

# EWALUACJA RPO WK-P

## NOWE WYZWANIA DLA EWALUACJI PROGRAMÓW OPERACYJNYCH W PERSPEKTYWIE 2014-2020

Praca zbiorowa

Urząd Marszałkowski  
Województwa Kujawsko-Pomorskiego  
w Toruniu



Fundusze Europejskie  
Program Regionalny



WOJEWÓDZTWO  
KUJAWSKO-POMORSKIE

Unia Europejska  
Europejskie Fundusze  
Strukturalne i Inwestycyjne





**Urząd Marszałkowski**  
**Województwa Kujawsko-Pomorskiego**

**EWALUACJA RPO WK-P**

**NOWE WYZWANIA DLA EWALUACJI  
PROGRAMÓW OPERACYJNYCH  
W PERSPEKTYWIE 2014-2020**

**Praca zbiorowa**

**Toruń 2016**

**Praca zbiorowa****Redakcja techniczna:**

Mirela Jamroży, Katarzyna Kasprzyk

**Przygotowanie, organizacja i prowadzenie Konferencji Ewaluacyjnej:**

Agata Czajkowska

Mirela Jamroży

Katarzyna Kasprzyk

Jolanta Konkel

Karol Kwiatkowski

Paweł Peńsko

Jolanta Rudnicka

© **Urząd Marszałkowski Województwa Kujawsko-Pomorskiego w Toruniu  
Toruń 2016**

**ISBN:** 978-83-65206-00-8

**Recenzent:**

prof. dr hab. Andrzej Miszczuk

**Wydawca:**

Urząd Marszałkowski Województwa Kujawsko-Pomorskiego w Toruniu

Plac Teatralny 2

87-100 Toruń

[www.mojregion.eu](http://www.mojregion.eu)

Poglądy i opinie wyrażone w publikacji nie odzwierciedlają stanowiska Zarządu Województwa Kujawsko-Pomorskiego, a jedynie stanowiska autorów.

Departament Rozwoju Regionalnego

Wydział Analiz i Ewaluacji

Biuro Ewaluacji

tel. 56 62 18 653

[www.mojregion.eu](http://www.mojregion.eu)

**Skład, przygotowanie do druku:**

Michał Wcisło

**Druk i oprawa:**

Alnus Sp. z o.o.

ul. Wróblowicka 63, 30-698 Kraków

**EGZEMPLARZ BEZPŁATNY**

# SPIS TREŚCI

Słowo Wstępne.....	5
WPROWADZENIE .....	7

## Aspekty teoretyczne i metodologiczne

### PAWEŁ CHURSKI

Polityka spójności a przemiany przestrzeni społeczno-ekonomicznej Polski w wynikach badań regionalnych.....	15
---	----

### KORNELIUSZ PYLAK

Czy badania ewaluacyjne mogą być bardziej naukowe a nauka bardziej użyteczna dla badań ewaluacyjnych? .....	45
---	----

### JAROSŁAW GÓRNIAK

Szacowanie przyczynowego wpływu interwencji: funkcje, walory, ograniczenia oraz relacje z innymi podejściami w ewaluacji.....	61
---	----

### RAFAŁ TRZCIŃSKI

Ewaluacja oparta na teorii w praktyce badawczej.....	73
--	----

### MONIKA KWIECIŃSKA-ZDRENKA

Jak nadążyć za zmianą? Rola ewaluacji rozwojowej w wypełnieniu luki w badaniach złożonych procesów interwencji publicznej.....	93
--	----

### DARIUSZ MAJEREK

Zastosowanie techniki Propensity Score Matching w badaniach ewaluacyjnych.....	105
--	-----

### BOGDAN JANKOWSKI I MAGDALENA BŁASZCZYK

Socjologia wizualna i big data jako nowe narzędzia ewaluacji .....	119
--	-----

### BARBARA LESZCZYŃSKA

W poszukiwaniu lepszych rekomendacji... Jak wykorzystać podejście Design Thinking w pracy ewaluatora.....	131
---	-----

## Aspekty praktyczne

### WERONIKA FELCIS

Standardy ewaluacji narzędziem profesjonalizacji procesu ewaluacyjnego – aktualizacja i rozwój.....	145
---	-----

### ANDRZEJ REGULSKI

Partycypacyjny model ewaluacji w perspektywie finansowej 2014-2020. Refleksje wykonawcy.....	159
--	-----

### PIOTR STRONKOWSKI

Ewaluacja polityki spójności w Polsce z perspektywy ewaluatora. Doświadczenia i perspektywy.....	169
--	-----

### AGNIESZKA ZAWRZYKRAJ I BARBARA PORĘBSKA-MACIOŁEK

Diagnoza problemów i środków zaradczych w perspektywie 2014-2020 na przykładzie