnego, *Finanse Komunalne, 11*.

Czapiński, J. (2004). Quo vadis homo?. In: M. Marody (ed.), *Wymiary życia społecznego. Polska na przełomie XX i XXI wieku*. Warsaw: Wydawnictwo Naukowe Scholar.

Czyżewska, A. (2001). Jak zwiększać udział społeczności lokalnej w zarządzaniu zmianami? Jak to robią w Sopocie?. In: M. Warowicki & Z. Woźniak (eds.), *Aktywność obywatelska w rozwoju społeczności lokalnej – od komunikacji do współpracy*. Warsaw: Municipium.

Czyżewska, A. (2012). *Jak planować proces rewitalizacji społeczno-gospodarczej przestrzeni miejskiej?*. Warsaw: Forum Rewitalizacji. Pracownia Badań i Innowacji Społecznych.

Defining and Measuring Capacity Development Results (2010). United Nations Development Programme.

Edwards, B. et al. (2001). *Beyond Tocqueville: Civil society and social capital debate in comparative perspective*. New Hampshire: University Press of New England, Hanover: Tufts University.

Effective Capacity Assessment for Non-Profit Organisations (2001). McKinsey and Company.

Feldman, R.S. (1990). *The Social Psychology of Education. Current research and theory*. Cambridge: Cambridge University Press.

Exploring Capacity Issues "On the Road to Seoul": Quality Training for Capacity Development, Discussion Note (2009). Paris: Organisation for Economic Co-operation and Development.

Fontana, D. (1998). *Psychologia dla nauczycieli.* Poznań: Zysk i S-ka.

Fourth High Level Forum on AID Effectiveness, 29 November – 1 December 2011. Busan, Korea, conference papers.

Evaluation Methodology & Baseline Study of European Commission. Particip-ADE–DRN-DIE–ECDPM-ODI, and Particip GmbH, leading company (2012), unpublished report.

Gajowiak, M. (2012). *Kapitał społeczny.* Warsaw: Wydawnictwo PWE.

Gliński, P. (2001). Dziesięć lat współpracy sektora publicznego i obywatelskiego: stan posiadania. In: M. Warowicki & Z. Woźniak (eds.), *Aktywność obywatelska w rozwoju społeczności lokalnej – – od komunikacji do współpracy.* Warsaw: Municipium.

Goodland, R. (1994). Environmental Sustainability and the Power Sector, *Impact Assessment*, *12*.

Goold, M. & Campbell, A. (2002). *Designing Effective Organizations. How to Create Structured Networks.* San Francisco: Jossey-Bass, A Wiley Company.

Grolnick, W. & Ryan, R. (1987). Autonomy Support in Education: Creating the Facilitating Environment. In: N. Hastings & J. Schwieso, *New Directions in Educational Psychology.* Philadelphia: The Palmer Press.

Harrison, M. (1986). *Diagnosing Organizations. Methods, Models, and Processes.* Applied Social Research Methods Series, Vol. 8. London: Sage.

Harré, R. & Secord, P. (1972). *The explanation of social behaviour.* Oxford: Blackwell.

Harré, R. (1979). *Social being. A theory for social psychology.* Oxford: Blackwell.

Harré, R. (1980a). Making social psychology scientific. In: R. Gilmour & S. Duck (eds.), *The Development of Social Psychology.* London: London Academic Press.

Harré, R (1980b). The Dramaturgical Model. In: A. Furnham & M. Argyle (eds.), *The Psychology of Social Situations. Selected Readings.* Oxford–New York: Pergamon Press.

Horton, D. (2011). Evaluating capacity development, *Voices of capacity development. Capacity, 43.*

Kazimierczak, T. (2014). *Organizowanie społeczności lokalnej: aplikacje, wdrażanie, przyszłość. Prace Laboratorium Innowacji Społecznej*. Warsaw: Institute of Public Affairs.

Kazimierczak, T. & Hernik, K. (2008). *Społeczność lokalna w działaniu. Kapitał społeczny. Potencjał społeczny*. Warsaw: Institute of Public Affairs.

Koopmans, T.J. (1957), *Three Essays on The State of Economic Science*, New York–Toronto––London: McGraw-Hill Book Company, INC.

Kretzmann, J. & McKnight, J. (1993). *Building Communities from the Inside Out. A Path Toward Finding and Mobilizing a Community's Assets*. Evanston: Northwestern University Center for Urban Affairs and Policy Research Neighborhood Innovations Network.

Lessons Learned on Capacity Assessment (2009). New York: UNDP Capacity Development Group.

Lewenstein, B. & Thesis, M. (2008). Kapitał społeczny, lokalne społeczeństwo obywatelskie, aktywizacja; zachodnie koncepcje, polskie doświadczenia. In: A. Kościański & W. Misztal (eds.), *Społeczeństwo obywatelskie. Między ideą a praktyką*. Warsaw: Wydawnictwo IFiS PAN.

Lewenstein, B. & Palska, H. (2004). Organizacje pozarządowe na scenie publicznej Polski okresu transformacji: dynamika rozwojowa i relacje z władzą – analiza badań jakościowych. In: P. Gliński et al. (eds.), *Samoorganizacja społeczeństwa polskiego: III sektor i wspólnoty lokalne w jednoczącej się Europie*. Warsaw: Wydawnictwo IFIS PAN.

Lewenstein, B. (2001). Obywatelskie strategie i modele wprowadzania zmiany społecznej w miejskich społecznościach lokalnych – doświadczenia amerykańskie. In: M. Warowicki & Z. Woźniak (eds.), *Aktywność obywatelska w rozwoju społeczności lokalnej – od komunikacji do współpracy*. Warsaw: Municipium.

Mondini, G. & Vale, M. (2007). Environmental Assessment within EU. In: M. Deakin et al. (eds.), *Sustainable Urban Development*, Vol. 2: *The Environmental Assessment Methods*. London: Routledge.

Morgan, P. (2006). *The concept of capacity*. European Centre for Development Policy Management.

Otoo, S. et al. (2009). *The Capacity Development Results Framework: A strategic and results--oriented approach to learning for capacity development*. New York: World Bank Institute.

Pearson, J. (2011a). *Section 1. Core concept revised*, unpublished.

Pearson, J. (2011b). *Training and Beyond: Seeking Better Practices for Capacity Development*. Organisation for Economic Co-operation and Development, Learning Network on Capacity Development.

Sange, P. et al. (2002). *Piąta dyscyplina. Materiały dla praktyka. Jak budować organizacje uczącą się*. Krakow: Oficyna Ekonomiczna.

Schwarz, R. (1994). *The Skilled Facilitator*. San Francisco: Jossey-Bass Publishers.

Somers, J. (1996). *Drama and Theatre in Education: Contemporary Research*. Ontario: Captus University Publications.

Somers, J. (1994). *Drama in the Curriculum*. London: Cassell.

Strengthening Training Institutions – Expert Group Meeting Report (2011). Nairobi: UN-HABITAT.

Supporting Capacity Development in PFM. A Practitioner's Guide (2011). Vol. 1, European Union, CIPFA, Organisation for Economic Co-operation and Development.

Swarts, P. et al. (2011). *Re-engineering Education and Training for Economic and Social Development in the 21st Century: A Focus on Technology and Skills Development in National Education and Training Systems in Africa*. Ministerial Round Table, Nairobi: GESCI.

Szarfenberg, R. (2013). *Nierówności społeczne*. Warsaw: Instytut Polityki Społecznej Pracownia Pomocy i Integracji Społecznej Uniwersytetu Warszawskiego.

Sztompka, P. (2016). *Kapitał społeczny*. Krakow: Znak.

Taylor, P. & Clarke, P. (2008). *Capacity for a change, a document based on outcomes of a Capacity Collective workshop*. Brighton: IDS.

Toolkit for Capacity Development, Tools and Methods Series (2009), No. 6. Brussels: European Commission.

The Challenge of Capacity Development: Working Towards Good Practice, DAC Guidelines and Reference Series (2006). Paris: Organisation for Economic Co-operation and Development.

Ubels, J. et al. (2010). *Capacity Development in Practice*. London: Earthscan Ltd., Dunstan House.

Walters, H. (2008). *Capacity Development, Institutional Change and Theory of Change: What do we mean and where are the linkages. A Conceptual background paper*. bit.ly/3ct1sFy [accessed: 15 February 2020].

Warowicki, M. (1985). Metodologiczne i teoretyczne uwarunkowania faktów empirycznych w psychologii społecznej. In: E. Paszkiewicz & T. Szustrowa (eds.), *Metody badań psychologicznych*, Series III, Vol. 4. Warsaw: Państwowe Wydawnictwo Naukowe.

Warowicki, M. (2008). Social Integration for Social Inclusion at the Local Level. In: *XXI Annual Conference of ENTO for dynamic local and regional public policies. Training for Innovation, Innovation in Training*. Desenzano del Garda, Italy, 2–3 October, conference papers.

Warowicki, M. (2012). *Capacity Self-Assessment Methodology for Training Institutions. Reading the Organization*. Nairobi: UN-HABITAT, unpublished report.

Warowicki, M. (2013a). *Polish CSOs Transition and State; Technical Assistance for Civic Society Organisation (TACSO)*. SIPU/FSLD. bit.ly/2NTmK4r [accessed: 15 January 2020].

Warowicki, M. (2013b). *Baseline study and project evaluation: Empowering marginalised girls in Faryab province of Afghanistan*. France: ACTED funded by DIFID, unpublished report.

Warowicki, M. et al. (2010). Zbieranie danych o sytuacji społecznej. In: N. Chojnowska-Ochnik & E. Singer (eds.), *Metodologia wspierania rozwoju społecznego gminy przez partycypację. Gminna Strategia Rozwiązywania Problemów Społecznych*. Warsaw: MPiPS.

Warowicki, M. & Woźniak, Z. (2001). *Aktywność obywatelska w rozwoju społeczności lokalnej*. Warsaw: Municipium.

Wieczorkowska-Wierzbińska, G. (2011). *Psychologiczne ograniczenia*. Warsaw: Wydawnictwo Naukowe Wydziału Zarządzania Uniwersytetu Warszawskiego.

Wieczorkowska, G. & Wierzbiński, J. (2012). *Statystyka od teorii do praktyki*. Warsaw: Wydawnictwo Naukowe Scholar.

Wignaraja, K. (2010). *Measuring Capacity*. New York: United Nations Development Programme.

Evaluation of support for competencies of people on the labour market and for employment promotion. The full picture of effects from the micro and macro perspective

DOI: 10.47050/66515246.88-106

Jacek Pokorski

Abstract

This chapter presents opportunities for combining participants' perspective with that of the market in the case of impact evaluation of various forms of support for the development of competencies of people on the labour market and for employment promotion. It analyses Polish evaluation studies which make use of the aforementioned micro and macro perspective (in particular, evaluation of the effects of the European Union's cohesion policy implemented in Poland) and which are based on the counterfactual approach. It presents the evolution of the approaches used so far, data sources and their limitations. It also points out the benefits of synergies arising from the combination of micro and macro research perspectives.

Introduction

For over a decade, as part of Polish evaluation studies, methodology for estimating the causal effects of various types of support activities, referring to the counterfactual approach, has been developed. In practice, such evaluations attempt to extract the impact of external factors (employee's features, employer's characteristics, economic situation, other individual, local and specific conditions, etc.) on the observed changes in people treated under the intervention and to estimate the net effect of support. For this purpose, appropriate comparison groups are used which illustrate the situation of beneficiaries who were not exposed to the support. On the basis of previous evaluations, and mainly using declarative methods to evaluate (gross) effects of training or other developmental services, it can be concluded that counterfactual impact evaluation, which has consolidated for some time, has become a milestone in the Polish evaluation system. The guidelines of individual directorates

Jacek Pokorski
Since 2004 he has worked at the Polish Agency for Enterprise Development (PARP) as Chief Expert at the Evaluation Unit. Since 2016 he has been Head of the Monitoring and Evaluation Section at the Analysis and Strategy Department. In the years 2010–2014 he was Vice President of the Polish Evaluation Society (PTE) and is currently the Head of PTE's supervisory body. He specialises in analyses of enterprises and the business environment and in evaluation of economic development programmes. He is an evaluator, trainer and consultant in the field of evaluation studies.

KEYWORDS:
LABOUR MARKET,
COUNTERFACTUAL IMPACT,
EVALUATION,
MICRO- AND MACROECONOMIC IMPACT,
COMPETENCIES,
EMPLOYMENT,
EUROPEAN FUNDS

of the European Commission (EC) – cf. *The Programming Period...*, 2013; *Programming Period...*, 2015; *Common methodology...*, 2014[1] – and general regulations concerning European funds also played a role in promoting this approach.

It should be stressed that various counterfactual impact evaluation options have recently been developed, which aim to appraise both absolute (net) effects and relative causal effects (Pokorski, 2016), and which indicate the advantage of one support instrument over another in terms of their effectiveness. Different data sources are tested as part of this approach. More and more often, data from public registers and statistical research are being reused for this purpose, and the scope of causal analyses is being extended. These cover participants, households, employers, and territorial government units responsible for implementing certain active labour market policies or supporting entrepreneurship. Nevertheless, the developed approach still seems imperfect due to a relatively narrow scope of effects analysed, limited to an individual perspective. Counterfactual impact evaluations made to date, based on matching participants and people who have not been exposed to the intervention – which is supposed to make these individuals statistically almost identical (theoretically they differ only in whether they participated in the programme) – and analysing the differences in their situation, changing over time, are in fact limited to effects seen from a micro perspective. Thus, they show how the provided support changed the situation of a person on the labour market in terms of their competences or behaviour (motivation, activation). This can translate into changes on the labour market in their immediate environment, yet such correlation may not exist at all. In large and comprehensive support schemes – such as, for example, the Operational Programme Knowledge Education Development 2014–2020 (OP KED), co-financed in Poland from the European Social Fund (ESF) – which focus on change not so much of an individual or local character, but of a more general one (at the level of many regions, sectors and the whole country, e.g. in the area of improving the adaptability of enterprises so that they can face current economic challenges), evaluation of effects from the perspective of programme participants is definitely insufficient. In order to obtain knowledge about the effectiveness and efficiency of supporting people's competences, a macroeconomic analysis is also needed, which would show how the change affecting beneficiaries translates and extends into the labour market, and the extent to which the programme/policy objectives are achieved. At the same time, as in the abovementioned micro-approach, this analysis must also refer to a zero (baseline, counterfactual) scenario or alternative directions for intervention.

Here, just like a decade ago, we have a considerable theoretical and empirical output in the area of macroeconometric modelling. However, attempts to use it in the evaluation of impact of labour market or other policies (e.g. evaluation of different types of instruments supporting investment in production or R&D in enterprises), which indirectly affect employment (unemployment, professional inactivity) and its sectoral structure or wages and labour productivity, remain rare. As in the case of counterfactual impact evaluation from the perspective of individuals, where the theoretical basis has been developed by world-renowned authorities in the field of microeconomics (e.g. James Heckman, Donald Rubin), current paradigms of studying the macroeconomic impact of various types of policies on the labour market also refer to the works of Nobel Prize winners. A good example here are the works of Dale Mortensen and Christopher Pissarides on labour market frictions and the resulting possibility

1 Cf. ec.europa.eu/regional_policy/sources/docoffic/ 2014/working/ wd_2014_en.pdf [accessed: 15 September 2019]; ec.europa.eu/regional_policy/sources/wikiguidance/gne165_esf_monitoring_and_evaluation_guidance.pdf [accessed: 15 September 2019]; ec.europa.eu/transparency/regdoc/rep/10102/2014/EN/10102-2014-179-EN-F1-1.PDF [accessed: 15 September 2019].

of analysing both flows from unemployment, employment and inactivity, and data on new jobs created as a result of a specific intervention.

Therefore, especially in the case of impact evaluation of larger initiatives aimed at the development of competences and employment promotion, an evaluation approach combining the advantages of both research perspectives – that of participants (employees/employers) and of the market – seems optimal. It should be stressed that the Polish Agency for Enterprise Development (PARP) – has significant experience in the scope of the abovementioned evaluations of micro- and macroeconomic impact of European funds, especially of the ESF and the European Fund for Regional Development in areas of the labour market (cf. Mogiła et al., 2019). This chapter presents Polish experience (including that of PARP) in the application of both approaches, their evolution and opportunities for synergy in the evaluation of EU cohesion policy, with particular emphasis on direct support for competences of individuals (employees) and the adaptability of enterprises (their employers).

Counterfactual evaluation of the impact of programmes on human resources of enterprises and on the situation of employees (microeconomic perspective)[2]

In Poland, the first initiatives regarding impact evaluation of competence support and employment promotion, based on the counterfactual approach and quasi-experimental schemes, concerned the pre-accession Phare Economic and Social Cohesion programme (2002, 2003). Analyses carried out by PARP at that time concerned the so-called active labour market policies (ALMP), *inter alia* training for the unemployed, as well as subsidy programmes for investments and consultancy services for entrepreneurs[3]. In the former case, with the agency of the then Ministry of Labour and Social Policy, data from the Register of the Unemployed operated by County Labour Offices[4] were used. Both information which made it possible to match the treatment group with the comparison group using the averaged index of similarity characteristics (propensity score matching – cf. Trzciński, 2009), and the programme effect indicator, i.e. taking up employment, were used. Information on the percentage of programme participants who, within a certain period of time after completing the training, still remain in the Register of the Unemployed, as well as on those whose professional situation changed significantly, was the key (and, in principle, the only) measured net effect of the intervention. Differentiating features were also taken into account (in the scope of the training programme, e.g. effectiveness of training for persons planning to set up a business vs. specialised training for future employees; or in the scope of key characteristics of participants, e.g. how long they remained in the Register of the Unemployed, their gender, age, etc.). Undoubtedly, at that time secondary use of administrative (registry) data was ground-breaking when compared with current trends in conducting this type of impact assessment. It contained population data, as opposed to the results

2 It should be pointed out that fragments of this section presenting the development of counterfactual impact evaluation of a programme on treated individuals (micro perspective) were taken from the report entitled *Badania wykorzystujące dane rejestrów publicznych*, drawn up in 2018 by the author on behalf of the Polish Economic Institute (IBRKiK-PIB).

3 When carrying out these analyses in the years 2006–2007, PARP cooperated, among others, with PBS, Pentor Research International and the Jagiellonian University.

4 The PULS system comprises broad characteristics of the unemployed, both beneficiaries and non-beneficiaries of the ESF – identified in the system on the basis of their personal identification number (*PESEL*) – as well as information on whether they have taken up a job/set up a business and, as a result, de-registered from the unemployed system.

of statistical research of the Statistics Poland (GUS) or *ad hoc* surveys carried out on samples of beneficiaries and a control group used for this purpose.

Similar analyses concerning effects in the area of employment, yet relating to support not for employees but for employers, were carried out by PARP as part of *ex post* evaluation of advisory (2.1) and investment (2.3) schemes of the Sectoral Operational Programme – Improvement of the Competitiveness of Enterprises 2004–2006. At that time, the analyses covered a group of applicants from whom final beneficiaries and the so-called inefficient ones were chosen. Using detailed characteristics of each group, determined on the basis of grant application forms, control cases were matched to beneficiaries. The PSM index was also used for this purpose, calculated based on these group characteristics. However, data from the Statistics Poland[5] (GUS) concerning cases from both groups were used as a source of information on the effects of the intervention and were compared using the "difference in differences" method[6]. As regards the effects of employment, the analysis included the indicator of the number of people employed, average employment in terms of full-time jobs and the salary level.

Since then (2006–2007), at least a few relatively similar projects have been implemented in the area of entrepreneurship and the labour market. They were based on a rigorous methodology of counterfactual analyses and on GUS microdata. The most important (with the greatest cognitive and developmental achievements for this type of methodology) include:

1. Analysis of net effects of selected measures of the Operational Programme Innovative Economy 2007–2013 using the counterfactual impact evaluation approach as part of the PARP evaluation project entitled BAROMETR INNOWACYJNOŚCI (Innovation Barometer) (PARP/Jagiellonian University/EGO, 2015)[7];

2. Determining the value of selected economic indicators for beneficiaries of Regional Operational Programmes (ROP) and the Operational Programme Innovative Economy (OP IE) 2007–2013 and for selected control groups (GUS, 2015)[8];

3. Determining the value of selected economic indicators for beneficiaries of the Human Capital Operational Programme 2007–2013 (HC OP) and for selected control groups (GUS, 2017)[9].

At the same time, it should be mentioned that since 2016 the responsibilities of GUS for conducting this type of analysis for the needs of counterfactual impact evaluation of the system for the implementation of European funds carried out by various institutions have been formalised.

5 To put it simply, after matching an appropriate control group from among the so-called ineffective applicants with beneficiaries, which was done on the basis of characteristics available in the programme applicant base, GUS was asked to calculate for the indicated list of entities – identified by the taxpayer's identification number (*NIP*) and the National Business Registry number (*REGON*) – the values of average measures for specific intervention effect indicators (e.g. export sales revenues, employment, investments for specific years, in which the programme effect was expected). The indicators referred to data collected by GUS and obtained from enterprises on the basis of F01 forms.

6 The difference in the value of the indicator at the target point, calculated after the difference between the cases in the base period has been eliminated.

7 See: badania.parp.gov.pl/images/badania/Raport_Barometr_netto_POIG.pdf [accessed: 15 January 2020].

8 See: stat.gov.pl/files/gfx/portalinformacyjny/pl/defaultstronaopisowa/5807/1/1/raport_rpo_poig.pdf [accessed: 15 January 2020].

9 See: stat.gov.pl/statystyka-regionalna/statystyka-dla-polityki-spojnosci/statystyka-dla-polityki-spojnosci-2016-2018/badania/badania-kontrfaktyczne/ [accessed: 15 January 2020].

This was done in accordance with the letter of intent signed by the Ministry of Regional Development and the Statistics Poland. It speaks of the need for GUS to provide counterfactual analyses for the purpose of evaluating operational programmes implemented in Poland under the EU cohesion policy in the years 2014–2020, as well as the need for appropriate financing of research works commissioned to the Office (under the Technical Assistance Operational Programme 2014–2020, see: *Statystyka dla polityki spójności...*, 2015, pp. 3, 110–115). At present (at the turn of 2020), at least three counterfactual analysis projects are underway, carried out by evaluation units of PARP, the National Centre for Research and Development and the Ministry of Investment and Development (National Evaluation Unit) in cooperation with the Statistics Poland (based on microdata from the Office's statistical research, using rigorous methodology and maintaining strict rules of statistical confidentiality).
The strengths of this counterfactual analysis model are as follows:

- → High analytical capacity of public statistics services, which can meet the challenges of counterfactual analysis methodologies;
- → Providing analyses with the observance of statistical confidentiality – microdata, e.g. ID of participants, are transferred to GUS from the outside; the Office makes estimates based on such microdata, if such entities are in its registers, it calculates the differences in indicators between individuals from the treatment group and the control group, expressed in average measures (central trend); only average measures, e.g. nominal values, percentages, y/y dynamics, etc., are sent outside;
- → Different models of cooperation: both broad, assuming a greater role of GUS, including the selection of samples in the matching model, taking into account the characteristics of companies available in the Office's registers, and narrow, assuming that GUS only makes calculations based on a ready-made match, e.g. according to a predesigned algorithm.

However, as far as weaknesses of the described model are concerned, the following issues come to the fore:

- → conducting analyses based on data of the Statistics Poland is possible, provided that the entities which are of interest (beneficiaries, control group representatives) participated in statistical research of the Office[10];
- → GUS data for specific reporting years can be used, but with an almost two-year delay[11];
- → as a rule, the model of cooperation with GUS assumes the obligation on the part of the unit commissioning work to the Office to provide experts (its own or independent evaluators) who will be responsible for adequate preparation of the set for GUS, taking into account relevant variants

10 For example, microenterprises (entities employing fewer than ten people) do not participate in the full SP census entitled *Roczna ankieta przedsiębiorstw* [*Annual questionnaire for enterprises*]. However, this group of entities is covered by the SP-3 census, which for many reasons is partial (representative), which reduces the probability of coverage (feasibility of analyses). What is more, for the target group of the SP census the Statistics Poland does not need to have a full data set for a given entity (ID) for specific years. If we want to calculate the dynamics of y/y pre- and post-index, we need a set of data for four reporting years (2 pre- and 2 post-) and the same for the "twin" from the control group. The lack of one year in a pair of analysed entities excludes the possibility of calculating the "difference" (DiD). In the case of analyses carried out within the framework of Project No. 1 – *Analiza efektów netto...* [*Analysis of net effects...*] *for Action 4.4 IE OP, 31 pairs of companies for which full information was available (measurement for a given indicator in two time points for both companies in a pair) were found in GUS, which constituted 31% of the initial pool. For Action 8.2 IE OP, 104 pairs (for the majority of indicators) of companies that submitted SP reports for the indicated years and had no data gaps were found* [...], *which constitutes 29% of the initial pool.* Cf. *Raport z "Analiz..."*, p. 96, footnote 36, badania.parp.gov.pl/images/badania/Raport_Barometr_netto_POIG.pdf [accessed: 26 September 2018].

11 This is so because of: a) time needed by GUS to collect relevant information, as the research start date is c. April of the year following the reporting year, i.e. in general right after the closing of main accounting processes at companies; b) restrictive procedures (consistent with relevant statistical standards) for the verification of data and preparing sets that can be used; c) the schedule of data publication, e.g. GUS annual reports *Działalność przedsiębiorstw niefinansowych* [*Activity of non-financial enterprises*]; as a rule, the use of data collected by GUS is not possible before their official publication (according to the schedule for the statistical research programme) within a given round of company research, for any other targeted (commissioned) analyses; the National Bank of Poland uses a similar approach.

of analysis models, depending on data coverage; interpreting the results of evaluation taking into account the intervention logic and evaluation criteria should also be the responsibility of such experts; GUS does not carry out evaluation, only provides data for it (calculations, without interpretation);

→ Depending on the model of cooperation, *the obtained results of analyses* [...] *should be interpreted with caution, because* [...] *it is not possible to assess the statistical significance of the observed differences between the experimental group (beneficiaries) and the control group (non-beneficiaries), due to the aggregated nature of data obtained from the Statistics Poland*[12], although in new projects implemented under the *letter of intent between the Ministry of Regional Development and GUS, this limitation is less frequent* (Hermann-Pawłowska et al., 2017, pp. 99–116)[13].

Many of the abovementioned limitations can be overcome by using data that are relatively complete and available in almost real time (just after the forms have been submitted by users or approved by GUS) in public registers (such as those of the National Revenue Administration, the Social Insurance Institution, and the National Court Register). Human resources (top-class analysts, data engineers, business intelligence experts), technical infrastructure resources enabling the processing of large and properly secured data sets and the effectiveness of enforcement of reporting obligations remain a certain challenge.

From among the abovementioned evaluation projects, in the context of the labour market issues, it is worth focussing on the third one – Ustalenie wartości wybranych wskaźników ekonomicznych dla beneficjentów Programu Operacyjnego Kapitał Ludzki 2007–2013 oraz dla dobranych grup kontrolnych [Establishing the value of selected economic indicators for beneficiaries of the Human Capital Operational Programme (HC OP) 2007–2013 and for selected control groups]. The aim of the HC OP actions selected for analysis was, *inter alia*, support and promotion of entrepreneurship and self-employment. In contrast to projects funded by the European Regional Development Fund – ERDF (Operational Programme Innovative Economy – IE OP, Regional Operational Programmes), this was soft support and microfinancing (broadly understood as training services and relatively small financial support related to starting a business). Therefore, the study focussed on establishing the value of economic indicators for start-ups.

It covered the participants of HC OP projects who, thanks to financial support from the ESF, set up business activity, as well as newly established non-beneficiary companies (entities selected for the control groups) (*Ustalenie wartości...*, 2017)[14]. Again, statistical research conducted by the Statistics Poland, and not business activity registers, remain the main source of data, although the National Business Registry (*REGON*) database was also used in this research to calculate the "survival rate" and provide characteristics that were controlled in the model (PKD code of business activity, company location). Moreover, it served as a link with the PESEL number (personal identification number of a person participating in the project, coming from the PEFS[15] database for OP HC). Without the matching of REGON and PESEL, this type of

12 Cf. Raport PARP z "Analiz...", p. 7, footnote 6, badania.parp.gov.pl/images/ badania/Raport_Barometr_netto_POIG.pdf [accessed: 15 September 2019].

13 See: www.parp.gov.pl. Publicationslibrary/ebook/719 [accessed: 15 September 2019].

14 See: stat.gov.pl/statystyka-regionalna/statystyka-dla-polityki-spojnosci/statystyka-dla-polityki-spojnosci-2016-2018/badania/badania-kontrfaktyczne/ [accessed: 15 September 2019].

15 *The PEFS 2007 system (European Social Fund Monitoring Subsystem 2007) is an IT system designed to monitor the effects of projects co-financed by the ESF. When submitting a payment request, a Beneficiary is obliged to provide* [...] *data concerning the participants of projects carried out under HC OP,* cf. pokl.parp.gov.pl/index/index/2242 [accessed: 15 September 2019].

analysis would not have been possible. At the same time, the possibility to match REGON with PESEL allowed for identifying only these participants who set up a business in the form of a "natural person conducting economic activity". Possible other forms of economic activity (companies) chosen by HC OP participants were not included in the analysis.

As the report shows, *with regard to interventions which are the subject of this study (HC OP instrument 2007–2013 addressed to persons starting a business: Actions 6.1.3, 6.2, 8.1.2), the survival rate of businesses founded by the beneficiaries was considered a starting point for effectiveness and sustainability.* Other indicators were of a complementary nature, including those *concerning the number of employees, the form of employment or annual salary costs* [...], *also the effects related to financial performance and investment activity (e.g. the volume of revenues and costs, the value of fixed assets and fixed capital formation)*. As opposed to the previous two studies (PARP, GUS) concerning grants under the Operational Programme Innovative Economy (OP IE) and Regional Operational Programmes (ROP) for enterprises with an established market position, where it was expected that thanks to the support these entities would significantly change in comparison to their "statistical twins" not targeted by the support, the situation in the described case was slightly different.

> [...] *the desired state (i.e. in line with the expected effects of the intervention) was considered to be achieved when the value of indicators for beneficiaries was close (not worse) to that observed in the control group.* [...] *Taking into account the specific nature of OP HC* [...], *the support granted was not aimed at creating enterprises, which are distinguished by their above-average market position, but rather at creating stable jobs, which would allow the beneficiaries to continue their professional activity in the long term. Lack of effectiveness can therefore only be said to exist if the supported companies achieve significantly worse results than the representatives of the control group. The study therefore sought answers to the following main research questions:*
>
> 1. *Does the survival rate of companies established by beneficiaries differ from that of similar entities established without the support of European funds (including the ESF)?*
> 2. *Does the profile of companies founded by beneficiaries differ from that of similar entities established without the support of European funds (including the ESF)?*
> 3. *What are the financial results, employment dynamics and investment outlays in the companies established by the beneficiaries in comparison to other entities established without the support of European Funds (including the ESF)?*[16]

The project *Establishing the value of selected economic indicators for OP HC beneficiaries and for selected control groups* also revealed various limitations. The main analytical challenge was to compare treated beneficiaries and the control group, i.e. enterprises established in the same period and possibly characterised by a similar profile, but not benefiting from subsidies.

16 Cf. ibidem. Finally, based on the SP-3 form, apart from the survival rate (REGON), the study also analysed the following indicators in the scope of employment-related effects:
– total number of employees (excluding students);
– average number of employees working under an employment contract in a given year calculated as full-time employment;
– total number of employees (excluding students): including those for whom the entity is the main workplace;
– total number of employees (excluding students): including those for whom the entity is the main workplace: including women;
– total number of employees working under an employment contract (full- and part-time) including seasonal workers;
– gross wage paid in a given year to employees working under an employment contract [...], including mandatory contributions.

The feasibility of the above analyses was conditioned by:

→ The availability of beneficiaries' data in the REGON register and in the sets coming from statistical research conducted by the Statistics Poland using SP-3 (Report on the economic activity of enterprises) an PL1/PL2 (Panel research of enterprises) forms;

→ The possibility of identifying in the abovementioned sets a sufficient number of enterprises meeting the criteria of the control group, i.e. characterised by a profile close enough to that of treated beneficiaries to make it possible to carry out counterfactual analyses[17].

All these limitations are connected with the fact that the analyses are not carried out on a full set (register) of newly established entities, but on partial data (representative – a very numerous[18], but nevertheless a random sample). In this case, using data from the Social Insurance Institution (ZUS) (information from payers' accounts – i.e. companies run by natural persons identified via the combination of PESEL and REGON numbers – in the scope of contribution payment, i.e. employment and remuneration indices) would make it possible to obtain more reliable data (unencumbered by an error resulting from random samples, not of an approximate nature). On the other hand, payments of taxes observed in the registers of the National Revenue Administration (KAS) would make it possible to present (and not estimate!) the effect on financial indicators. Of course, we would also have to deal with certain limitations of the scope of information collected, in this case by Tax Offices (in particular – depending on the form of company settlements – the revenue and expense ledger, revenue register, lack of separate register/flat tax amount).

It should be stressed, however, that today, for similar purposes, the OP KED Managing Authority (the ESF Management Department at the Ministry of Investment and Development) uses the data of the Social Insurance Institution by analysing contributions and payers (in this case the employees of enterprises covered by the training programme), monitors the "economic fate of programme participants", e.g. in the form of a change in the level of remuneration of training participants, as well as their remaining at or leaving the place of employment. From a microeconomic perspective, it seems that in the area of employment promotion and labour market analyses, ZUS data has the greatest potential for being used in counterfactual impact assessment.

It is of note that recently we have faced a revolutionary change in the area of secondary use of ZUS data for, *inter alia*, the needs of programming and analyses of development policies (especially evaluation of social programmes). Thanks to the amendment of the Act on the social insurance system,[19] ZUS can provide, *inter alia*, the minister competent for regional development with a relatively broad catalogue of data (referred to in Article 40 of the Act) related to the recording of social insurance contributions (including information on the code of the form of employment and the amount of contributions paid on this account). Although this is a breakthrough solution based on the possibility of exchanging microdata[20] of persons and entities, the scope of its use allowed by law is relatively narrow (closed).

17 *Due to the connectivity of databases and the requirements concerning the adequate sample size, it was necessary to narrow down the data for effect indicators to those coming from SP-3 forms (a much smaller sample of data coming from PL forms resulted in too many data gaps). Data collected on the basis of SP-3 forms* [representative analysis on a random sample of entities employing fewer than ten people – author's note] *does not allow for tracing the same specific beneficiaries*, as opposed to PL1/PL2 panel research covering the same newly established entities each year. *As the samples in this study are not panels, it is impossible to track changes in the effect indicator of a specific beneficiary in subsequent years and therefore to estimate the DiD* (cf. ibidem).

18 For example, for Sub-action 6.1.3 of HC OP, a data set including 150,000 companies was used, where for each reporting year (year of activity) the control group of companies was over ten times larger than that of project participants.

19 Act of 13 October 1998 on the social insurance system, uniform text, par. 50(3a, b).

20 On the other hand, the Ministry is also able to transfer data to ZUS. Cf. Act of 11 July 2014 on the rules for the implementation of cohesion policy programmes financed under the financial perspective 2014–2020, uniform text, par. 71(2) – this is the so-called implementation act.

We are talking about ZUS data collected on the accounts of the insured who are "ESF participants" and on the accounts of their payers, including all necessary identification data of payers, such as: NIP, REGON, PESEL, PKD, etc. (paragraph 3a) as well as, which is worth emphasising, "non-participants in the ESF", which also refers to both the insured/employees or payers/employers (paragraph 3b). The aim of using this option, which is provided for in the law, is only (unfortunately) to meet the requirements of the ESF Regulation and the General Regulation for the European Structural and Investment Funds. In the case of the former, these are indicators concerning the obligation to measure output and long-term results (including those related to employability) for ESF investments and the Youth Employment Initiative. The latter is even more widely used to evaluate the impact of operational programmes in the years 2014–2020 using support from the ESF. Access to data of "non-ESF participants" is to be used for analyses carried out by means of quasi-experimental methods (comparing ESF participants to an appropriately selected control group of non-participants).

Significantly, this solution breaks down barriers (legal and mental) that have so far been impossible to remove and allows for the use of a vast scope of ZUS data in the analyses that indirectly serve the purpose of strategic management of ESF interventions in Poland. At the same time, it has great potential for use in many other areas, all the more so when combined with other registers. They can be used, for example, in analyses in the area of education, labour market policy, entrepreneurship development, especially in the context of personnel and hitherto unidentified determinants and dependencies in this respect, which concern socio-demographic characteristics, costs/wage structure, movement of people between companies, in connection with economic factors, etc. Outside the ESF, objective (non-declarative) data could be used for the evaluation of interventions affecting employability and future remuneration of workers (e.g. the Erasmus+ programme) and the performance of employers (e.g. ERDF).

ZUS data was used in evaluation reports prepared on behalf of the Ministry of Investment and Development in the years 2018–2019 for the purpose of carrying out:

1. Studies on the effects of support provided to young people under OP KED 2014–2020[21];
2. A meta-analysis of the results of evaluation studies concerning the evaluation of support provided as part of the ESF[22].

An unquestionable advantage of using ZUS data is its full coverage for all support participants. Thus, the problem of sample randomness (and the declarative nature of information obtained by means of surveys) when calculating the values of long-term result indicators was eliminated. However, the use of ZUS data to calculate indicators also involves certain limitations. In some forms of employment, social security contributions are not always paid. This applies to persons working on the basis of contracts for a specific task, students under 26 years of age employed under contracts of mandate, persons working in the grey economy, and those who run a farm and are insured in the Agricultural Social Insurance

21 Cf. *Badanie efektów wsparcia...*, 2018. It was a long-term study of the effects of support provided to young people (PO KED, Youth Employment Initiative – YEI). The persons covered by the support (study) are NEET (not in employment, education or training), and are 15–29 years old. The aim of the study was, among others, to calculate the value of four long-term result indicators (from the Common List of Key Indicators), measuring the effects of support provided to young people as part of OP KED, six months after the end of treatment. Cf. www.power.gov.pl/media/56886/III_raport_wskaznikowy_27_03_18.pdf [accessed: 15 January 2020].

22 *Metaanaliza wyników badań ewaluacyjnych...* 2018, pp. 207–208. This was a long-term study aimed at synthesising the results of the 2017 MA and IB evaluation reports on programmes using support from the ESF (OP KED, ROP) and estimating the value of selected indicators (from the WLWK list). The persons covered by the support (study) are the participants of OP KED who are inactive in the labour market (not in employment, education or training), belonging to different social groups, six months after receiving ESF support. Cf. www.power.gov.pl/media/60892/Metaanaliza_EFS_RCz_2018.pdf [accessed: 15 September 2019].

Fund (KRUS)[23]. Also, the analysis of earnings using ZUS data is limited. Apart from situations where ZUS contributions are not paid, contributions paid by the self-employed which are not income-related are also problematic[24].

It is of note, however, that ZUS data indicating the survival rate of start-ups that received support (i.e. continuation of business activity by persons who started self-employment), together with parameters concerning the situation of their personnel (the amount of contributions paid, approximate size of the remuneration fund) remain unavailable from other data sources used so far in evaluation studies (GUS statistical research, REGON register, KRS).

On the basis of the report entitled *Metaanaliza wyników badań ewaluacyjnych dotyczących oceny wsparcia w ramach EFS* [*Meta-analysis of the results of evaluation studies on the evaluation of support under the ESF*] and documenting the multi-annual (2016–2018) counterfactual analyses of ESF impact carried out on the basis of ZUS microdata, we can tell how many people were still running their newly started business 30 months after project completion (53%) and what the financial effectiveness of the support was (cf. *Metaanaliza wyników badań ewaluacyjnych... Raport cząstkowy...*, 2018, pp. 3–4)[25]. It is worth noting, however, that in the final report on the study, the micro approach was complemented by a mezzo perspective on the impact of the ESF on the local (county-level) labour markets. It tells us that:

> *1. The support provided to the unemployed has contributed to reducing registered unemployment at the local level;*
> *2. The award of one thousand start-up grants in the years 2015–2017 contributed to an increase in the number of natural persons conducting business activity in 2018 in the county of residence of project participants (on average by 760);*
> *3. Reducing the number of the unemployed in the county by one person required eligible expenditure on support for registered unemployed persons amounting to PLN 7,000 on average* (Tarsa, 2019, p. 16)[26].

> *As the authors of the report indicate, the weakness of the counterfactual approach so far is that: it ignores the indirect effects on non-participants, especially the substitution effect. For example, thanks to the support, a participant could get a job at the expense of a non-participant, which means a neutral impact of support on the employment level, but is interpreted as a positive net effect in the form of improvement of the participant's professional situation. Likewise, setting up a new business with the help of ESF funds can contribute to driving existing businesses from the market. Therefore, the results of typical analyses help to assess the impact of support on the future fate of a participant, but do not provide information on the impact of the intervention on the labour market, including e.g. on the level of unemployment* [at the local level – author's note]. *The approach used in this report helps to fill the knowledge gaps, as the impact of ESF interventions on* [...] *the number of registered unemployed people*

23 *The scale of underestimating employment success rate calculated on the basis of ZUS data is probably small* [...] *and reaches between 0.5* [...] *and 3%*. (cf. ibidem).

24 *Of those working six months after the end of their participation in support, 5% worked under more than one contract at the same time. In the case of persons insured under several different titles, the majority were concurrent employment contracts and contracts of mandate, which may result in the contract of mandate not requiring social security contributions* (*Badanie efektów wsparcia...*, 2018, pp. 30–31).

25 Cf. www.power.gov.pl/media/60892/Metaanaliza_EFS_RCz_2018.pdf [accessed: 15 September 2019].

26 Cf. www.ewaluacja.gov.pl/media/79605/Metaanaliza_wynikow_badań_ewaluacyjnych_dotyczacych_ oceny_wsparcia_z_EFS.pdf [accessed: 15 September 2019].

> *in a participant's county of residence is analysed. The results, therefore, take into account substitution and other indirect effects* […][27].

The evolution of the approach noted in the last edition of this study – i.e. the combination of participant and local market perspectives – should therefore be assessed positively. Nevertheless, it can be treated as an intermediate stage between the evaluation from the microeconomic perspective and evaluation of ESF impact on the labour market in Poland and employment indicators at a macro level.

Counterfactual evaluation of programmes' impact on the labour market (macroeconomic perspective)

As indicated in the introduction, the evaluation of effects from a microeconomic perspective is valuable, as it allows for checking whether an intervention mechanism is at all effective for participants and their individual situation on the labour market. However, if evaluation units and evaluators stop at the micro perspective, giving up the observation of the target area from "a bird's eye view" (macro perspective), they accept lack of knowledge about indirect impact of the support on the wider programme environment (including on the market of service providers and other labour market actors). They will also have no knowledge of possible flows of effects and social change expressed in the programme objectives. This means that such evaluations in a sense reduce the social utility of the programme to only one perspective and provide an incomplete picture of the effects.

Analyses using macroeconometric modelling have a long history in economic research. Their role has also been highlighted in the area of monitoring and evaluation of the use of European funds. Starting from the years immediately preceding Poland's accession to the EU, various types of research works based on macromodelling and concerning ex ante estimated impact of EU funds on the level of basic macroeconomic indicators of the Polish economy (GDP, employment/unemployment, investments, etc.) were systematically commissioned by ministers responsible for regional development. The latest such analysis was commissioned in 2019 by the Development Strategy Department of the Ministry of Investment and Development. It concerns the impact of cohesion policy on the social and economic development of Poland and its regions in the years 2004–2017[28]. It also takes into account estimates of the expected impact of funds awarded as part of the current perspective for 2014–2020, both at the national and regional level. Moreover, the analyses document historical impact of the funds since 2004 and provide forecasts until 2023 concerning, among others, employment and income levels of the residents. For example:

> *In 2018, the employment rate of people aged 20–64 in Poland was 72.2%* […]*, and if not for EU funds, the employment rate* […] *would have been 70.3%.* […] *The strongest impact was recorded in the Warmińsko-Mazurskie Province (3.0%) and the Podkarpackie Province (2.6%)* […];

27 *This approach also has certain limitations. It ignores effects occurring outside a beneficiary's county of residence* […]. *None of the approaches described, neither these focussed on the participant's future nor these concentrating on the local labour market, are ideal, as they do not present the full picture. They are, however, complementary and, combined, allow for a much more complete and multi-faceted evaluation of the intervention* (*Metaanaliza wyników badań ewaluacyjnych… Raport cząstkowy…*, 2018, pp. 253–254).

28 Cf. www.ewaluacja.gov.pl/media/79609/Wplyw_polityki_spojnosci_na_rozwoj_spoleczno-gospodarczy_Polski_i_regionow_ w_latach_2004-2018.pdf [accessed: 15 September 2019].

1. *When Poland joined the EU, it had the highest unemployment rate (19.1%). In subsequent years, it fell rapidly to 3.8% in 2018. At that point, it was 3% lower than the EU-28 average* [...]; *and 1% lower due to EU funds. In absolute terms, this means a decrease of the number of the unemployed by more than 400,000* [...];
2. *Compared to 2004, a per capita income in comparison to the EU average increased by 19.3%*[29].

Initially, such works were connected with relatively detailed requirements of the European Commission (EC) and concerned extensive assistance measures (entire perspectives of the NDP/CSF 2004–2006, NCS/NSRF 2007–2013). In recent years (2015–2019), macroeconometric literature – in particular commissioned by PARP, the Educational Research Institute or selected ministry departments, which act in the capacity of Operational Programme Managing Authorities (SG OP, EP OP) as an element of evaluation of a given programme – is also starting to focus on selected target areas of fund utilisation. These areas are e.g. social and economic cohesion of the Eastern Poland macroregion, impact on the competitiveness and innovativeness of economy, and impact on the labour market. Aside from the key issue of the demand for modelling results for cognitive and practical purposes (e.g. the implementation of a formative, conclusive or socio-political function of evaluation in the process of programme or policy management), an essential prerequisite for using macromodelling tools when evaluating the impact of a programme should be access to wide public statistics (cross-sectional, long time series) and, above all, a substantial budget for the development activities under analysis. The amount of planned public resources must be large enough to have a real chance to impact (shock) the recorded macroeconomic values (total Polish GDP or its y/y dynamics, employment/unemployment rate, etc.).

Various modelling tools are used in evaluation. These include both relatively simple single equation econometric models, and advanced multi-equation and general equilibrium models, which assume the reflection of market processes occurring in the whole economy and their interaction, taking into account the dynamics and uncertainty of market economy phenomena. The most advanced tools of this type are DSGE (dynamic stochastic general equilibrium) models (cf. Mogiła et al., 2019; *EduMod...*, 2015)[30]. From the evaluation perspective, the key function of a model is *to assess the impact that exogenous economic changes and public interventions, as set out by the model user, will have on all modelled variables* [...] [a purely evaluative function – author's note]. *An additional purpose of a model is to make conditional forecasts of the future social and economic situation in Poland and to identify factors behind the economic phenomena observed in the past* (foresight function) – cf. *EduMod...*, 2015, p. 11. Referring to the EC's guidelines concerning the 2000–2006 and 2007–2013 perspectives, we can expect robust macroeconomic models to:

1. reflect the demand and supply structure of the economy, and in the case of the latter to include public expenditure related to the analysed policy or programme;
2. include a selected list of socio-economic indicators to be evaluated, as well as relevant explanatory variables;
3. include various functions of counterfactual/scenario analysis and sensitivity, making it possible to simulate how the value of effect indicators changes after the modification of shock parameters (e.g. outlays on education policy, macroeconomic changes), how these indicators behave with and without intervention, and how they would develop in the case of an alternative allocation

29 Adjusted real gross disposable income of households per capita in PPS. Cf. ibidem.

30 The systematics and properties of individual model types can be found in literature and in various methodological reports devoted to individual models.

of resources (alternative costs) (cf. Zaleski, 2019; T*he Ex-Ante Evaluation...*, 2000; *Indicative Guidelines...*, 2006).

On the other hand, models are inherently designed to simplify a complex economic reality. Defining them in an overly comprehensive and, at the same time, overly detailed way based on thousands of equations or processes of assumption parameterisation (estimating as well as calibrating, which are not devoid of the arbitrariness of modelling teams), which are difficult for users to understand, may discourage units responsible for shaping regional or sectoral policy from reaching for them and, as a result, limit the use of macroeconomic simulations in setting directions for intervention. In literature, we can find many critical opinions about the most advanced general equilibrium models used by international organisations and leading financial market think-tanks to analyse the business climate and the economic impact of policies. They are often accused of unrealistic assumptions – such as perfect competition on the market, rapid adaptation of prices to changes in supply/demand, balance in access to information or model homogeneity of businesses – which may result in detachment from the analysed economic reality. The consequence of scepticism towards these models was the "reform movement" leading to the modification of the model (purely theoretical) parts of the assumptions. This resulted in a stronger reliance on microeconomic fundamentals, but also in balancing – provided by observation models – with evidence from a microeconomic perspective (quantitative and qualitative data obtained from entities, analyses in experimental and quasi-experimental schemes). Above all, it was agreed that in reality the market is not ideal, and various asymmetries and frictions can be observed there. Then, attempts were made to implement them in DSGE models.

> *Frictions in the labour markets seem to be crucial from the point of view of the economy. An efficiently functioning labour market forms the basis of economic development, and its condition translates into parameters of the economy, as well as the condition of employees and households* [...]. *Such contemporary DSGE models, where frictional labour markets are linked by a range of political conditions,* [...] *are based on microeconomic assumptions* [...] *of the preferences and decisions of consumers and business, i.e. participants in the economic system* [...]. *The mechanism proposed by Peter Diamond from the Massachusetts Institute of Technology, Dale Mortensen from Northwestern University and Christopher Pissarides from the London School of Economics was introduced here* [...]. *An axial element of the theory is the assumption of the existence of costly frictions on the labour market, which delay the process of matching a job seeker with a corresponding vacancy.* [...] *The second module of the frictional labour market model are wage negotiations based on the Nash bargaining solution borrowed from game theory.* [...] *During the negotiations, the participants in the process divide a certain surplus among themselves, which in the case of a job seeker becomes their salary, and in the case of an employer – the profit of the company. Cooperation continues as long as it is profitable for both parties.* [...] *In this case, households actively participate in the labour market: they provide work in return for remuneration. This takes place in a dynamic environment, which is exposed to external economic shocks* (Woźniak, 2019, pp. 131–148).

Considering these labour market interdependencies in advanced general equilibrium models allows for a more reliable and accurate analysis of the market and interventions directed at it. In particular, this is possible thanks to taking into account the specificities of human capital in the model, as well as the mechanisms which it is actually subject to, and how it responds to external "shocks" – both to direct

instruments of active labour market policies, and to passive actions or interventions concerning the production potential of employers, which indirectly affect labour.

From among few Polish evaluation studies in which the DSGE model was used in relation to frictional labour markets it is worth mentioning analyses at the sectoral and macroeconomic level concerning the effects of selected actions of operational programmes for the years 2014–2020, i.e. Smart Growth and Eastern Poland[31]. The use of the VESPA model allows for analysing the impact of OP SG and OP EP actions on the labour market: employment, unemployment, sectoral structure of employment, as well as the level of wages and labour productivity. It considers conditions where *the labour market simultaneously includes unemployment, employment and professional inactivity, and the model can simulate all flows between these situations in the labour market (e.g. the stream of people becoming active and deactivated in a given research period, as well as the flow of unemployed people finding a job and people losing employment). This makes it possible to estimate both inflows and outflows from unemployment, employment and inactivity and the net number of jobs created.* [...] *The parameters of the labour market module were determined on the basis of actual historical data, which makes it possible to reflect the differences between the flexibility of the labour market in individual regions of Poland and to differentiate the impact of Programme funds spending on labour markets in individual regions* (cf. *Analiza efektów...*, 2019) and by business activity (PKD) sectors.

An analogous evaluation study by PARP carried out with the use of the previous generation of the VESPA model focussing on macroeconomic effects of OP IE 2007–2013 in the area of indirect impact on the labour market and competences provided the following results:

- → *The effect* [...] *of measures taken by PARP under OP IE is an annual increase in employment in the analysed period (2007–2025). The number of people working in the economy increases thanks to the support by 0.16% on average. In the period 2007–2017, each year on average 45,000 more people were employed than if the intervention had not been implemented. The largest increase in the number of employees was observed in the last years of the intervention and in 2016 and 2017. After 2017, a negative deviation from the baseline scenario is forecast (phasing out of intervention effects), but it is insignificant.*
- → *PARP's interventions also caused a slight increase in labour productivity (on average +0.03% for the entire period) and a growth in wages (on average +0.08% for the entire period). The simultaneous growth of these two indicators shows a reorientation of the economy towards ever greater technological advancement*[32] [...].
- → *As a result of support for R&D&I activities, the productivity of Polish enterprises increased in the long term, due to the dissemination of knowledge and industrial commons*[33]. [...]

To sum up, [...] *labour productivity, company profits and wages increased. The main reason for this was increase in capital-labour ratio (improvement in capital productivity in treated companies and, more broadly, in the economy as a whole). As a result, more skilled jobs were created and an additional stimulus*

31 Evaluation commissioned by PARP in the years 2019–2022 is carried out by the consortium WiseEuropa – Warsaw Institute of Economic and European Studies Foundation and Ecorys Polska.

32 High-tech economies create many well-paid jobs (characterised by high productivity).

33 Knowledge base and (technical, design and operational) capabilities common for a given industry sector, such as know-how, R&D&I, advanced process development and engineering skills or production competence related to a specific technology.

> *for the accumulation of human and organisational capital in companies (modern know-how) emerged* (*Analiza wybranych działań...*, 2017, pp. 11–12)[34].

These conclusions and various valuable recommendations, which were not mentioned here, were formulated by complementing data from the model with qualitative research with the participation of entrepreneurs (microeconomic perspective) and experts in macroeconomics. The conducted interviews indicated spillover effects of assistance provided to entrepreneurs which affected a wider group of economically related organisations, households and individuals as a key success factor – cf. Bukowski & Pokorski, 2019. They occurred in particular in R&D and innovation projects carried out by companies.

Another example of a valuable combination in evaluation studies of the micro- and macroeconomic as well as quantitative and qualitative perspective was provided by evaluation carried out by PARP in 2016. It concerned the impact of the Operational Programme Development of Eastern Poland 2007–2013 (DEP OP) on the development of entrepreneurship in the Eastern Poland macroregion[35]. Although labour market issues were not brought to the fore in that study, indicators concerning the quantity and quality of local labour market resources were included in the macro- and microeconometric models used in it (PSM and Stratified-PSM quasi-experimental schemes – cf. Wojtowicz & Widła-Domaradzki, 2017). Evaluation was carried out in accordance with the methodological approach of counterfactual impact evaluation and theory-based evaluation (cf. Hermann-Pawłowska et al., 2017). At the same time, the counterfactual component was filled in synergistically by micro- and macroregional perspectives (reference to changes observed in the macroregion in the modelled situation without the support of DEP OP). In this case, macroanalyses *were performed using econometric modelling (panel models for data at the level of provinces; cross-sectional linear and spatial models at the level of subregions and cross-sectional models at the level of municipalities and poviats). The analyses were carried out on the basis of absolute values and standardised variables obtained from the Statistics Poland and the KSI SIMIK 07-13 system* [...]. *For the best suited models, an analysis of the remaining models was carried out on a sample of territorial units which were beneficiaries of the DEP OP in order to indicate which units were more effective in stimulating entrepreneurship using funds from the DEP OP than the estimated model would indicate. In addition, based on the models at the sub-regional level, the net impact of axis 1 of DEP OP on the economy of individual sub-regions was estimated* (*Ewaluacja wpływu Programu Operacyjnego Rozwój Polski Wschodniej...*, 2016, p. 29)[36].

Considering that advanced modelling of the labour market – in particular using DSGE macromodels, which take into account Mortensen and Pissarides's friction mechanism – is a relatively new methodological achievement in economic research, the author is not familiar with its application in evaluation studies strictly oriented at supporting the development of human capital in Poland (e.g. OP HC 2007–2013, OP KED 2014–2020, Individual Funding Systems for Development Services under Regional Operational Programmes 2014–2020). The issues of the labour market and employment promotion were usually only an element of macroanalyses of a wider scope (e.g. the impact of the use of European funds in Poland under all operational programmes for 2007–2013). Evaluation studies of ESF

34 Cf. www.parp.gov.pl/component/publications/publication/analiza-wybranych-dzialan-po-ig-na-poziomie-sektorowym-i-makroekonomicznym-za-pomoca-modelu-przeplywow-miedzygaleziowych [accessed: 15 September 2019].

35 The evaluation was commissioned by PARP and carried out in 2016 by a consortium of Eval, WYG PSDB and EGO companies.

36 Cf. www.parp.gov.pl/storage/publications/pdf/20190211135235sbsjd.pdf [accessed: 15 September 2019].

support, although often very advanced[37], have so far focussed mainly on assessing the effectiveness of the support and its impact from a micro perspective (changes in the situation of people in the labour market and not changes in the market situation as a result of the programme). Although some modelling tools are in place in this sector[38], they have not been widely used in evaluation studies. In the future, complementing this perspective would certainly provide valuable conclusions and recommendations for labour market policies and the use of the ESF. This also remains a challenge for ex-post evaluation of the use of EU funds in the 2014–2020 perspective.

Conclusions

In Poland, from the methodological point of view, the practice of carrying out evaluation of the impact of competencies and of employment promotion support is highly advanced. It often provides a great deal of valuable evidence and recommendations concerning the effectiveness of implemented labour market policy instruments. Moreover, it does not differ from the practice of leading European centres conducting this type of research work. However, in evaluations of support provided to competencies development (especially financed by the ESF under EU cohesion policy), among sophisticated counterfactual techniques a microeconomic perspective predominates, focussing on changes that primarily affect individuals. As shown in this chapter, combining the perspective of employees and of the market in evaluation can provide a more complete picture by showing how much change at the level of numerous participants in a given programme (individuals) actually responds to a social need (the reason for the programme) or solves a structural problem for the national labour market. The macroeconomic perspective, going beyond direct participants of support programmes, can also capture the flow of values from the participant to the market, as well as verify a possible spillover of support effects onto the results of employers and incomes of households, and, as a result, onto the socio-economic condition of the whole country. As has been shown, both approaches have strengths, but also involve numerous limitations. Thanks to more than a decade of applying them in Poland, they are well identified, especially in counterfactual evaluations of microeconomic impact. This means that individual evaluations may lead to erroneous or incomplete judgements of reality and estimates of values resulting from the implementation of a given public programme. Following the example of methodological triangulation, well-established in evaluation studies, combining participant-oriented (micro) and market-oriented (macro) research perspectives should make it possible to emphasise the strengths of each – for the benefit of evaluation recipients (policy-makers) – and at the same time to reduce doubts concerning specific perspectives (approaches) and related research and analytical techniques.

37 Cf. counterfactual evaluations of the impact of ESF support carried out using ZUS microdata concerning payers and the insured (cf. *Badanie efektów wsparcia...*, 2018; *Metaanaliza wyników badań ewaluacyjnych...*, 2018), evaluation of causal relative effects of post-graduate studies financed by the ESF (cf. Trzciński, 2013; *Efekty dofinansowanego kształcenia...*, 2014).

38 For example, *experts from the Educational Research Institute have prepared a structural simulation and prognostic model of the Polish economy. It takes into account phenomena related to education, adult education, operation of the training market, demand of particular professional groups, etc. This model is intended to simulate the socio-economic impact of public policies, with particular emphasis on educational choices and labour market policies*. Cf. eduentuzjasci.pl/publikacje-ee-lista/inne-publikacje/1230-edumod-model-symulacyjno-prognostyczny-polskiej-gospodarki-raport-podsumowujacy-projekt.html [accessed: 15 September 2019].

Bibliography

Analiza efektów netto wybranych działań Programu Operacyjnego Innowacyjna Gospodarka 2007–2013, z wykorzystaniem podejścia counterfactual impact evaluation w ramach projektu ewaluacyjnego PARP "BAROMETR INNOWACYJNOŚCI" (2015). Warsaw: Uniwersytet Jagielloński, EGO, commissioned by Polska Agencja Rozwoju Przedsiębiorczości.

Analiza efektów wybranych działań POIR i POPW na poziomie sektorowym i makroekonomicznym za pomocą modelu makroekonomicznego. Raport metodologiczny (2019). Warsaw: WISEEuropa & Ecorys Polska, commissioned by Polska Agencja Rozwoju Przedsiębiorczości.

Analiza wybranych działań PO IG na poziomie sektorowym i makroekonomicznym za pomocą modelu przepływów międzygałęziowych (2017). Warsaw: WISEEuropa & OPI PIB, commissioned by Polska Agencja Rozwoju Przedsiębiorczości.

Badanie efektów wsparcia zrealizowanego na rzecz osób młodych w ramach Programu Wiedza Edukacja Rozwój. III Raport wskaźnikowy (2018). Warsaw: IBS, IMAPP & IQS, commissioned by the Ministry of Investment and Development.

Blanchet, T. et al. (2019). *How Unequal is Europe? Evidence from Distributional National Accounts, WID.world WORKING PAPER No. 2019/06*. Paris: World Inequity Lab.

Bukowski, M. & Pokorski, J. (2019). Polityka innowacyjności i konkurencyjności w świetle modelu VESPA 3. In: Z. Mogiła et al. (eds.), *Wykorzystanie modelowania ekonometrycznego w ewaluacji wpływu polityk publicznych i programów*. Warsaw: Polska Agencja Rozwoju Przedsiębiorczości.

Common methodology for State aid evaluation (2014). Brussels: European Commission.

EduMod: Model symulacyjno-prognostyczny polskiej gospodarki uwzględniający zjawiska związane z kapitałem ludzkim (2015). Warsaw: IBS, commissioned by the Educational Research Institute.

Efekty dofinansowanego kształcenia w grupie wiekowej 50 plus (2014). Warsaw: Fundacja Instytut Społeczno-Ekonomicznych Ekspertyz, commissioned by Polska Agencja Rozwoju Przedsiębiorczości.

Ewaluacja wpływu Programu Operacyjnego Rozwój Polski Wschodniej 2007–2013 (PO RPW) na rozwój przedsiębiorczości w Polsce Wschodniej (2016). Warsaw: Evalu, WYG PSDB & EGO, commissioned by Polska Agencja Rozwoju Przedsiębiorczości.

Hermann-Pawłowska, K. et al. (2017). *Ewaluacja oparta na teorii w złożonym otoczeniu społeczno-ekonomicznym*. Warsaw: Polska Agencja Rozwoju Przedsiębiorczości.

Indicative Guidelines on Evaluation Methods: Ex Ante Evaluation (2006). Brussels: European Commission.

Mackiewicz, M. (2019). *Wpływ polityki spójności na rozwój społeczno-gospodarczy Polski i regionów w latach 2004–2018*. Zespół Sterujący Ewaluacją Polityki Spójności (26 September 2019). Gdańsk: Ministerstwo Inwestycji i Rozwoju.

Metaanaliza wyników badań ewaluacyjnych dotyczących oceny wsparcia z Europejskiego Funduszu Społecznego (2018). Warsaw: Evalu, commissioned by the Ministry of Investment and Development.

Metaanaliza wyników badań ewaluacyjnych dotyczących oceny wsparcia w ramach EFS – –Raport cząstkowy 2018 (2018),. Warsaw: Evalu, commissioned by the Ministry of Investment and Development.

Mogiła, Z. et al, (2019). *Wykorzystanie modelowania ekonometrycznego w ewaluacji wpływu polityk publicznych i programów*. Warsaw: Polska Agencja Rozwoju Przedsiębiorczości.

Pokorski, J. (2016). *Relative Causal Effects in Evaluation Studies of Public Programmes – – Some Inspiration from PARP's Evaluation of Pro-Innovative Business Support*. 12th Biennial Conference – Evaluation Futures in Europe and beyond Connectivity, Innovation and Use. Maastricht: European Evaluation Society.

Pokorski, J. (2018). *Badania wykorzystujące dane rejestrów publicznych (inwentaryzacja). Raport*. Warsaw: IBRKiK-PIB.

Programming period 2014–2020: Monitoring and Evaluation of European Cohesion Policy – – European Social Fund. Guidance document (2015). Brussels: European Commission.

Rogalińska, D. et al. (eds.) (2015). *Statystyka dla polityki spójności*. Warsaw: Statistics Poland.

Tarsa, P. (2019). *Metaanaliza wyników badań ewaluacyjnych dotyczących oceny wsparcia w ramach EFS*. Grupa Sterująca Ewaluacją POWER (26 September 2019). Gdańsk: Ministerstwo Inwestycji i Rozwoju.

The Ex-Ante Evaluation of the Structural Funds interventions, The New Programming period 2000–2006: methodological working papers (2000). Brussels: European Commission.

The Programming Period 2014–2020: Guidance document on monitoring and evaluation – European Regional Development Fund and Cohesion Fund (2013). Brussels: European Commission.

Trzciński, R. (2009). *Wykorzystanie techniki propensity score matching w badaniach ewaluacyjnych*. Warsaw: Polska Agencja Rozwoju Przedsiębiorczości.

Trzciński, R. (2013). *Efekty (netto) dofinansowanych studiów podyplomowych*. Warsaw: Polska Agencja Rozwoju Przedsiębiorczości.

Ustalenie wartości wybranych wskaźników ekonomicznych dla beneficjentów Regionalnych Programów Operacyjnych (RPO) i Programu Operacyjnego Innowacyjna Gospodarka (POIG) 2007–2013 oraz dla dobranych grup kontrolnych (2015). Warsaw: Statistics Poland.

Ustalenie wartości wybranych wskaźników ekonomicznych dla beneficjentów Programu Operacyjnego Kapitał Ludzki 2007–2013 (POKL) i dla dobranych grup kontrolnych (2017). Warsaw: Statistics Poland.

How to analyse sustainability and long-term effects of transnational learning mobility

DOI: 10.47050/66515246.108-117

Jadwiga Fila
Michał Pachocki

Abstract

This chapter presents the assumptions underlying tracer studies of educational mobility beneficiaries – former HEI and vocational school students. Such studies were one of the methods of evaluating the impact of European educational programmes managed by the Foundation for the Development of the Education System in the years 2007–2016. As part of the studies, some of the project participants were reached and asked about their opinions on completed mobilities. They also shared information about their current professional situation and what they had been guided by when making choices related to work and further learning.

Introduction

The Foundation for the Development of the Education System (FRSE) has managed European educational programmes for over thirty years. During this time, the Lifelong Learning and Erasmus+ programmes have supported more than 150,000 participants in work placements and more than 200,000 students in undertaking studies and placements abroad. The great popularity of these programmes is related to the high level of EU funds allocated for international learning mobilities of Polish students. It is, therefore, reasonable to investigate how to evaluate not only the quality and sustainability of supported mobilities, but also their long-term effects. *Ex post* evaluation evaluation on VET and HE schools' graduates with learning mobility experience should help to provide such information. .

This chapter presents an example of the application of cross-sectional and longitudinal studies in the area of graduate tracer studies. Taking their limitations into consideration, these methods have proven to comply with the requirements and assumptions for undertaking a deferred evaluation. When carrying out these research projects, FRSE for the first time made an attempt to evaluate the impact of mobility experience on participants' educational and professional choices. The graduate tracer studies made

Jadwiga Fila
Graduate of sociology at the Jagiellonian University, where she specialised in social research and data analysis. Researcher of the Analytical and Research Team at the Foundation for the Development of the Education System. She specialises in educational research, recently mainly in the area of higher education and international mobility. She is the author of qualitative and quantitative research, and co-author of research reports, articles and other works published in Poland and abroad.

KEYWORDS:
LEARNING MOBILITY, GRADUATES' CHOICES, HIGHER EDUCATION, VOCATIONAL EDUCATION AND TRAINING, PANEL RESEARCH

Michał Pachocki
Graduate of modern languages and cultural sciences, as well as of the School of Evaluation at the Faculty of Sociology, University of Warsaw. Member of the Polish Evaluation Society and a doctoral student at the School of Social Sciences of the Polish Academy of Sciences. He worked in development programmes carried out by public and non-governmental organisations. Since 2009, he has worked at the Foundation for the Development of the Education System, where he conducts research on the impact of international mobility on the Polish education sector. He has published in the following periodicals: *Cogito*, *Dyrektor Szkoły*, *Kształcenie Zawodowe* and *Języki Obce w Szkole*.

it possible to reach out to some of the project participants and not only learn their opinions about the completed learning mobility, but also find out more about their current situation and career paths. Their feedback also contributed to identifying the profile of former participants, their professional motivations and other reasons behind their choices related to work and further study.

Key research areas and methodological approach

The study focussed on graduates of vocational schools and higher education institutions who during their studies benefited from transnational learning mobilities financed as part of European projects. In the case of graduates of vocational and secondary technical schools, the mobilities consisted in internships and apprenticeships in companies and training centres abroad, while in the case of students, support consisted in undertaking part of their studies at universities and serving internships or placements at companies abroad. The study was anonymised, and the data were collected mainly via online surveys. In the case of the tracer study of participants in traineeships in the VET sector, individual and group interviews were also conducted.

The respondents took part in the following educational programmes:

- → Lifelong Learning Programme, which was implemented in the years 2007–2013 and targeted, among others, HEI (Erasmus) and VET (Leonardo da Vinci) students;
- → Erasmus+, which in the years 2014–2020 continues actions taken as part of the previous EU funds programming phase (support for HEI and VET students' mobilities);
- → Systemic projects financed by the European Social Fund, which offer internships and placements abroad for students in accordance with the principles of the Leonardo da Vinci (2007–2013) and Erasmus+ programmes (since 2014).

The respondents were asked to evaluate their mobilities in their entirety and to assess potential long-term benefits of participating in the programme. This was to allow for analysing their opinions on the usefulness of foreign learning mobilities for their further education and professional careers. The main focus was on the evaluation of the intervention in retrospect, as well as on the situation of the participants, with particular emphasis on the changes that took place in their lives in terms of education and career progression. When developing research assumptions and tools, not only purely vocational skills acquired in the framework of formal

education, but also soft skills obtained in the process of non-formal education as part of the agreed vocational mobility programme were taken into consideration.

Table 1. Target groups of former mobility participants divided by European programme (survey of VET students)

	Lifelong Learning Programme – Leonardo da Vinci	Erasmus+ Vocational Education and Training
Timeframe	2007–2013	2014–2016
Minimum duration	2 weeks	
Maximum duration	39 weeks (approx. 9 months)	1 year (12 months)
Complementary ESF support under scrutiny	2012–2013 (as part of OP HC)	2014–2016 (as part OP KED)

Source: authors' own work.

The main substantive areas of research concerned specific competences that former trainees and apprentices acquired, as well as their usefulness on the labour market and the impact international mobility had on their careers and personal development. In the case of the tracer study of HEI graduates, such issues as continuing education at subsequent cycles of studies, taking up studies in other programmes, as well as serving internships and placements at further stages of education (both obligatory and not required by the study programme) were also important. Potential impact of completed mobilities on making decisions concerning subsequent mobilities for work, studies or placements funded from sources other than the European programmes, was also investigated.

Table 2. Target groups of former mobility participants divided by European programme (survey of HEI students)

Year of study	Tracer study – analysed group	Panel research – analysed group
2017	Mobility participants in the years 2007–2015 (Erasmus and Erasmus+)	---------
2018	Mobility participants in the years 2016–2017 (Erasmus+)	Graduated in: 2017, 2015, 2013
2019	Mobility participants in 2018 (Erasmus+)	Graduated in: 2018, 2016, 2014

Source: authors' own work.

Research methods and tools

In the case of the tracer study of HEI graduates, the research scope was very broad, so the decision was made to apply two interrelated research schemes in parallel.

The first is **a cross-sectional study** (called **a tracer study**). Each year, it covers new groups of respondents who meet the preconditions for participation in the study. The prerequisite to be included in the tracer study is meeting both of the following criteria:

→ mobility must have ended in the studied year; and

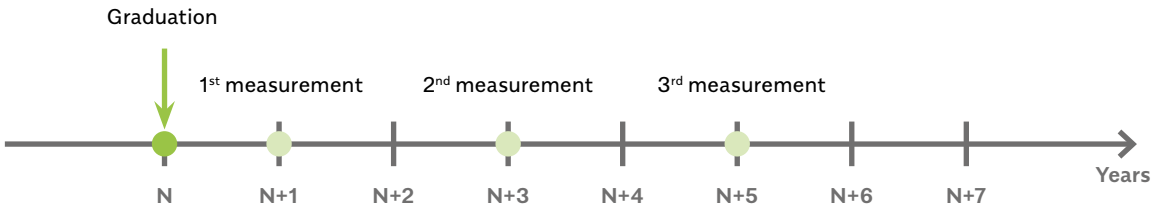
→ students must have graduated (completed that stage of their studies at which the mobility took place).

The study addresses a comprehensive sample, i.e. everyone who has benefited from an international mobility for studies and/or placements under the Erasmus+ programme. The questionnaire is sent out periodically, once a year (usually in November[1]) to people who have completed mobility in the year selected for analysis. Respondents' data come from the programme database available to FRSE, which is the National Agency for the Erasmus+ programme in Poland. Upon return, every mobility participant is obliged to fill in a Participant Report, where he/she agrees to be contacted at a later date and provide contact details . This is where the email addresses to which the questionnaire is sent come from.

The second scheme is **a longitudinal study** (called **a panel research**), where the same individuals are surveyed (the so-called panel of respondents) several times, at intervals determined in the methodology. The main feature of the panel research is that the same group of respondents with specific characteristics is surveyed at fixed intervals. This means that from time to time the same people fill in a questionnaire, which in each subsequent measurement is very similar or identical in form in order to ensure comparability of results. Such data allows to determine for each respondent the direction and dynamics of changes taking place in their opinions and attitudes. This also makes it possible to trace the dynamics of changes in a given group and to identify cause and effect mechanisms associated with observed phenomena. Furthermore, it will be possible to verify if any long-term effects of international mobility can be identified. The aim of the panel research is to analyse in detail the professional and educational choices of mobility participants who have completed at least first-cycle studies. The main focus is on how, in the opinion of the respondents, mobility under the Erasmus+ programme translated into the development of competences, whether it impacted their educational and professional choices, and whether it helped them to enter the labour market.

The measurements are being made three times: one year, three years and five years after graduation. In this case, the date of receiving a certificate of completion of the studies during which the respondents participated in mobility is considered as the time of graduation. The same research will cover the graduates of all cycles of studies.

Figure 1. Diagram illustrating individual measurements as part of panel research (HEI graduates)



Source: authors' own work.

1 Administering the survey in November makes it possible to include in it all students who completed their studies in a given year, including those who received their certificate of completion after the summer break.

The multitude of factors affecting professional and educational decisions of the respondents over a period longer than five years weakens impact of international mobility. Therefore it was decided to introduce such intervals in the survey.

Tracer studies and panel research are a continuous process. Tracer study measurement is the initial stage for panel research measurement (respondents give their consent to participate in subsequent stages, i.e. the panel research, and provide their e-mail address to be contacted later on).

Table 3. Research plan followed in the analysis of HEI graduates

	Tracer study	Panel research
Type of measurement	Cross-sectional measurement aimed at including in the research group successive graduate cohorts during successive editions of the study	Longitudinal study covering the same group of graduates at specific time intervals
Respondents	Prerequisites for the respondents:	
	→ They carried out a mobility in the year subject to analysis; and → They completed the studies during which the mobility took place	→ They took part in the tracer study; and → They gave their consent to participate in the panel research and provided their e-mail address for further contact
Research method	Online survey (CAWI*)	
Type of questionnaire	Basic questionnaire	Panel research questionnaire
Time of measurement	Every measurement year, around November	One, three and five years after the graduation of the given group of students

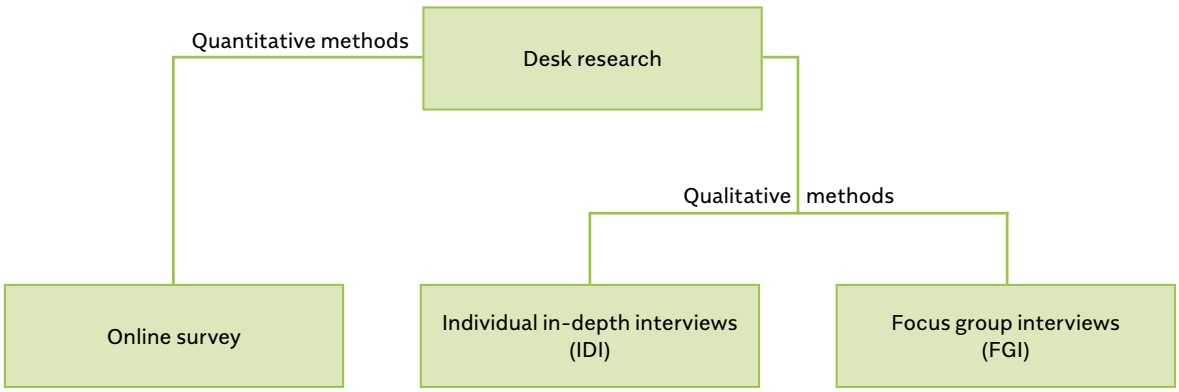
* CAWI – computer-assisted web interview.

Source: authors' own work.

In the two abovementioned research schemes, the method of data collection is an online questionnaire (computer-assisted web interview – CAWI). The tracer study requires the use of a slightly more extensive questionnaire (the majority of questions concern educational choices, mobility assessment and entering the labour market) than in the case of the panel research. The panel research allows to fill in a shorter questionnaire concerning the professional situation of the respondent and assessing mobility after a longer time. In both cases, the data collected are mostly quantitative, although the questionnaires contain several open questions allowing for longer, more in-depth answers.

Both quantitative and qualitative techniques were used to monitor former participants of internships and placements abroad. Quantitative techniques were used to administer the questionnaire (online survey), while qualitative methods concerned individual in-depth (IDI) and focus group (FGI) interviews. At all stages of the study, participants were asked both to evaluate the intervention itself and to assess the usefulness of the acquired competences at a later stage of their educational and vocational path. The results obtained were also complemented by "existing data" (mainly contextual information and statistical data reflecting the stage of implementation of analysed European educational programmes).

Figure 2. Research scheme concerning VET graduates



Source: authors' own work.

The triangulation of adopted tools resulted from differing characteristics of the two respondent groups (vocational school students and graduates), while the triangulation of the planned research methods primarily aimed at collecting in-depth information on experiences related to mobility and its effects.

Research implementation

The tracer study of participants in apprenticeships and internships in the VET sector was carried out from March to September 2017. It included both quantitative and qualitative methods.
In the quantitative part, an online questionnaire was used as a basic research tool. It was sent out to the e-mail addresses of former participants of internships and traineeships funded under the programmes covered by the study. A link to the questionnaire was also sent to the schools with the request to distribute it among former students and mobility participants still being educated. The quantitative part of the study covered the entire population of former internship and traineeship participants, and the only criterion for being included in the survey was participation in mobility abroad carried out under the programmes covered by the study. In effect, FRSE received more than 2,600 anonymously filled questionnaires, and a large majority of respondents met the study criteria. The respondents most often served internships at later stages of their education at school (the questionnaire was completed by less than 3% of those who left for mobility in the first year of study) – more than half of the respondents were people who participated in mobilities in their third year and a significant percentage was formed by those who carried out mobility in the last year of study.

The respondents of individual in-depth interviews were school graduates who participated in transnational internships and traineeships organised as part of the programmes. In total, 20 interviews were held with people living in different regions of Poland. Meetings were arranged in the place of residence, the workplace or the school of former trainees. A randomly selected sample was varied in terms ofgender, fields of training, vocational qualifications and target countries of mobilities.

In the framework of the study, five focus group interviews were also conducted with students who participated in transnational mobilities. The study was carried out in five schools actively implementing educational mobility. A total of 65 students of the second, third and fourth year of technical upper secondary schools and of the second year of basic vocational schools were interviewed. The focus group interviewees also included recent graduates who had completed their education in the school year preceding the study.

In the case of the tracer study of HEI graduates, three editions of the study and two waves of the panel research have been conducted so far. The first edition took place in 2017 (it covered participants in mobilities carried out from 2007 to 2015[2]), the second in 2018 (graduates who benefited from mobility programmes in 2016 and/or 2017), and the third started in November 2019.

The first wave of the panel research took place in 2018 and covered a group of 1,317 former mobility participants who, at the time of conducting the research, had graduated one, three and five years earlier, respectively. The questionnaire was answered by 45% of those invited to take part in the study. The distribution of the sample was not even, which is understandable. Those who have just graduated are more willing to take part in such a study, as their mobility experience is fresh in their memory. In accordance with the methodology, no weights were used in this sample, as the study was about obtaining "raw data" to observe trends within individual cohorts[3], as well as within groups which graduated in a given year (change of attitudes and opinions). Thus, the data obtained can be analysed both by cohorts and by a specific graduation year.

In order to ensure anonymity of the respondents and continuation of the measurement, individual respondent identifiers (tokens) were used in the study. Thanks to using specialist software, it was possible to identify individual respondents without the need to store their personal data. At the analysis stage, the identifier made it possible to combine responses from three panel measurements. Such a solution also allowed to monitor the number of questionnaires collected and for sending e-mails with survey reminders. The second wave of the panel research is under way. The questionnaire was sent out to respondents in December 2019. The first results from this wave of measurements will be analysed in the first quarter of 2020.

Publications of results

The tracer study of VET graduates who completed internships and apprenticeships abroad has already been completed and its results have been presented in the report entitled *Is mobility the key to a successful career?* published in 2018 in Polish and English (Pachocki, 2018)[4]. The report presented a summary of the most important results based on quantitative and qualitative data. The methodology developed for this study is also used in an international study, which has been carried out since the beginning of 2018. The international tracer study of graduates has allowed similar research activities

2 These were beneficiaries of both the Erasmus programme (carried out until 2013) and its successor, Erasmus+, who completed mobilities for studies and/or placements/internships and have already graduated from their studies. This study was preceded by a pilot study to test and develop the tool.

3 The cohorts included are groups of graduates who completed their studies: one year earlier, three years earlier, and five years earlier.

4 See: czytelnia.frse.org.pl/mobility-key-successful-career [accessed: 20 January 2020].

to be carried out in nine European countries[5], and FRSE will publish an international report on the study at the end of 2020.

The first results of the cross-cutting tracer study were published in 2019 in the report entitled *Erasmus... and what next? Tracer study on education and careers of Polish students participating in Erasmus mobility projects*. It features an analysis of data coming from more than seven thousand respondents who carried out international mobility in the years 2007–2015. The publication is available in print and on the FRSE website[6] (Dąbrowska-Resiak, 2019).

In 2020, for the first time, a repeated measurement will take place among the same group of participants in the panel research (they were previously analysed in 2018). It will cover students who graduated in 2017 and 2015, who will be three and five years after graduation, respectively. When the data are collected, the first analyses as part of the panel research will be possible, as this research scheme requires at least two measurements on the same group of participants. More complete and complex analyses will only be possible in 2022, when we will have acquired measurements from three waves of the panel research for the first group of respondents.

Limitations and potential risks

Longitudinal research is an interesting method that provides sample opportunity for data analysis in the long term. At the same time, like any test method, it involves certain limitations. Below, the authors have identified key risks and suggested solutions to help minimise them.

Reaching the respondents was the most important limitation related to the research. This was of particular importance in the case of quantitative research, where, in theory, the target group was the entire population of former mobility participants. As the primary source of information about the participants were individual reports filled in by them immediately upon their return from mobility, in many cases it turned out that contact details contained therein were out of date and it was not possible to obtain new e-mail addresses. In the case of the tracer study of former participant in internships abroad, similar limitations accompanied the process of selecting participants in individual in-depth interviews. However, it should be stressed that because of a much smaller sample and intentional selection of respondents, finding members of this group proved to be considerably easier than in the case of respondents of the questionnaire.

In the case of the panel research, the main difficulty was to maintain the respondents' motivation to participate in the three measurements. In each of them, the questionnaire is quite similar in form, which may tire the participants and make them reluctant to continue participating in the study. This leads to a phenomenon called "panel attrition", which consists in people gradually dropping out of the group for various reasons, often out of researchers' control. This is a natural phenomenon in the case of longitudinal studies, and a few percent reduction in the research group in subsequent measurement waves is an effect that is difficult to prevent. In order to minimise its scale, the number of questions in subsequent waves of the study was reduced to the bare minimum and a two-year measurement interval was applied.

5 The international survey was carried out by National Agencies of the Erasmus+ programme in Austria, Belgium, The Czech Republic, Macedonia, Ireland, Luxembourg, Latvia, Slovakia and The United Kingdom.

6 See: czytelnia.frse.org.pl/erasmus-and-what-next [accessed: 05 June 2020]

Another limitation related to keeping the respondents in the panel is the risk of a participant changing their e-mail address and the consequent loss of contact with them during the study. In order to reduce this risk, participants were asked to provide their private e-mail addresses, assuming that a private address is less frequently changed than a school or business one. Lack of participants' current e-mail addresses may therefore constitute a serious obstacle in implementing the survey, especially in terms of ensuring a satisfactory response rate

Another important limitation is the avoidance of the collection and storage of participants' personal data, which is in accordance with applicable laws. The need to have contact with the same groups of people in the panel research requires storing their e-mail addresses. In order to maintain confidentiality of the research, individual tokens were used to identify participants, which are linked to e-mail addresses stored in a separate database. This ensures that the respondents remain fully anonymous during the analysis of their answers.

It should also be noted that data obtained are of a declarative nature and take into account the point of view of the analysed mobility participants only. This is due to the assumed objectives of the research, which consist primarily in collecting opinions on the implemented interventions.

Conclusion

Ex-post evaluations allow to assess the effects of intervention in the long term. The studies carried out so far have allowed for monitoring the lives and careers of former VET and HEI students with the experience of transnational learning mobility. Thanks to the research, it was not only possible to learn about the effects of the programmes that were subject to evaluation, but also to identify the needs of the participants and their expectations related to further learning and entering the labour market. Such knowledge can not only be used to evaluate the quality of the programmes but can also be of help in better adapting their offer to the needs of participants. However, it should be remembered that, as in the case of any other research method, it is impossible to avoid certain limitations related, for example, to the declarative character of the answers provided, the difficulty in reaching former participants of the programme, or the "panel wearing thin" phenomenon. It should also be kept in mind that monitoring graduates, regardless of their level of education, is only effective if it is carried out continuously and allows for the comparison of collected data in the long-term perspective.

Bibliography

Babbie, E. (2013). *Podstawy badań społecznych*. Warsaw: Wydawnictwo Naukowe PWN.

Dąbrowska-Resiak, J. (2019). *Erasmus... and what next? Tracer study on education and careers of Polish students participating in Erasmus mobility projects*. Warsaw: Foundation for the Development of the Education System.

Erasmus+ Programme Guide. Valid as of 1 January 2014 (2014), version 3: 9 April 2014. Brussels: European Commission.

Koniewski, M. & Lisek, K. (2017). *Badanie absolwentów programu Erasmus+. Wskazania do realizacji pilotażu ilościowego oraz badania panelowego*. Warsaw: Foundation for the Development of the Education System (in-house document).

Menard, W. (1991). *Longitudinal Research*. Newbury Park, CA: Sage.

Pachocki, M. (2018). *Is mobility the key to a successful career? Tracer study report on vocational learners with transnational mobility experience*. Warsaw: Foundation for the Development of the Education System.

Desing Thinking and evaluation – differences, similarities, and examples of their application in education

DOI: 10.47050/66515246.118-129

Agnieszka Skowrońska

Abstract

Evaluation is an area of applied social research closely related to projects or, more broadly, to support activities. In recent years, a methodology called Design Thinking (DT), which in the author's opinion shares several features with evaluation, has also been developed. In a broader sense, the ultimate goal of evaluation is to lead to the design of activities aimed at solving problems in a specific area. Design Thinking also focusses on the design of products and services (Service Design) that meet people's needs or solve their problems. This chapter identifies common features of the two approaches, as well as differences between them. It presents opportunities for using methods applied in the Design Thinking methodology in evaluation, and features practical examples in the area of education. Such knowledge can contribute to the development of evaluators' work.

Assumptions of Design Thinking[1]

The assumptions of Design Thinking (DT) focus on people, so the methodologies that emerge from them are called human-centred design (HCD). Everyone has needs they want to satisfy or problems they need to solve. By using services or products, we satisfy those needs or solve those problems, and we should have positive experiences related to this. Unfortunately, many products and services generate negative experiences, because they do not meet our expectations. Let us take training as an example. We participate in a training course with the intention of gaining specific knowledge or skills, because we need to develop in that particular area. If it turns out that, contrary to what was announced, the level of presented knowledge is lower than expected, and we do not

1 Design Thinking is a generic name of the methodology for carrying out projects. It covers various methods focussing on a specific element of a process or on a given area, e.g. User Interface Design, User Experience Design (UX), Service Design, SCRUM, and Agile.

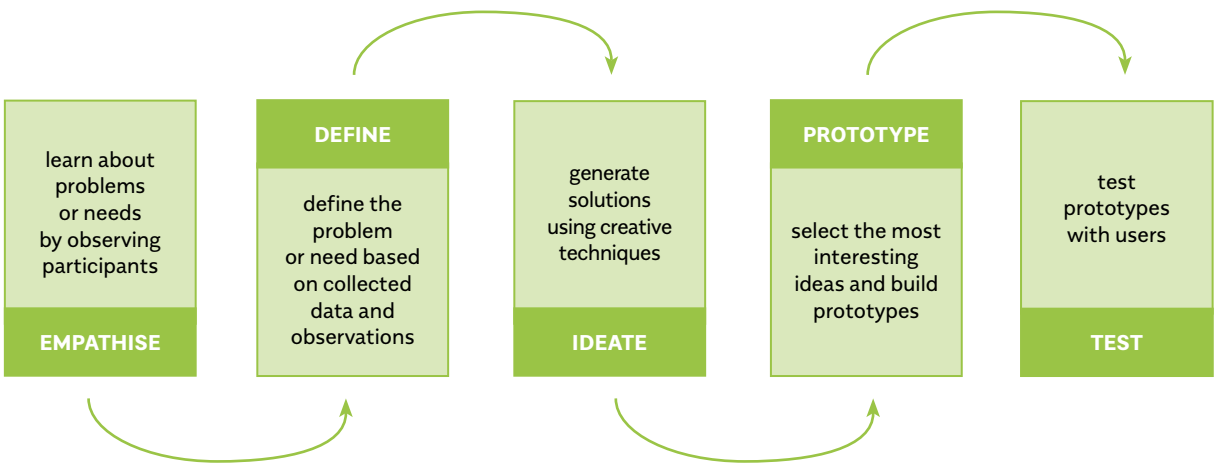
Agnieszka Skowrońska
Sociologist with a PhD in humanities, she runs Pracownia Innowacji, which analyses needs, designs services and social innovations, as well as offers training in user needs research and innovation design. She has 15 years of professional experience in the field of social policy, conducting research and carrying out evaluations. She is the author of several publications, among others: *Badania w pomocy społecznej* [Research in social services], *Pomoc społeczna w liczbach* [Social services in numbers] and *Potrzeby to podstawa* [Needs are the basis].

KEYWORDS:
EVALUATION,
DESIGN THINKING,
HUMAN-CENTRED DESIGN,
PROJECT PROCESS,
NEEDS

learn anything new, we feel that we wasted time and money. This means that the service was not designed in accordance with the assumptions of DT, or at least at some stage (i.e. communication about the service), a mistake was made, because in effect the training was attended by a person who had broader knowledge than that which was presented. At the evaluation stage a training participant will certainly give a low score for relevance of training.

A thorough examination of the needs of people for whom a product or service is designed forms the basis of DT. Thanks to this approach, people who are to use the product or service to be developed are from the very beginning involved in the design process. Ideation and prototyping services or products are the subsequent stages of the process, in which their future users also participate. What is important in this approach is to constantly test the solutions, i.e. try using prototype products and services and improve them according to users' expectations. In this way, services or products which are required by users, which solve their important problems or which meet their needs are created. The basic elements of the described process are shown in the figure below.

Figure 1. Stages of Design Thinking



Source: author's own work based on *The Field Guide to Human-Centered Design*.

Apart from focussing on people, (what means designing solutions primarily for them), DT also involves project work in multidisciplinary teams. This is to provide an outlook on a particular problem from a different perspective. For example, not only methodologists, andragogues and education professionals should take part in the process of developing new forms of adult education, but also their future participants, trainers and communication specialists, whose task will be to communicate new possibilities for improving skills. It is also worth noting that, according to DT, a designed service (e.g. adult training cycle) comes into being much earlier than the first training session. It starts at the time when the need to improve a future trainee's specific skills arises. And the process ends when a training participant receives a certificate attesting the completion of training or skills acquired as a result of it, and, if necessary, after they share opinions on the training received.

In the case of a product, things get even more complicated, because we can design the process of manufacturing the product itself, then its delivery to the user, various stages of its use, and finally what happens with it at the end of its life cycle. Another element distinguishing DT processes from other design methods is prototyping instead of discussing possible solutions. The essence of prototyping is to create fast and cost-effective yet tangible visualisations of the solution, which can be shown to future users, and thus to obtain feedback on them. This process is supported by the ability to present solutions to users, which can be summarised in the slogan: *Show it to me instead of talking about it!* Using DT we can design single products, e.g. books or educational materials, more comprehensive services, e.g. e-learning courses or innovative study programmes, as well as entire systems. An example of a system designed using DT is the Innova School network in Peru[2]. In reaction to the poor quality of education in that country, a network of private Innova schools and a company called IDEO designed an ideal school and a new model of education based on research and prototyping. The school includes a campus, which can be easily adapted to different purposes. It comprises laboratories, amphitheatres, cafés, and developed roofs. The interiors can also be easily rearranged, by moving sliding walls and furniture on casters. Learning at the school is a combination of teacher-directed individual and group education, project work, experimentation and the use of the latest communication technologies. Also, a Teacher Resource Centre was created, where teachers can find 18,000 lesson scenarios developed in accordance with the latest trends in pedagogy. Thanks to the introduction of these solutions, Innova Schools, originally comprising several schools, has expanded into the largest private school network in Peru and its students score up to three times higher than the national average in exams.

Design Thinking and evaluation – differences and similarities

In the Design Thinking process, the needs of stakeholders who will use the service or product are examined. This is an equivalent of an *ex ante* evaluation, which also analyses the needs of future users and appraises, among others, whether the project assumptions are in line with those needs. Several iterations in the DT process are *de facto* evaluations of different versions of solutions at different stages of their development. The whole DT process resembles an ongoing evaluation, assuming that there is a possibility of making changes on an ongoing basis (which does not happen in many projects). The goal of both processes is to solve problems more effectively. If it is assumed that the ultimate goal of evaluation is not to write and present a report, but to implement the recommendations developed (preferably with the participation of many stakeholders) on its basis, then ultimately it should always lead to the implementation of solutions effectively addressing the problem.

Although both evaluation and DT are processes, the basic difference between them is that the former is a research process, and the latter is a design process. The former usually ends with a report featuring recommendations, and the latter with an advanced project or launch of a product or service. In DT, the emphasis is placed on qualitative data collection methods, creative exercises and techniques, while evaluation still focusses on quantitative methods, although qualitative methods are also used. There are also different reasons why an evaluation or design process is carried out in accordance with DT principles.

2 See: www.ideo.com/work/designing-a-school-system [access: 7 February 2020].

Design Thinking is used in order to:
1. raise the quality of services offered,
2. personalise services in line with user/customer needs,
3. introduce innovative solutions,
4. invent new services and products,
5. solve problems.

Evaluation is used to:
1. evaluate structured projects/programmes,
2. improve the functioning of an organisation/institution,
3. evaluate the quality of an initiative.

Evaluation and DT processes are not the same but have some features in common. Each new solution is based on the evaluation of previous experiences, so DT is inseparable from evaluation. We can say that the DT process is intertwined with traditional evaluation, while participatory evaluation uses many elements of the DT process. With a modern concept of evaluation – especially when using participatory processes – we can assume that Design Thinking is an extended and advanced participatory ongoing evaluation that leads to the implementation of new and effective solutions in response to complex problems. The participatory process should be based on involving stakeholders, with particular emphasis on the users of the solution (service or product) at each stage of design or evaluation.

Design Thinking methods which can be used during evaluation

Design stages

The Design Thinking process can be useful in designing an evaluation, and also during its implementation. It provides an opportunity to develop new methods, to think outside the box and, above all, to foster greater stakeholder involvement in the project under evaluation. The first stage of the DT process is to discover needs, which can be compared to identifying the needs of parties commissioning evaluation or, more broadly, to the needs of other stakeholders in a given intervention. It consists in careful observation of behaviours and listening what they say. We can infer about people's needs not necessarily from what they say about them, but from what they say when we ask them more general questions. For example, if we want to design an innovative way of encouraging young people to take up physical activity, we should not only talk to physical education teachers, but above all look at the kinds of activity young people take up, and analyse under what conditions they are most willing to move, what motivates them to do it, and what they complain about.

At the second stage, we select the collected information, analyse it and determine which challenge we want to address. In evaluation, this is the time to formulate evaluation questions and to verify previously adopted assumptions and objectives. After a preliminary identification of the problem and the needs of various parties, it may be necessary to extend the scope of evaluation (or the design process). For example, if young people often complain about the lack of a suitable place to wash and change after a physical education class, we have to precisely define this need, and not just deal with the way the classes are taught.

Creative idea generation is the time to write down as many ideas as possible that come to our mind. It is important not to assess their feasibility at this stage. When designing the evaluation process, it is possible to invent new research tools or apply traditional tools in a new context. During evaluation at this stage, ideas can be gathered on how to solve a given problem (recommendation). With reference to the previous example, a workshop could be organised to design a place where young people could wash and change after physical education classes.

The final stage is to create a prototype and test it with prospective users. An example of a prototype is a model of a room or a tool. When designing evaluation, these can be prototypes of research tools that can be tested at the pilot stage. In the course of evaluation, it is possible to "test" developed recommendations in this way. Prospective users can verify, for example, the ease of use of the research tool (understanding of the instructions, willingness to participate in an interview under certain conditions) or the usefulness of the proposed solution (e.g. what would attract them to a place designed in this way, and what would make it difficult for them to use it). It is probable that the solution would have to be improved or supplemented with other functionalities. This process must be repeated until an optimal solution is achieved.

Methods of discovering needs

Discovering needs (empathising) consists in "putting ourselves in users' shoes" and a profound understanding of their problems. It cannot be based solely on statistics or a cursory review of documents and existing data related to the subject matter of the research object. *Ex ante* evaluation is the stage at which it is worth discovering the needs of different stakeholders of the designed solution, but this can also be done at other evaluation stages. Below a few methods that can be used with little financial expenditure are presented.

Service Safari

Service Safari is participatory observation, the essence of which is to experience a given service. If we evaluate solutions that already exist on the market, we should first use them ourselves. If this is, for example, an online learning service for adults, it is worth taking part in such a course. Apart from observing, it is necessary to write down our insights, and photograph or record what has been observed. Experiencing the service during such an observation allows us to understand exactly what the service is about, and what its strengths and weaknesses are. Usually, when experiencing the service, a lot of ideas for improvements come up, along with research questions that can be explored at further stages of evaluation.

User Shadowing

This is a form of a non-participatory (but usually overt) observation, and it consists in "following" the user and observing everything that happens around them. It is useful especially when we cannot use the service ourselves, e.g. we want to observe how lessons are taught at a primary school, and we cannot pretend to be a student or teacher. The purpose of such observation is to understand certain behaviours, e.g. what children do after school, what teachers do during breaks. Thanks to this method, the observer can see things people usually do not talk about during interviews or those that cannot be included in questionnaires. The observer often uses a video camera, takes photos or notes.

Contextual, Photo-Interview

During evaluation, an in-depth individual interview or a group interview is a frequently used method. Another type of interview is a contextual interview, which is conducted in the place and at the time related to the subject of the conversation. Such an interview is often conducted in users' homes while they perform daily activities, e.g. help their child with homework. The interviewer asks questions about particular activities and reasons for applying a specific method of performing them. They can also observe what emotions accompany these activities, and what makes them easier or more difficult. This way, we can find out what emotions are involved in helping the child to learn, what parents lack in providing this help, etc.

The method that is especially recommended for interviews with children and young people is photo interview. During the interview, we can use photos taken earlier (e.g. during school events) or presenting a given topic, e.g. "show me what you do after school". The respondent tells the interviewer when and why the photo was taken, thus telling a story related to it.

User Diary

This is a form of self-observation and recording our insights, e.g. in a diary or notebook, or with the use of a video or photos. Its advantage is that people can do it when and where it is most suitable for them. The observation should last at least a few days, so it is particularly useful if we want to observe habitual behaviour. For example, if we want to evaluate periodic training or classes offered by third age universities, we can suggest that people taking part in such classes record when and where they participate in them and how they choose them. They can also attach photos of the places where the classes are held, the aids that are used, etc., as well as briefly describe what they have learnt during the classes.

The analysis of collected data makes it possible to learn about behaviour patterns and the motivation behind them. The aim of the needs test is to expand knowledge about the users themselves or about the service, and not to focus on collecting opinions about the service. Therefore, all qualitative methods are recommended.

Methods of analysing information and defining the problem

A lot of information is available after a thorough examination of the needs of prospective or current users of a service. Yet organising and defining the need that we want to meet with our products or services is just as important as data collection. This is the second stage of the Design Thinking process. Organising information also takes place at different stages of evaluation (after the initial identification of the evaluation area or after the collection of relevant data). Stakeholder and context maps are often used at this stage. These are also known in evaluation, so they will not be described here. However, tools not used in evaluation, i.e. customer journey map and personas, will be presented. Finally, a simple method to formulate a problem will be shown.

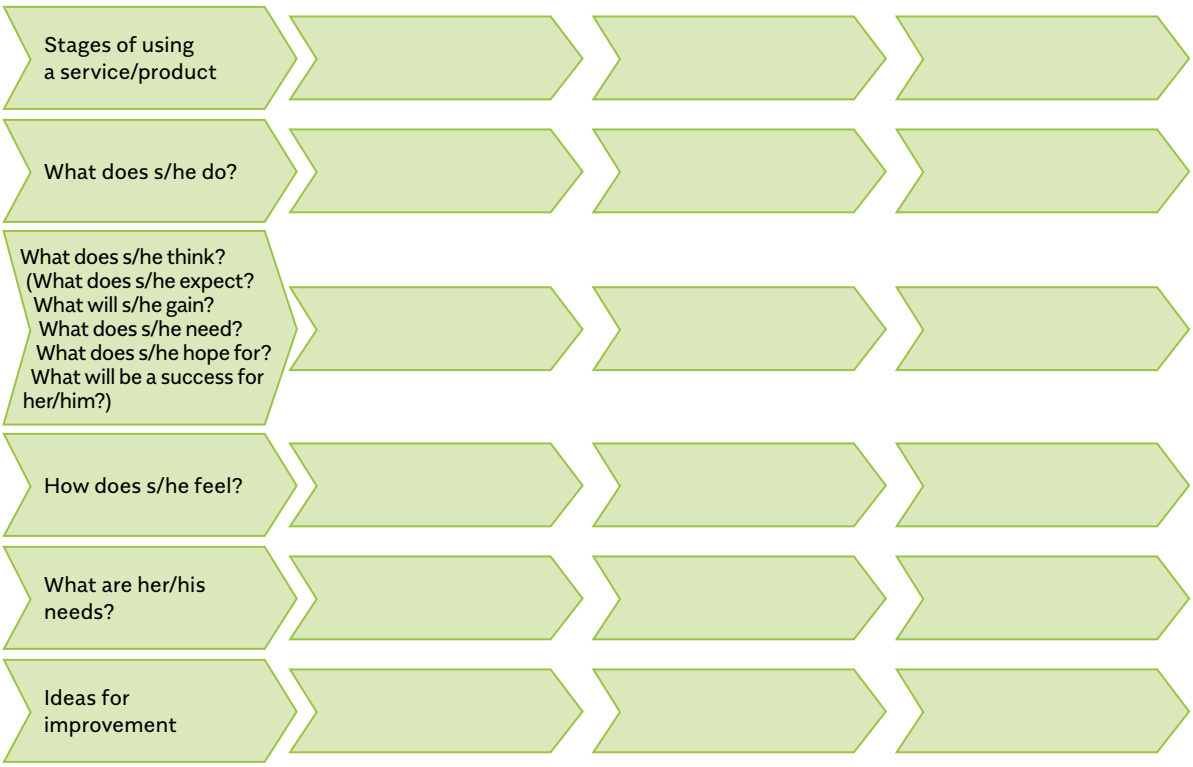
Customer journey map

A customer journey map illustrates the path a person takes when they want to satisfy a specific need or solve a problem. It consists in drawing subsequent steps that a user needs to take to use the service. This journey begins when the person becomes aware of the given need, e.g. "I need to speak English better to communicate with my grandchildren (born and living in the UK)", and ends when they share their opinion about the English course taken or continue their learning at a higher level.

During this journey they have to make various decisions (e.g. which form of learning to choose, whether to take an intensive course or study over a longer period of time, whether to learn individually or in a group, etc.). They have to choose a specific course and enrol in it, pay for it, get to the place where the classes are held or use a computer if it is an online course. Last but not least, they have to participate in classes, memorise vocabulary at home, and use it in real life situations.

Such a journey can be drawn in many different ways (e.g. in the form of tables, diagrams, mind mapping). Sometimes drawing the user's path (writing down subsequent user/client actions) itself leads to awareness that an element is lacking in a given service (e.g. the possibility of enrolling in a course online). The maps representing each stage of the service should include the results of research (including quotes of users' statements), the most important observation insights, information obtained during interviews, etc. It is worth organising this information into different categories – the main are actions, thoughts, and feelings. It is also worth allocating some space on the map to needs and ways of satisfying them. The map can feature photographs and drawings. An example of a "map" diagram is presented below.

Figure 2. Customer journey map diagram



Source: author's own work.

The most important thing when drawing such a map, just like when analysing the whole material, is to identify significant elements concerning needs. This should lead to defining them.

Personas

Another way to organise information obtained during user research is to create a persona, i.e. a representation of a present or future user of a service or product. The aim of creating a persona is to determine their needs and motivations. In evaluation, a persona can be created in relation to those who are to participate in research to propose the best method and also as a summary of the results, e.g. to present different types of e-learning course participants. When creating a persona, we give them a name and surname (always fictional), determine their age, professional activity (e.g. student, pensioner or employee in a particular industry), and place of residence. We often describe their interests, family status, favourite catchphrase (motto) and channels from which they obtain information (e.g. television, supermarket leaflets). It is important to describe both the general needs and objectives of the persona, and those relating to the area that is analysed. We need to determine what they like and dislike. We can also draw such a persona. On the Internet, we can find many templates for describing a persona. One of the simplest ones is presented below.

Figure 3. User persona template

Source: *Świat według FUZERS...*, 2018.

A persona or personas are meant to remind us that solutions are created for people with specific characteristics and not for institutions, companies or the general public. During teamwork on creating a persona, a common vision of the person for whom the solution is designed is constructed. The aim of the data analysis process, which is visualised through user journey maps or personas, is to narrow down the problem we want to focus on (at least initially) or to indicate areas of activity aimed at addressing the problem in a comprehensive manner. For example, a general idea of designing modern forms of continuing education for adults can, after examining their needs, be narrowed down to the development of education methods which do not require commuting. At this stage, the needs and a solution are formulated. The following formula may be helpful here: How to help (whom?)... to solve a problem/satisfy a need (related to what?)..., so that (what is the desired effect?)...

Example 1: How to help adults (e.g. 30–45 years old) to meet their need for development so that their qualifications are up to date on the labour market.

Example 2: How to help adults (e.g. 30–45 years old) to meet the need to develop their communication skills in a way that does not require them to commute to a "school".

Methods for generating solutions

Having precisely defined a need or a problem, we can proceed to working on solutions. The DT process is usually about innovative solutions, but this isn't innovation for the sake of innovation. The essence of this phase is to open up to new ways of solving a problem which has not yet been solved. We do not have to try to develop new research techniques if the ones we have been using are effective. But if previous evaluations show that the questionnaire response rate is low, it may be worth considering other methods or tools. At this stage, different techniques are used to stimulate creativity, which are to lead to pattern breaking. Two of them are presented below.

Individual brainstorming

Brainstorming is a well-known and often used technique in teamwork. However, individual brainstorming gives everyone (even when working in a group) an opportunity to write down their ideas. Then we can pass them on to another person in the group, and maybe our colleague will come up with a new idea based on the one we have produced. This working method makes it possible to generate several ideas in a short time, and to avoid schematic solutions that go in the same direction.

Pattern breaking – SCAMPER

SCAMPER is a method of pattern breaking by means of questions, and its name is an acronym of the first letters of words identifying those questions: S – substitute, C – combine, A – adapt, M – modify, P – put to other uses, E – eliminate, and R – rearrange. With answers to the questions (listed below), we can define new features of the idea we want to develop. It can also be applied to products or services that already exist. Some of the questions may seem absurd, but it is their absurdity that contributes to breaking patterns.

Figure 4. SCAMPER questions

Substitute	→ Can we substitute any of the elements of our idea with something else? → Can we change the rules? → Can we replace components or materials with others? → What would happen if we changed the position of individual parts/elements/activities? → Can we use this product somewhere else or replace something else with it? → What would happen if we changed our attitude to this product or its image?
Combine	→ What would happen if we combined our idea with another one? → Can we combine or recombine the parts that we have? → What materials can we combine? → Can we add other items? → How can talent and resources be combined to create a new image of this product?
Adapt	→ How can we adapt the idea so that it is suitable for other purposes? → Who or what can be replaced by this idea? → How can existing solutions be adapted to the predefined problem? → What can we copy, borrow, steal for this purpose? → In what contexts can we use our product? → Does the problem exist in a different context?
Modify	→ How could we change the shape, look or feel of our idea? → What can we add to it? → What could be taller, bigger or stronger? Lower, smaller or weaker? → What could we emphasise or highlight to create more value? → What element of this product could we strengthen to create something new?
Put to other uses	→ Can we use the idea somewhere else or which element can we use in a new way to solve the problem? → What are other applications for it? → Can it be used by people for whom it was not intended? → How would a child use it?
Eliminate	→ How could we streamline or simplify this product? → What is unnecessary? → Which rules/steps/elements can be skipped? → Should it be divided?
Rearrange or reverse	→ What would happen if we reversed it or sequenced things differently? → What other setting would be better? → Can we do something in reverse/another order? → What idea is the opposite of ours?

Source: author's own work[3].

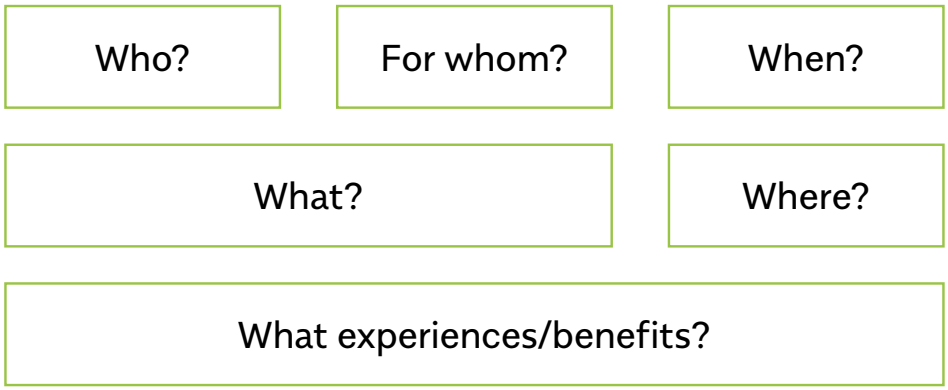
Methods for idea prototyping

Once we have generated our ideas, we should pick a few of them for testing. Selection methods are usually based on a variety of criteria, including whether it could be implemented over a short period of time, whether it would be expensive, and whether it would comply with applicable law and ethical standards. It is also important that future recipients of a given service participate in the selection of the idea to be implemented. In evaluation, prototyping can be applied to both research tools and

3 Based on *SCAMPER. Improving Products and Services*, www.mindtools.com/pages/article/newCT_02.htm [accessed: 10 February 2020].

solutions described in recommendations. Every idea can be described by means of a simple diagram, as presented below. At further stages, it can be developed using various prototyping techniques.

Figure 5. Basic service description



Source: author's own work.

Storyboard

The first, quite simple, way of testing our idea is talking about it to people who are going to use it. One of the methods of presenting subsequent stages of a service is using illustrations, which form a storyboard. They can also be presented as a comic book. These images can be very simple and range from schematic drawings to illustrations made using computer programmes. The most important thing is that our audience finds them clear. A storyboard is intended to help us present the milestones of the service and to collect opinions on how much the presented idea appeals to people, whether (and to what extent) they understand what the service is about, and whether it responds to their needs. The result of the conversation is a confirmation of the validity of the idea or, if necessary, its supplementation with further elements that were omitted in the original idea.

Models

Another way is to create a model of the solution, especially if it concerns space. Different objects can be used to create a model – cardboard, bricks, plasticine, toys – depending on what we have at our disposal. Using these objects, it is possible to arrange a small-scale space, in which the service is to be provided (e.g. arrangement of benches in a classroom) or to present subsequent stages of the service. The work on the model itself is an opportunity to discuss the details of the solution and look for common methods.

Mock-ups

When designing online services, it is worth creating mock-ups, i.e. paper versions of websites or applications. In the simplest versions, the pages are drawn on a piece of paper or pasted together using post-it notes, and in the more advanced versions, they are created using computer programmes.

The idea testing phase is to show whether prospective customers will use the solution, so they need to participate in the testing. It is a process of collecting feedback, as well as correcting and improving prototypes until the version best suited to the needs of the recipients is created.

What elements of Design Thinking are worth using in evaluation: summary

Design Thinking serves as an opportunity to develop competences used in evaluation. The author is not saying that from now on every evaluation must be carried out like a design process, but there are several elements that can and should be "transplanted" from DT to evaluation. The first one is to make evaluation processes more focussed on end users than on institutions. This means including users in evaluation design and also creating tools tailored to the capabilities of a particular group – schematic surveys do not perform this task. It also means that recommendations should be worked out not only with representatives of those who commission the evaluation, but above all with those who will use the solutions (e.g. students, parents). Bringing to light users' hidden needs is another element that can be developed in evaluation. Discovering needs that users do not talk about directly will allow us to design solutions that will solve their real problems and improve their experience in using the services. Design Thinking also provides an opportunity to complement the practice of evaluation with iterations, i.e. "quick and short" evaluations of subsequent prototypes of solutions, as well as to use prototyping to develop recommendations.

Bibliography

Grocholiński, P. & Michalska-Dominiak, B. (2019). *Poradnik design thinking, czyli jak wykorzystać myślenie projektowe w biznesie*. Gliwice: Onepress.

Rudkin Ingle, B. (2015). *Design Thinking dla przedsiębiorców i małych firm. Potęga myślenia projektowego w codziennej pracy*. Gliwice: Wydawnictwo Helion.

Świat według FUZERS, czyli słownik service design. fuzers.com/pl/swiat-wedlug-fuzers-czyli-slownik-service-design-cz-2/ [accessed: 10 December 2019].

The Field Guide to Human-Centered Design. IDEO. www.designkit.org/resources/1 [accessed: 10 December 2019].

Creative evaluators or how to make evaluation more appealing[1]

DOI: 10.47050/66515246.130-142

Beata Ciężka

Abstract

In the evaluation of developmental activities, such as training, coaching, counselling and guidance, addressed to adults, questionnaire surveys, which often bring superfluous and useless information, predominate. This chapter presents evaluation concepts and techniques focussing on the development-oriented role of evaluation. They aim at a more creative approach based on deeper reflection, dialogue and learning outcome ownership. The purpose of this chapter is to present methods of attracting participants to the evaluation process and to allow them to collect relevant and useful information.

Why is it worth looking for new evaluation techniques in the case of development and training projects?

Evaluation research methodology draws on the high methodological achievements of such fields as sociology, pedagogy, psychology, economics and management sciences. It applies and adapts approaches and research methods used in them for the purposes of evaluation research and analysis.

Despite many years of evaluation practice in Poland, the question of methodology – how to collect information from participants to make it as useful as possible in order to improve the development service and actually give valuable feedback to its providers – is still relevant. In the evaluation of development projects, especially training projects addressed to adults, questionnaire surveys predominate. But is this method always (or at all) the best tool for this? The majority of people involved in adult development projects have experience in conducting surveys, especially these carried out *ex post*. However, everyone knows how hard it is to encourage trainees to complete the questionnaires. Inasmuch as closed questions, in which participants mark a selected option, are answered more often (although the questionnaires can still be returned not fully completed), open

1 This article is a compilation and extension of the author's materials published on the Electronic Platform for Adult Learning in Europe (EPALE) at epale.ec.europa.eu/pl.

Beata Ciężka
Independent evaluator since 1995. She has extensive experience in the implementation of research and evaluation projects conducted, among others, for the European Commission, the European Parliament, public administration bodies, non-governmental, as well as private sector organisations. Author of training programmes in evaluation, many of which she conducted herself. She was also a lecturer, teaching postgraduate courses in evaluation at the Institute of Sociology of the University of Warsaw and at the Warsaw School of Economics. Founding member of the Polish Evaluation Society, of which she was President in the years 2004–2010 and is currently a Member of the Management Board.

KEYWORDS
EVALUATION,
EVALUATION CONCEPTS,
EVALUATION TECHNIQUES,
PARTICIPATION

questions, where participants need to formulate an answer on their own, usually remain unanswered or the answers provided are composed of one, maybe two words, and have little informative value. The reluctance to fill in questionnaires is caused by their widespread use – we are asked to complete them time and time again, be it in a shop, on the street, or at a service outlet. What is more, in the case of trainings, they are usually conducted at the very end of the session, on the run, when the majority of participants are about to leave. How, then, can we provide favourable conditions for reflection, which is so important for evaluators and their clients?

Of course, questionnaires have their advantages – above all, they can cover a large group of people, are relatively easy and quick to conduct, and their results are countable. This way we obtain information about the frequency/percentage distribution of participants' traits and opinions, and we know which concern the majority and which are marginal (the latter can be very important from the perspective of improving the evaluated service). Participants have freedom in responding to the questionnaires, their anonymity is also ensured. These advantages make us systematically use this tool. However, more and more often the obtained results disappoint us. Frequently, the collected information is perfunctory and of poor quality, which mainly results from the participants' reluctance to provide it. Consequently, it does not form a basis on which we can found the quality of our work.

What can we do in such a situation? This chapter proposes a more creative approach involving greater reflection, dialogue and learning outcome ownership. It assumes that information is collected from participants using active methods, in a way that engages them to such an extent that fatigue after training and reluctance to participate in evaluation are overcome.

Concepts used in the evaluation of development and training projects

Evaluation strategies supporting the quality of development services increasingly emphasise the socialisation functions of evaluation and a more open approach to its implementation. As an example of such approaches, it is worth mentioning participatory evaluation and developmental evaluation, as well as the concept of action evaluation (Ciężka, 2018a). Michael Quinn Patton (1997, 2011) is the creator and advocate of the two former approaches, whereas action evaluation is based on the assumptions of action research (Červinková & Gołębniak, 2010).

Participatory evaluation (Patton, 1997) assumes that all participants in an evaluated project are involved in the whole process, at each of its stages – from the construction of the evaluation concept through the collection of results to their interpretation. The evaluation process provides for real participation, a practical acquisition of evaluation skills by the participants and joint work and collective research supported by the evaluator. People involved in the project being evaluated choose the principles they want to adapt in the process of evaluation research. They define categories they believe to be of key importance, and the evaluator supports them with methodological knowledge. The focus of participatory evaluation is to search for opportunities to use evaluation data to find practical solutions to problems within the current organisational context, which can enhance the results of educational activities. The evaluator becomes a facilitator and collaborator – they have knowledge and research skills, but their role is not crucial or dominant. This is because work is carried out in a team, with the assumption that the evaluator fosters team cohesion and ensures the methodological correctness of the study. At the same time, they explain its principles so that all aspects of evaluation are understood and justified by the participants. The difference in status between the evaluator

and the participants is minimised. It is the participants who make decisions and act as evaluators, focussing on the process and the results they consider important. The purpose of such evaluation is to strengthen the responsibility of participants for themselves and the community in which they work. This is why this model is useful in educational projects involving a change in the behaviour of participants and in the work culture of their organisations.

A more advanced strategy, which takes into account the indications of the participatory model, is the concept of **developmental evaluation**, formulated by the same author (Patton, 2011). It assumes long-term, partner relations between persons conducting the evaluation and designing an organisation's development programme. In accordance with this approach, evaluation consists in asking research questions and using evaluation logic as a tool supporting the process of change. An evaluator is a member of the team working together to develop, design and test new approaches for the process of long-term, continuous improvement, adaptation and purposeful introduction of change. The primary role of the evaluator is to stimulate team discussions, provide explanations, and present evaluation questions, methods, data analysis and logic. In this way, the evaluator supports decision-making based on reliable sources and, consequently, the development process. It can therefore be said that the evaluation process is also the result of the evaluation. In this approach, change is not seen as progress, but as continuous adaptation to what is happening inside and outside the institution. Developmental evaluation often includes elements of participatory evaluation, especially when staff and participants are encouraged to formulate personal objectives and monitor progress in achieving them. Similarly, all members of the team interpret the conclusions of the evaluation, analyse the implications and apply the results at the subsequent stage of the organisation's development.

This trend also includes the popular concept of **action evaluation**, which refers to the sought-after action research methodology (Červinková & Gołębiak, 2010). Here, the purpose of collecting information is not only to describe a given situation or problem, but above all to bring about a specific change. The main value of this approach is the participatory character of the whole process, which means analysing actions and their effects with the involvement of a given community and its orientation on activity throughout the process. The very term "action research" and the assumptions of this concept were formulated as early as the 1940s. It was Kurt Lewin who defined this new approach to social research. It assumed the creation of social theories combined with the production of social change thanks to actions taken by the researcher.

Action research can thus be defined as a s*ocial situation study aimed at improving the quality of action within it* (Elliott, 1997). Its goal is to provide practical information in specific situations and, consequently, to ensure smarter and more efficient action (ibidem). The key elements of action research are cyclicality, the emergence of successive research questions, co-creation of further steps, using the research to bring about change, and continually introducing change. The researched community, the initial idea and evaluation itself are here constantly evolving. In this approach, evaluation does not constitute an independent process, but forms an integral part of project management, in particular change management, and focusses on analysing the effects of such changes and on supporting the process by identifying whether or not a specific action is producing the desired results (Ciężka, 2018a).

What techniques should be used to interest participants in evaluation?

Evaluation not only plays a research role, but can also perform important social functions. Above all, it significantly strengthens the sense of ownership of project outcomes. This is particularly important in development and educational projects. Evaluation can also impact forming teams by participants and staff carrying out and managing the project. However, in order for such functions to become apparent, the evaluation process itself, including the techniques used in it, should be dialogical and should provide space for building relationships between project participants, reduce the fear of evaluation and support the feeling that evaluation forms an important part of the decision-making process. Those taking part in evaluation should be willing to participate in it, and should not only feel that they are involved in an important activity, but also be curious about it, want to be engaged in it, and be motivated to be active. The techniques presented below promote such engagement. They stimulate participants' interest in the evaluation process and allow them to collect relevant information, deepen their reflection, and formulate useful and revealing recommendations. Evaluations carried out using these techniques are an alternative to common questionnaires conducted after a training or project often in a non-reflective and automatic way.

Poster sessions organised in the Talking Wall formula

Use

Talking Wall is a method which works well in both small and large groups. On poster-format sheets of paper (e.g. a flipchart) attached to a wall (but also placed on tables or even on the floor), participants can write their statements or stick post-it notes with them. Posters have titles that indicate topics on which participants can formulate their opinions. Such headings can cover all questions that are usually asked in evaluation questionnaires. They can refer to substantive, methodical or organisational aspects of the project or training.

Examples of poster headings (tried and tested during many training sessions):

→ During the training, I learned that...
→ The most interesting thing for me was...
→ The most useful thing for me was...
→ The least interesting thing for me was...
→ The least useful thing for me was...
→ What I'll take from the training is the idea that...
→ During the training, I was able to make contact with...
→ The training was organised...
→ Next time, the organisers should...
→ I'd like to tell the coaches that...
→ I see a need for continuing development in...

There shouldn't be too many posters – 4 or 5 topics will provide organisers and coaches with the most useful information. The topics can relate to project or training assessment, outline the perspective of the use of knowledge and skills or identify further training needs. This format forces the participants to leave their tables, changes their perspective, allows them to summarise their experience and confronts them with the opinions of others. This is important as it strengthens the effects of a project or training. Participants in such sessions often say that getting to know the statements and opinions of others allows them to realise how they can use knowledge and skills gained during the training. What is important, this form of evaluation is interesting for the participants and activates them to an extent which enables us to obtain a lot of information they would never enter in a questionnaire.

The advantage of the Talking Wall method is that it can also be used at mass events, e.g. conferences, even those attended by several hundreds of people from different countries. We can hang posters in the lobby where participants have coffee during breaks. At that time, they can write their opinions and comments. Very often, we can see that participants stop by individual posters and have lively discussions. The abovementioned feedback effect appears here.

The advantage of a questionnaire is that it allows us to diagnose the occurrence of a given opinion (i.e. summarise the answers to a given question – we can calculate what percentage of participants have given a specific answer). The Talking Wall is a qualitative method. This means that we can get to know participants' views, but we will not know whether a given opinion pertains to the majority or to only a few individuals. In the evaluation of training or other forms of development, we often do not need any information about the frequency of the occurrence of a given perspective. Such evaluations primarily focus on people's (individual) experience and on how we can improve our offer. Usually, answers that inspire change can be heard from one or two people, but certainly not from the majority of participants. Nevertheless, sometimes we need to know the scale or prevalence of some opinions in the group. In such a case, we can ask the participants to mark on the poster (e.g. with a dot) viewpoints which have already been written down and with which they agree. This will allow us to know how many people express a similar opinion and the quality of such data will be no poorer than that obtained from a questionnaire.

Another frequently cited advantage of questionnaires is respondents' anonymity. It is based on the assumption that thanks to the lack of the author's identification, the respondents will be more inclined to express critical opinions and will feel more secure. However, experience shows that questionnaires do not encourage unfavourable statements. The open nature of writing opinions on posters does not prevent criticism – on the contrary, it even opens up space to indicate what can be improved in a project or training and how this can be changed so that it better responds to the needs of participants.

Preparation and implementation

At the beginning of a training/conference, the trainer/moderator informs the participants about evaluation and its purpose, which is to obtain their feedback. They stress the usefulness and anonymity of opinions to be given and argues that the viewpoints (including critical ones) make it possible to assess the quality and usefulness of training/conference in a reliable manner. They tell the audience that evaluation will take place using the poster format, i.e. unfinished sentences displayed on large sheets of paper.

During the first break, posters featuring headings as indicated above are hung in the training/conference room or in the hall/lobby (depending on the capacity of the premises). During the breaks and at the end of the training/conference, participants are asked to walk around and write down their

opinions by finishing the sentences in the headings. The trainer/moderator accompanies the participants during breaks and after the training/conference, asks questions to obtain detailed information and explanations concerning the views written on the poster and keeps a record of them. They encourage sincere and critical opinions. It is important to note which sessions are followed by specific attitudes, and which statements appear at the end of the training/conference.

As far as possible, at the end of the training/conference, the trainer/moderator summarises the information obtained during the poster sessions by discussing individual poster headings. It is worth to initiate a short summary discussion, aiming at making recommendations to the trainer or organisers, but also to the participants on how they can make full use of having taken part in the training/conference to make it as valuable and useful as possible (this is about creating ownership of achieved outcomes).

Hexagon method

Use

The hexagon method[2] can be used to collect information and strengthen educational outcomes. In many development projects, the outcome (apart from influencing beneficiaries) is to build durable cooperation mechanisms between the persons or organisations involved. The hexagon method can help to identify key categories of project implementation, e.g. cooperation between partners (from the perspective of those involved in the project), and discuss the quality and effectiveness of cooperation. Here, hexagons are used as moderation cards for discussion. Using them, the participants create a map, which is a pretext for discussion. It can focus, for example, on mechanisms of effective (and ineffective) cooperation between partners.

Why use hexagons? Firstly, to define the most important categories for the participants. These describe the evaluated reality, e.g. cooperation environment. And secondly, to hold a discussion organised by these categories and aimed at identifying logical and cause-effect links between the categories (phenomena) being discussed.

Preparation and implementation

Such evaluation consists of several stages. At the first stage, the participants, e.g. representatives of partner organisations in a project, list (writing them down independently on small sheets of paper) all their associations with the word "cooperation". This is how categories for hexagons are formulated. Several such categories can emerge, for instance "division of tasks", "responsibility", "transparency", "information flow", "experience", "being on the same team", "who guides us?", "atmosphere", "decision", "trust", "work organisation", "limited resources", "learning", "rivalry", "money" or "what's going on?". As we can see, the categories can vary, but their very definition already serves as a *de facto* first step in identifying and selecting the topics that are most relevant to the subject being evaluated (in this case the quality of cooperation).

At the second stage, the categories are written down on hexagons. Each participant in the discussion draws a few hexagons displaying individual categories (the hexagons are placed face down).

2 The use of hexagons in evaluation was based on the method of educational hexagons, described on the website of the Centre for Citizenship Education: ceo.org.pl/publikacje/ksiazki-metodyczne-i-pomoce-dydaktyczne/heksy-edukacyjne [accessed: 1 February 2020].

Each participant also receives one blank hexagon – to write down their own category, if they think that a relevant issue is missing and, based on the discussion, they decide it should be introduced.

At the third stage, the participants create a map (in the case of this evaluation – a map of the cooperation environment during project implementation). It should consist of hexagons that fit together. Each hexagon position needs to be justified: Why am I putting this hexagon right here? How does it connect with the hexagons adjacent to it? How do categories in adjacent hexagons interact?

As a result of such an evaluation, many interesting and nonobvious issues concerning the implementation of the project can come to light. In the quoted example of evaluation of cooperation between project partners, its participants admitted that evaluation had never given them so many useful solutions they could apply in subsequent joint projects.

One such nonobvious issue for the participants of the evaluation was the observation that a good flow of information between project partners is correlated with thorough preparation of individual stages of project implementation, already during its planning (the focus of which was to plan activities within the project rather than ponder how the partners will work together). What is also important, effective flow of information not only relates to the division of tasks between the organisations involved, but also requires learning the working style of the partners. A very important event promoting the building of cooperation mechanisms was the kick-off meeting, which was held to establish the principles of project management and team building (although initially most partnership members considered it a waste of time and were reluctant to participate in it). During the evaluation, they understood that the kick-off meeting was the foundation on which they built the whole mechanism of cooperation in the project, and the rules established then became the canon to which all members of the partnership referred. This proved to be very effective due to the fact that the project staff changed many times during the project.

The Success Story method

Use

This method consists in identifying, investigating and describing the most, and sometimes the least, successful cases in a project, e.g. a training. A case can be the story of one person, a team or even a department in the workplace, and any distinguishable unit can be described here (Stake, 2009).

This method is most useful for assessing the impact of development activities such as training, coaching and consulting, although its application can be much wider. It can, for example, be useful when testing the effectiveness of newly implemented processes or even purchasing new equipment.

The purpose of this method is not to assess medium or average efficiency. Its strength is rather to show spectacular results and participant experience. It is useful for documenting the history of a programme's impact and for understanding the factors that strengthen or hinder it. It provides answers to the following questions: When does the programme work? What are its most effective elements? Under what conditions does the programme work best? Who can benefit most from participating in it? As a qualitative method, it can be an excellent complementary material for quantitative research or can be used as a self-contained technique.

Preparation and implementation

When planning research using the Success Story method, we can start with quantitative research to identify the extent to which the business goal of the initiative was achieved throughout the organisation. This data can then be used to find outstanding cases, in which the initiative was particularly successful (or ineffective), in order to study them more closely.

In training evaluation, the Success Story approach can be used in addition to the Kirkpatrick model for the evaluation the results of training (Kirkpatrick, 2001; Kirkpatrick, 2015), which is largely based on quantitative methods. As Andrew Downes aptly notices in *Learning Evaluation Theory: Brinkerhoff's Success Case Method*:

> *Ongoing quantitative analysis of large data sets provides the latest indicators for your initiative's success, as well as trends related to learning and results in your organisation. The qualitative analysis makes it possible to become acquainted with these results and examine their potential causes. You may have heard the statement: 'Correlation does not imply a cause and effect relationship, but suggestively moves the eyebrows and gestures stealthily when saying: look there.' Qualitative research, and especially the Success Story method, allows us to 'look here' with more attention. At this stage of the study, you collect data using in-depth individual interviews. The stories told should cover all four levels of the Kirkpatrick model of evaluation with evidence at each level to show learning outcomes for both individuals and organisations.*

An example of evaluation which used the Success Story method is that of post-graduate studies for managers and staff of enterprises offered by one of Warsaw's universities. As part of the evaluation, participants' career paths were described along with accompanying conditions, including the role of completed post-graduate studies. The obtained material – thanks to taking into account the real perspective of an individual participant and elements of their biography – was not only interesting for research and useful for evaluation, but also constituted one of the most effective elements of promotion and communicating the educational offer of the university. This example shows that the results of such an evaluation can be used not only to improve a programme, but also for marketing purposes.

Short evaluation formats

Below, we can find the techniques for collecting evaluation information useful during training, especially when it is demanding in terms of the content discussed and requires a lot of focus and long--term attention. In such a case, evaluation becomes a break, redirects attention, and is energising. At the same time, it is an excellent way to obtain feedback, as well as reflect on the learning process and the outcomes achieved. It can also be a pleasant element of working together on the quality of training.

BIN, SUITCASE, SACK

Use

Summary of the results of a training at its end.

Preparation and implementation

On a flipchart/board we draw a bin, a suitcase and a sack. Next to the drawings, we can write explanations: **suitcase** – what do I take with me from the training? (what will I use in my work? what do I find particularly useful?); **bin** – what was unnecessary for me during the training? (useless, superfluous information I will certainly forget), and **a sack** – what was missing? (what should have been discussed? what can be included in the next training?).

Participants can write down their opinions on post-it notes and stick them on the drawings or write directly on them. The results should be discussed. It is worth using the discussion to strengthen the outcomes of the training, e.g. by identifying participants' plans to put into practice the knowledge from the suitcase.

SHOOTING TARGET

Use

Quick feedback from trainees (the method can be used as an interim or final evaluation).

Preparation and implementation

On a flipchart/board we draw a circle representing a shooting target. It is a good idea to insert in it numbers from five (centre of the target) to zero (circumference of the target), which will represent rating. With this method we can evaluate one category, e.g. usefulness of the training, as well as a number of areas, for instance by dividing the target into quarters we can evaluate four categories, e.g. usefulness of the training, opportunities for exchanging experiences, atmosphere, organisation of the training (we can, of course, apply other evaluation categories). We can even divide the target into eights (like a cake), however larger numbers of categories will only work in very small groups.

The participants approach the target and indicate their score with a dot (when we have e.g. quarters indicating four different categories, a dot should be placed in each of the quarters). If the target is attached to a cork board, darts or push pins can be used. The results should be discussed paying attention to the distribution of scores; an attempt at interpreting the results should be made.

HAND

Use

Summary of the training and its results at its end.

Preparation and implementation

Participants receive sheets of paper (thicker paper works better here) on which they are asked to draw the outline of their hand. Each finger is assigned one evaluation category, for example:

→ Thumb – best part of the training,
→ Index finger – information that I will pass on to my colleagues,
→ Middle finger – weakest part of the training,
→ Ring finger – an element of training that requires improvement,
→ Little finger – information that I'll keep for myself.

Categories can be defined freely, depending on the needs of a given evaluation. Each participant writes one answer on each finger of their outline, in accordance with the defined categories. This method can be limited to self-reflection by the participants. We can also create a hand exhibition and use the participants' answers to guide the discussion.

PROJECTIVE METHOD OF TEVYE THE DAIRYMAN (who sang "If I were a rich man..." in the play *Fiddler on the Roof*)

Use

Making recommendations for subsequent training sessions or the entire development project.

Preparation and implementation

Participants are asked to put themselves in the coach's or project coordinator's shoes and indicate what they would do in this role:

→ If I were a project coordinator, next time I'd pay attention to...
→ If I were a coach, I should consider...

The roles can be adapted according to needs. The participants write down their proposals for the action they would take in a given role. The results obtained are a pretext to discuss recommendations for the training or project.

Development-related functions of evaluation engaging its participants

The evaluation concepts and techniques presented above assume the subjectivity of evaluation participants and aim at overcoming negative phenomena related to the evaluation practice in which those taking part in it are treated as research subjects. The agency of the participants is based on a slightly different understanding of evaluation from the commonly used one emphasising its assessment and control functions. The presented concepts focus primarily on the pro-developmental function, in accordance with Helen Simons' definition of evaluation, which she terms *a call for development* (Simons, 1987). It redirects the emphasis from independent evaluators to self-evaluation involving those involved in the process – here external evaluators become facilitators of the process.

This approach uses contemporary learning theories, which emphasise learners' responsibility for outcomes, where the learners are largely responsible for the acquisition of knowledge. Therefore, evaluation is to assist them in creating a learning process that is as effective as possible.

The very process of evaluation as critical review is transformed into accompanying evaluation, which focusses on learning and change, rather than on the measurement of outcomes. Here post-evaluation activities are an integral part of the process. Also, *ex post* decisions should be participatory to the maximum. In such an approach, participation is the key word applied at as many stages of the process as possible (from the establishment of evaluation criteria to the development of recommendations).

Currently, in the evaluation of development projects, i.e. training, coaching, advisory and counselling projects addressed to adults, the most commonly used method is a questionnaire. However, it is not always the most appropriate methodological perspective. It does not support integration, participation and subjectivity of development project participants. Socialisation-related functions of evaluation – which are the key to understanding and practical implementation of such approaches as participatory evaluation, developmental evaluation or action evaluation – are achievable thanks to more participatory, creative methods of collecting information presented in this article. In these approaches, evaluation not only plays a research role, but also fulfils an important function of strengthening the ownership of project outcomes, which is particularly important in developmental and educational projects. The techniques for acquiring evaluation information presented above stimulate participants' interest in the evaluation process and make it possible to collect relevant information, deepen reflection, and formulate useful and revealing recommendations. Evaluations carried out using these techniques are an alternative to commonly used questionnaires, and the range of information they help to collect is often much more extensive and deeper than that obtained by means of questionnaires.

The evaluation process is carried out with respect for all concerned. This means that the needs, expectations and opinions of all stakeholders are taken into account, even if they contradict one another. Evaluation is carried out ensuring a sense of trust and security of those who take part in it. Even if critical opinions are expressed, they do not affect the relationship between the participants, but serve to reach a compromise that will satisfy everyone.

Evaluation is carried out with the conviction that the results will actually be used and will contribute to quality- and development-oriented change. Evaluation is not done for the sake of itself – it is the starting point for activities aimed at improving work. This is why it is important to discuss and interpret evaluation results in order to better plan future activities of the organisation.

It is worth involving as broad a range of stakeholders as possible (including managers and other staff) in the process of improving the institution/organisation. They should take part in discussing the results of evaluation and making recommendations for further actions. Such a participatory approach will contribute to building consensus and ownership of change.

Bibliography

Angrosino, M. (2010). *Badania etnograficzne i obserwacyjne*. Warsaw: Wydawnictwo Naukowe PWN.

Barbour, R. (2011). *Badania fokusowe*. Warsaw: Wydawnictwo Naukowe PWN.

Borek, A. & Tędziagolska, M. (2015). *Ewaluacja w szkole i przedszkolu*. Warsaw: Wydawnictwo Ery Ewaluacji.

Borek, A. & Kowalczyk-Rumak, E. (2015). *Różne drogi ewaluacji – poradnik dla dyrektorów szkół i placówek*. Warsaw: Wydawnictwo Ery Ewaluacji.

Brinkerhoff, R.O. (2003). *The Success Case Method: Find Out Quickly What's Working and What's Not*. San Francisco: BK Publishers.

Ciężka, B. (2018a). Ewaluacja w edukacji dorosłych. Nowe trendy, wyzwania, oczekiwania. In: M. Gromadzka (ed.), *Przestrzenie i miejsca edukacji dorosłych w Polsce*. Warsaw: Foundation for the Development of the Education System.

Ciężka, B. (2018b). Ewaluacja szyta na miarę potrzeb, *ROZWÓJ/Miesięcznik Benefit, 1*(70).

Červinková, H. & Gołębniak, B.D. (2010). *Badania w działaniu. Pedagogika i antropologia zaangażowane*. Wrocław: Wydawnictwo Naukowe Dolnośląskiej Szkoły Wyższej.

Denzin, N.K. & Lincoln, Y.S. (2009). *Metody badań jakościowych*. Warsaw: Wydawnictwo Naukowe PWN.

Downes, A. (no year of publication). *Learning Evaluation Theory: Brinkerhoff's Success Case Method*. bit.ly/317RFj0 [accessed: 16 July 2019].

Elliott, J. (1997). Praktyczny poradnik prowadzenia "action research". In: H. Mizerek (ed.), *Ewaluacja w szkole. Wybór tekstów*. Olsztyn: Wydawnictwo MG.

Flick, U. (2010). *Projektowanie badania jakościowego*. Warsaw: Państwowe Wydawnictwo Naukowe.

Flick, U. (2011). *Jakość w badaniach jakościowych*. Warsaw: Państwowe Wydawnictwo Naukowe.

Kirckpatrick, D.L. (2001). *Ocena efektywności szkoleń*. Warsaw: Wydawnictwo Studio Emka.

Kirkpatrick, J. & Kirkpatrick, W. (2015). *An Introduction to the NewWorldKirkpatrick® Model*. Newnan, GA: Kirkpatrick Partners, LLC.

Kvale, S. (2010). *Prowadzenie wywiadów*. Warsaw: Państwowe Wydawnictwo Naukowe.

Mizerek, H. (2017). *Ewaluacja edukacyjna. Interdyskursywne dialogi i konfrontacje*. Krakow: Oficyna Wydawnicza Impuls.

Patton, M.Q. (1997). *Utilization-Focused Evaluation*. Thousand Oaks: Sage.

Patton, M.Q. (2011). *Developmental Evaluation*. New York–London: The Guilford Press.

Simons, H. (1987). *Getting to Know Schools in a Democracy*. London–New York–Philadelphia: The Falmer Press.

Stake, R. (2009). Jakościowe studium przypadku. In: N.K., Denzin & Y.S. Lincoln, *Metody badań jakościowych*. Warsaw: Wydawnictwo Naukowe PWN.

Effectiveness of executive coaching in the light of research findings

DOI: 10.47050/66515246.144-159

Seweryn Krupnik
Ewa Krupnik

Abstract

The aim of this chapter is to provide inspiration and tools for designing and conducting research into the effectiveness of executive coaching. To achieve this, a review of eleven studies using different research approaches, ranging from qualitative to quasi-experimental ones, is presented. This text provides information on the effects found, the results obtained and the factors affecting them, as well as research methodologies and tools. Disseminating executive coaching evaluation methods can contribute to a more evidence-based discussion about it.

The examination of the effectiveness of executive coaching is an important issue in the context of supporting competency development. As executive coaching is becoming more and more popular, there is a need for reflection on its effectiveness and research into this area. This article in instrumental in this regard by providing a description of selected, most frequently quoted studies into executive coaching effectiveness and answers to the following questions about it:

- → What effects are being analysed?
- → What factors influence the effects?
- → What research schemes are used?
- → What are the research findings?

The authors of this chapter want to provide inspiration and tools for designing and conducting research into the effectiveness of executive coaching.

Seweryn Krupnik
Sociologist, evaluator (member of the Polish Evaluation Society), coach (co-founder of O!Coaching), and Assistant Professor at the Department of Sociology of Economy, Education and Social Research Methods of the Jagiellonian University. Project coordinator and expert at the Centre for Evaluation and Analysis of Public Policies at that university. Specialist in the methodology of evaluation research, including the analysis of public support effects on enterprise innovation.

KEYWORDS:
EXECUTIVE COACHING,
COACHING,
EVALUATION,
EFFECTIVENESS

What is coaching?

Some consider it an ideal method of personal development, others treat it as a tool of little use. In recent years, many papers, articles and publications on coaching have been published in Poland (Wujec, 2007;

Ewa Krupnik
ICF coach, trainer, public relations specialist, founder of the web portals Promuj NGO! and O!Pracownia. Originator of the O!Coaching portal, which introduces the idea of coaching to Poland. She coaches individual clients, as well as the staff of public institutions, companies and NGOs. For several years, she has worked as an expert and coordinator of social projects in the third sector in Poland. Since 2015, she has been Vice President of the Laboratory of Social Inspirations Foundation.

Sidor-Rządkowska, 2009; Bennewicz, 2011; Czarkowska & Wujec, 2014). There have also appeared new study and training programmes for coaches, as well as new coaching jobs. All these with the aim to raise the company employee spirit. Although in Poland coaching is still a fledgling profession, in other parts of the world it has been developing for several decades.

Coaching is based on humanistic psychology and its origins date back to the end of the 19th century. However, its proper beginning is dated on early 1970s. At that time, it was used with the aim of improving sporting performance (although the word "coaching" in relation to sports appeared for the first time already in 1861 – *Online Etymology Dictionary*[1], 2019). We can certainly prove that coaching has its roots in sports psychology and that it drew a great deal from this science. In the 1970s and 1980s, other varieties of coaching, such as life coaching, managerial coaching and organisational coaching, developed rapidly in the United States. At the end of the 20th century, both in the United States and in Europe, it was deemed a new tool for facing various types of challenges (Whitmore, 2002).

At present, coaching is a buzz word in personal development and acquisition of new competencies. Coaches are said to change lives, as they motivate their clients to take action and achieve goals. There is some truth in it, but also some degree of imprecision. Let us then start from the beginning, that is with a definition.

Although since the 1970s an intensive development of coaching has been observed, still several definitions are applied to it. Hence, many coaches and trainers postulated that both the vocabulary and definitions be coordinated. The International Coach Federation (ICF) deserves credit for its work in this area. It has developed standards of coaching work and competencies related to it, as well as a definition which is one of the most frequently quoted in literature:

> *Coaching is partnering with Clients in a thought-provoking and creative process that inspires them to maximise their personal and professional potential* (*Code of Ethics...*, 2015).

This definition indicates that coaching can be considered in the context of personal, team and group development. In coaching it is important that there is a change on the client's part. It is a process of providing support in achieving goals and discovering potential, but the most important thing is the commitment and work of the client.

Therefore, it can be noted that coaching differs from counselling, mentoring, therapy or public speeches given by motivational speakers. A coach does not advise, train or pass on new knowledge. A coach is not a therapist. A coach will never say to their client: "you must change your

1 See: www.etymonline.com/word/coach [accessed: 20 February 2020].

life, and here's how you must do it..." Coaching is about regular meetings (about one hour-long sessions), during which the client works on a given subject. A coach is a demanding listener who summarises what the client says and asks questions that help them precisely define their goal and the way of achieving it. A coach helps distinguish what is feasible and what is not. As a result, the client can eliminate elements that can stand in the way to achieving the goal.

How to make the definition of coaching even simpler? We can compare it to **a journey that can be made from point A to point B. This journey produces change. This is done in a professional relationship with a coach who provides support and asks questions.**

And here a sensitive issue arises, namely the abovementioned questions, which are a coach's important and most visible tool. From this perspective, coaching may seem a trivial method – **if it is only about asking questions, then I can ask them myself.**

But this is not the key to coaching. Coaching is not about simply asking questions. Thanks to asking the RIGHT questions and using many tools suitable to the topic the client works on, the coach makes it easier for them to achieve their goal (Bennewicz, 2011; Bird & Gornall, 2016). This way the coach builds a client's awareness and together they strive for change – always in the client, not in the coach. Coaching is a process within a session. **It is a training in awareness that leads to a change in a given person's situation**. It is a change for the better. It is the client who decides what is to be changed, what the current situation is, and why the change is needed. Then, one by one, decisions are made about the desired result of the joint work and about how to determine if the outcome meets the client's expectations. Next, individual work on planning activities is conducted. In this case, available resources are used. At the last stage, it is verified whether these actions lead to a favourable solution. And finally, the goal is achieved.

It is worth adding that coaching can only be successful when both parties are committed to the process. Therefore, it is important for the client to strive for change, and for the coach to support the client in such a way that they can achieve the goal (Whitmore, 2002).

What is executive coaching?

One of the many difficulties faced by companies in Poland is adapting to frequently changing market conditions. For many executives, this is a challenge, as it entails a constant development of competencies and knowledge, as well as a change in mentality and company management methods. To introduce new activities, managers need to acquire new personal and social competencies. In addition, changing the old model of company management and attitudes is often an urgent need that cannot be met within e.g. a month. It requires much more time and awareness on the part of managers (Penc, 2005).

There are different types of coaching: life coaching, business coaching, career coaching, etc. One of them is executive coaching, i.e. coaching for the managerial and executive staff. It is most often used by key managers, senior directors, presidents and owners of companies. Executive coaching is considered *the science and art of the facilitation of personal and professional development, a manager's learning and acting by expanding opportunities offered by their natural behaviour* (Dembkowski et al., 2010). Therefore, unlike coaching activities aimed at improving management, this process aims to impact on the development of the manager's

potential. Executive coaching is a specific type of coaching and a coach needs to have experience in working with managers. The clients are senior staff members and business professionals who need to focus on resources and skills allowing them to meet business challenges. Here, a coach can support such activities as: strategic planning, strategy monitoring, operational implementation, as well as dealing with loneliness in decision-making or excessive stress. Often companies also employ coaches to support the implementation of organisational changes, including preparations for staff reductions (ibidem).

Effects of executive coaching

Thanks to improving their competencies, executive coaching makes it easier for managers to create a new working environment and well-coordinated teams, and to manage them with the observance of high standards. Presidents and managers can learn both new behaviours towards employees and new ways of managing the company. This is a key competency in adult education, as such superiors are role models for their subordinates and influence their behaviour.

Executive coaching focusses on working with managerial staff using analysis of their needs, capabilities, resources, and way of thinking and acting. All the aspects of executive coaching discussed above often consist in changing the attitude of the coached leader and helping them develop a new model of work and cooperation with subordinates. Just like in the case of the traditional coaching process, also in executive coaching it is important to go through six stages:

1. define the subject and purpose of coaching,
2. agree on the process, results and developmental needs,
3. develop a coaching plan,
4. implement the planned activities,
5. monitor the activities,
6. verify the results at the end of the process (Thorpe & Clifford, 2004).

The effects of executive coaching can also be analysed using the Kirkpatrick model (Kirkpatrick & Kirkpatrick, 2006). It includes four levels of training results evaluation:

Level one – Reaction. How does the client feel about the process? Is there anything they find difficult? Do they have energy and motivation to act? Or, on the contrary, are they tired, frustrated, etc.?

Level two – Learning. What has the client learnt? What have they practised? Have they completed all the tasks?

Level three – Behaviour. How has the client's work been put into practice? For instance, have they changed their management model?

Level four – Results. What are the practical results of coaching? For example, does the client manage their time more effectively? Has the expected change in terms of establishing relationships within the company taken place?

In order to achieve a result at a given level, it is necessary to have an effect at the previous level.

Effectiveness of executive coaching

Effectiveness is one of the possible criteria for the evaluation of executive coaching as a developmental activity. It refers to the extent to which the activity leads to the desired effects (objectives). Gil Bozer, James C. Sarros and Joseph C. Santora (2012) define it as the extent to which effects are achieved at an individual level thanks to participation in the coaching process. One of the main challenges in analysing effectiveness is identifying and measuring the effects. Approaches proposed by Kirkpatrick (after: MacKie, 2007), as described above in this chapter, and Thomas R. Guskey (1998) provide some framework for this. However, they need to be operationalised in the context of a given research. This is so because the coaching process can concern various effects, and a specific study can only present a certain limited scope. For example, Anthony M. Grant (2012) proposes to distinguish between direct and indirect effects (after: Bozer et al., 2014). The former take into account direct individual changes at the level of attitudes and behaviours, such as positive emotions associated with the organisation, increased self-awareness and better competence development. The latter refer to better functioning as an individual and can lead to the success of both the individual and the organisation. Bozer, Sarros and Santor (2014) included in their study such variables as career satisfaction, commitment to work and results of evaluations conducted by superiors.

It is worth noting that it is also possible to examine the coaching process from the perspective of other evaluation criteria, such as:

- → relevance – to what extent does coaching, as a form of developmental activity, respond to the client's needs?
- → usefulness – to what extent do the effects of coaching meet the client's needs?
- → effectiveness – does coaching constitute the most effective use of resources (time, money) to enhance specific competencies?
- → sustainability – to what extent has coaching led to permanent changes (e.g. one year after the process was completed)?

Another important aspect – after properly defining coaching effects – is measuring them (MacKie, 2007). There are two most important challenges typical for evaluation research. The first is relying solely on evaluations conducted by clients, which can lead to bias in the evaluation of results. Clients can misjudge the change. Also, feeling the need to show gratitude to the coach, they can – more or less consciously – give a better rating in order to make them feel good. As a way to overcome this, we can use other sources of data (a superior's evaluation or 360-degree evaluation, which also takes into account the opinions of colleagues and subordinates). The second challenge relates to changes which occurred during the coaching yet can result from factors other than the process itself, e.g. from other developmental or organisational activities. This risk can be minimised by introducing a quasi-experimental research scheme, which examines other people who are as similar to the clients as possible.

Selection of articles for the review

The table below presents an outlay of articles summarising executive coaching effectiveness studies. The articles were selected based on how many times they were quoted in the Google Scholar search

engine. The following keywords were used in the search: *effectiveness* and *executive coaching*. Major limitations of this method should be clearly highlighted, as they make the review neither exhaustive, nor representative. There already exist systematic reviews and even meta-analyses of this type of research (De Meuse et al., 2009). The purpose of the presented list is to illustrate the most frequently cited research. This way, articles describing studies from five countries and one international study were selected. Because of the selection criterion, the collected articles did not feature the latest and most methodologically advanced publications. To overcome this limitation, one research review (meta-analysis) was added to the list. The presented studies use different research approaches, ranging from in-depth interviews with clients through survey questionnaire and encoding of actual behaviour to quasi-experimental research schemes and random selection of the control group.

Results of the review

In the vast majority of the research, the developmental process itself is defined as executive coaching. In some cases, this description has been made more detailed by indicating the specific coaching approach used. At the same time, the diversity of analysed processes makes it possible to much better show the effects that these processes were supposed to lead to. We can indicate approaches in which the effect was defined very broadly (e.g. transformational leadership behaviour), those in which categories of effects were presented (e.g. awareness-raising, better setting of own goals) or those in which very specific skills were aspired to (moderating meetings). In some studies, effects were not assumed, but the coached clients were asked about them. This diversity of effects confirms great variety of coaching processes, which on the one hand is valuable, as it stands for following the client's goals and needs, but on the other hand forms a research challenge. It indicates that these are diversified processes, hard to reduce to a common denominator.

At the same time, once the effect has been determined, the way it is measured is quite often based on a standard procedure used in social research, i.e. scales developed and tested by other researchers. Sometimes researchers create their own measuring tools. Exceptions include the situation in which the client is asked about the effect using open questions (which is a standard technique used within the coaching process itself).

The results almost always confirm the positive effects of coaching. Interestingly, in the case of more advanced, quasi-experimental schemes, positive effects are still identified, but often not for all coaching goals. Less spectacular effects achieved in counterfactual impact research are typical in evaluation studies. This phenomenon is described by the iron law of evaluation, according to which the more accurate the measurement, the greater the chance that the measured value is zero (Rossi, 1987).

The research studies also identified factors affecting the volume of the effect. They refer to:

- → the client's attitude (including expectations and opinion on the credibility of the coach);
- → the coach's characteristics (skills, including these related to relationship building and pushing the client out of their comfort zone, whether the coach is internal or external, knowledge of various techniques, background in psychology);
- → the degree of matching between coach and client;
- → the quality of the relationship between the coach and the client (e.g. trust);

- → The process itself (number of sessions, compliance with procedures defining the process, time adjustment);
- → Support from the client's superiors in implementing the effects;
- → The organisational culture (and the coach's knowledge of it).

It is also worth commenting on methodological aspects of the research. As has already been indicated, they are very diverse. At the same time, even the most advanced ones – described as quasi-experimental – involve quite significant limitations related both to small samples and to the not-so-obvious and unverified assumption that members of the control group are comparable to clients actually taking part in the coaching. These limitations can significantly disturb research findings.

Table 1. Review of executive coaching effectiveness studies

No.	Article	Action	Effect	Method of measuring the effect	Result	Factors affecting effectiveness	Methodology	Country	Sample
1	Thach, 2002	Executive coaching	Effectiveness of a leader's actions	Own evaluation, comparison of 360-degree pre- and post--evaluation results, qualitative data	Positive, increase in ratings by about 50%	Number of sessions, coach matching	Action research, the scale for evaluating a leader's effectiveness developed for the company's needs is the main benchmark	United States	281 senior management staff members in a telecommunications company
2	Perkins, 2009	Executive coaching	Moderation of meetings (more attention paid by the leader to the meeting process, and not only to its substantive contents)	The main tool is the Meeting Leadership Measurement System (MLMS), qualitative data	Positive, statistically significant differences, as planned, in pre- and post-coaching behaviour	No impact of age and IQ	Encoding the number of specific behaviours during pre- and post--coaching meetings	United States	21 senior management staff members from various companies
3	MacKie, 2014	Strength--based executive coaching	Transformational leadership behaviour	Measurement tools (full range leadership model), 360-degree assessments at three points in time	Positive	Compliance with the procedures for strength-based coaching (verification using a checklist)	The research scheme included two cohorts: one was coached first and the other was coached second. When members of the first group were involved in the process, members of the second were treated as a control group. Next, the cohorts changed their roles	Australia	37 senior management staff members in an organisation operating in the not-for-profit sector
4	Paige, 2002	Executive coaching	Reported by the clients, both positive and negative ones	Clients' statements	Qualitative data	Organisational culture (and its understanding by the coach), external vs. internal coach, trust, matching in time, coach's characteristics: building and maintaining a relationship based on trust, ability to push the client out of their comfort zone	2 in-depth interviews with each client, using the Guskey evaluation model as a conceptual framework	Australia	5 respondents from the public and private sector who used executive coaching
5	Evers, et al., 2006	Executive coaching based on the GROW model	Setting own goals, acting in a balanced manner and mindful working and living	Clients' responses to a questionnaire	Positive for some of the results (setting own goals, acting in a balanced way)	None	Quasi-experimental scheme: treated and control group, measurements at two points in time	The Netherlands	30 managers

No.	Article	Action	Effect	Method of measuring the effect	Result	Factors affecting effectiveness	Methodology	Country	Sample
6	Bozer et al., 2014	Executive coaching	Performance at work, evaluation of performance conducted by a superior, self-awareness, emotional involvement at work, job satisfaction	Previously created by other researchers and scales tested for each effect	Positive	Coach's background in psychology, their perceived credibility	Quasi-experimental scheme: treated and control group, measurements at two points in time	Israel	Clients of 4 companies offering executive coaching services, 72 people were in the treated group. Moreover, 68 coaches, 28 superiors of the clients, and 29 people in the control group took part in the survey
7	Bozer & Sarros, 2012	Executive coaching	Performance at work, evaluation of performance conducted by a superior, self-awareness, emotional involvement at work, job satisfaction	Previously created by other researchers and scales tested for each effect	Positive, higher job satisfaction in the treated group	None	Quasi-experimental scheme: treated and control group, measurements at two points in time	Israel	197 clients of 4 companies offering executive coaching services, 68 coaches, 29 colleagues (control group) and 28 direct supervisors
8	Kombarakaran et al., 2008	Executive coaching	5 areas of change identified using statistical analysis (analysis of the main components): effective HR management, better relationships with managers, improved target setting and identification of priorities, greater commitment and productivity, more effective dialogue and communication	Questionnaire for coaches and clients with answers measured on a 5-point Likert scale	Positive	Selection of appropriate coaches, support for the process in the organisation, involvement of superiors		Various	42 coaches, 114 clients from one company

No.	Article	Action	Effect	Method of measuring the effect	Result	Factors affecting effectiveness	Methodology	Country	Sample
9	Boyce et al., 2010	Executive coaching	Various effects identified using the Kirkpatrick model	Measurement on a 7-point Likert scale	Positive	Quality of relationships, cognitive and behavioural match, the coach's credibility	Online questionnaire filled in by the clients and coaches after the process was completed	United States	74 pairs (client and coach), clients were U.S. Army staff members
10	de Haan et al., 2011	Executive coaching	Has not been investigated	N/A	N/A	Effectiveness does not depend on the choice of a specific technique or approach, but on the quality of the relationship and positive expectations, at the same time the coaches' knowledge of various techniques is important from the point of view of choosing the right one for a given client	Questionnaire (163 closed and 3 open questions)	United Kingdom	71 clients at the start of the coaching process and 31 of them 6 months after its completion, the clients analysed in the survey came from various backgrounds
11	Burt & Talati, 2017	Executive coaching	Quality of life, depression, achievement of goals, performance at work	Various measuring tools	Moderate effect, statistically significant	No statistically significant factors	Meta-analysis based on publications that met the following conditions: 1) inclusion of a random control group in the research scheme, 2) reporting of relevant statistical data enabling comparison, 3) written in English	United States, Australia	662 respondents (11 studies in total)

Conclusions

More and more often, managers focus on improving their competencies. Thanks to coaching, managerial staff can continuously develop, and thus prepare for new challenges that are associated with the volatility of contemporary organisations.

Evaluation of effects is an essential element of each coaching process. At the same time, it is embedded in the process itself and in the coach-client relationship. This results in relevant constraints and the need to conduct research using different perspectives and methodologies. The aim of this chapter was to provide inspiration and tools for designing and conducting research into the effectiveness of coaching. To achieve this aim, 11 articles with different research approaches were presented. This text provides information on the effects found, the results obtained and the factors affecting them, as well as on research methodologies and tools.

The results of the review also allow for some important conclusions to be drawn. Research findings show that professionally conducted coaching is effective. What is important, this result is based on effects evaluated not only by clients themselves, but also by their colleagues, and on comparing clients' evaluations with these of other people who did not take part in the coaching process at the same time (with random selection for the coaching process).

Simultaneously, the use of more precise research methods shows lower effectiveness than in the case of research based solely on evaluations made by clients. Delivering the least risk-incurred results requires the use of advanced research schemes and measurement tools. The table below presents a proposal for their division into categories adapted from a different area of evaluation research.

Table 2. Proposed division into categorise of executive coaching evaluation studies

No.	Subject of the study	Examples of questions	Method of conducting the study	Major challenges
1	Process	What was the clients' profile? How many hours of support did they obtain?	Reporting data analysis	Does not apply to effects
2	Feedback from clients	Are they satisfied with the change after coaching? What changes can they see?	Interviews (in-depth or questionnaire-based) with clients after the coaching process	Limited client objectivity, not taking into account other factors that could have influenced the effects, not considering the situation prior to the coaching process
3	Clients' opinions on the changes resulting solely from the coaching process	According to the clients, what changes occurred solely thanks to the coaching process?	Questionnaire (before and after)	Limited client objectivity
4	Comparing clients' performance with that of other employees in the organisation	How did clients' ratings change as compared to other employees?	Questionnaire (before and after) for clients and other employees of the company (or e.g. comparison of ratings from 360-degree feedback)	Clients are rarely typical employees
5	Comparing clients' results with those of similar employees who were not involved in developmental activities at the same time	How did clients' ratings change as compared to other employees?	Questionnaire (before and after) for clients and other employees of the company (or e.g. comparison of ratings from 360-degree feedback)	It can be challenging to ensure that the employees are actually similar
6	Comparing clients' results with those of employees who were almost identical from the statistical point of view and who were not involved in developmental activities at the same time	How did clients' ratings change as compared to the control group?	Questionnaire (before and after) for clients and other employees of the company (or e.g. comparison of ratings from 360-degree feedback) Random selection for the coaching process or use of statistical matching to select members of the reference group (e.g. propensity score matching)	Major logistical challenges in conducting studies designed in this way

Source: authors' own work based on the division into categories prepared for the evaluation of enterprise support by Storey, 2008.

It is clear that when deciding on a survey scheme, we need to strike a compromise between the accuracy of measurement and its logistics. Therefore, when conducting each individual study, it is a common-sense solution to apply the most advanced research methods while taking into account logistical conditions.

At the same time, it should be stressed that this review shows that a rigorous examination of the effectiveness of coaching – despite the uniqueness of each process and the diversity of its objectives – is possible. Such research can provide important feedback to the whole coaching community. It can give coaches significant information, which can contribute to improving the quality of their activities and their observance of the highest standards.

It is also worth noting the limitations of this review. Firstly, it did not aspire to provide a comprehensive description of research conducted in a given area. Secondly, because of the way the articles were selected, relatively little attention was paid to the latest research using the most advanced schemes.

This chapter presents recommendations for coaches and researchers interested in analysing the effects of coaching for purely practical as well as cognitive reasons. It shows research methods, presents their evaluation from the point of view of eliminating errors in measuring the effects, and justifies the need to conduct such studies. At the same time, in the context of the presented research, the potential for raising the quality of studies of the effects of coaching is shown.

Dissemination of executive coaching evaluation methods is also important due to the controversy which accompanies this approach to supporting competencies. Thanks to the use of reliable methods for measuring its effects, we can discuss coaching in the context of evidence, and not just opinions.

Bibliography

Bennewicz, M. (2011). *Coaching i mentoring w praktyce*. Warsaw: G+J Gruner + Jahr Polska.

Bennewicz, M. (2018). *Coaching. Złote zasady*. Warsaw: Onepress.

Bird, J. & Gornall, S. (2017). *Sztuka coachingu. Zbiór narzędzi i wskazówek*. Łódź: Galaktyka.

Boyce, L.A. et al. (2010). Building successful leadership coaching relationships: Examining impact of matching criteria in a leadership coaching program, *Journal of Management Development*, *29*(10), pp. 914–931. DOI: 10.1108/02621711011084231

Bozer, G. & Sarros, J.C. (2012). Examining the Effectiveness of Executive Coaching on Coachees' Performance in the Israeli Context, *International Journal of Evidence Based Coaching and Mentoring*, *10*(1), pp. 14–32.

Bozer, G. et al. (2014). Academic background and credibility in executive coaching effectiveness, *Personnel Review*, *43*(6), pp. 881–897. DOI: 10.1108/PR-10-2013-0171

Burt, D. & Talati, Z. (2017). The unsolved value of executive coaching: A meta-analysis of outcomes using randomised control trial studies, *International Journal of Evidence Based Coaching and Mentoring*, *15*(2), pp. 17–24. DOI: 10.24384/000248

Czarkowska, L.D. & Wujec, B. (2011). Kiedy możliwa jest zmiana? Od struktury sesji do energii zmiany w Coachingu, *Coaching Review*, *3*.

Dembkowski, S. et al. (2010). *Coaching kadry kierowniczej*. Warsaw: Wydawnictwa Profesjonalne PWN.

De Haan, E. et al. (2011)."Executive coaching in practice: What determines helpfulness for clients of coaching?, *Personnel Review*, *40*(1), pp. 24–44. DOI: 10.1108/00483481111095500

De Meuse, K.P. et al. (2009). Evaluating the effectiveness of executive coaching: beyond ROI?, *Coaching: An International Journal of Theory, Research and Practice*, *2*(2), pp. 117–134.

Evers, W.J.G. et al. (2006). A quasi-experimental study on management coaching effectiveness, *Consulting Psychology Journal*, *58*(3), pp. 174–182. DOI: 10.1037/1065-9293.58.3.174

Grant, A.M. (2012). An integrated model of goal-focussed coaching: an evidence-based framework for teaching and practice, *International Coaching Psychology Review*, *7*(2), pp. 146–165.

Guskey, T.R. (1998). The age of our accountability, *Journal of Staff Development*, *19*(4), pp. 36–44.

Kirkpatrick, D.L. & Kirkpatrick, J.D. (2006). *Evaluating training programs: The four levels.* San Francisco: Berrett-Koehler.

Code of Ethics (2015). *International Coach Federation*. coachfederation.org/code-of-ethics [accessed: 26 August 2019].

Kombarakaran, F.A. et al. (2008). Executive Coaching: It Works!, *Consulting Psychology Journal*, *60*(1), pp. 78–90. DOI: 10.1037/1065-9293.60.1.78

Kukiełka-Pucher, D. (2009). Coach-klient: etapy budowania relacji. In: M. Sidor-Rządkowska (ed.), *Coaching: teoria, praktyka, studia przypadku*. Krakow–Warsaw: Wolters Kluwer Polska.

Law, H. et al. (2010). *Psychologia coachingu*. Warsaw: Wydawnictwo Naukowe PWN.

MacKie, D. (2007). Evaluating the effectiveness of executive coaching: Where are we now and where do we need to be?, *Australian Psychologist*, *42*(4), pp. 310–318.

MacKie, D. (2014). The Effectiveness of Strength-Based Executive Coaching in Enhancing Full Range Leadership, *Consulting Psychology Journal: Practice & Research, 66*(2), pp. 118–137. DOI: 10.1017/CBO9781107415324.004

Online Etymology Dictionary (2019). "coach", www.etymonline.com/word/coach [accessed: 26 August 2019].

Paige, H. (2002). The impact of executive coaching on leadership effectiveness, *Leadership & Organization Development Journal, 23*(4), pp. 205–214.DOI: 10.1108/01437730210429070

Penc, J. (2005). *Role i umiejętności menedżerskie*. Warsaw: Difin Counselling and Information Centre.

Perkins, R.D. (2009). How Executive Coaching Can Change Leader Behavior and Improve Meeting Effectiveness: An Exploratory Study, *Consulting Psychology Journal, 61*(4), pp. 298–318. DOI: 10.1037/a0017842

Ramirez-Cyzio, K. (2013). Leadership Coaching w podejściu systemowym. In: L.D. Czarkowska (ed.), *Leadership coaching jako odpowiedź na wyzwania współczesnego świata*. Warsaw: Poltext.

Rossi, P. (1987). The Iron Law Of Evaluation And Other Metallic Rules, *Research in Social Problems and Public Policy, 4*, pp. 3–20.

Scherer, J. & Barnes, A. (2013). Coaching liderów jest inny. In: L.D. Czarkowska (ed.), *Leadership coaching jako odpowiedź na wyzwania współczesnego świata*. Warsaw: Poltext.

Sidor-Rządkowska, M. (2009). *Coaching. Teoria, praktyka, studia przypadków*. Warsaw: Oficyna Wolters Kluwer Business.

Storey, D. (1998). *Six Steps to Heaven: Evaluating the Impact of Public Policies to Support Small Business in Developed Economies*, Warwick University Business School, Small and Medium--sized Enterprise Centre, Working Paper No. 59.

Thach, E.C. (2002). The impact of executive coaching and 360 feedback on leadership effectiveness, *Leadership & Organization Development Journal, 23*(4), pp. 205–214. DOI: 10.1108/01437730210429070

Thorpe, S. & Clifford, J. (2004). *Podręcznik coachingu. Kompendium wiedzy dla trenerów i menedżerów*. Poznań: Dom Wydawniczy REBIS.

Whitmore, J. (2002). *Coaching for Performance: Growing People, Performance and Purpose*. London: Nicholas Brealey Publishing Ltd.

Wilczyńska, M. et al. (2011). *Moc coachingu. Poznaj narzędzia rozwijające umiejętności i kompetencje osobiste*. Gliwice: Helion.

Wujec, B. (2007). *Kadry w Polsce. Nowoczesna Firma*. Warsaw: Instytut Zarządzania.


frse
Foundation for the Development
of the Education System


Erasmus+


European
Funds
Knowledge Education Development


Republic
of Poland


European Union
European Social Fund

This publication is the result of cooperation between the Foundation for the Development of the Education System (FRSE) and the Polish Evaluation Society (PTE). Texts gathered in this volume raise questions about the condition of evaluation in Poland. They also cover methodological issues and other aspects of evaluators' work and may support the development of their professional competences and work techniques. The 20th anniversary of the Polish Evaluation Society is also an opportunity to inspect evaluation practices carried out in Poland currently. The authors of the articles hope that they will inspire the development of evaluators' work, broaden the spectrum of research methods that can be used in the evaluation of educational programs or activities, and contribute to the further development of the evaluation culture in Poland.

The Polish Evaluation Society is an independent non-governmental organization that for last 20 years has been creating the space for the discussion on evaluation and its role in the modern world. Among experts contributing to the Polish Evaluation Society research there are scientists representing the academic world, companies and research agencies, as well as central and regional publicadministration institutions. PTE also conducts training and takes part in the public debate on the quality of public policies and directions of improving administration, scientific institutions and non-governmental organizations.

Foundation for the Development of the Education System operates since 1993 and is the Polish National Agency (NA) of the Erasmus+ Programme, implemented in the years 2014–2020, as well as the Polish NA of the European Solidarity Corps. FRSE is responsible for other European educational and information initiatives in Poland, such as eTwinning, Eurodesk, Eurydice, Europass, ECVET and EPALE. The Foundation also supports cooperation with countries in the East via the Polish Lithuanian Youth Exchange Fund, the Polish Ukrainian Council of Youth Exchange, SALTO-EECA Eastern Europe and Caucasus Resource Centre. Since 2014, FRSE has been involved in the implementation of the Operational Programme Knowledge Education Development.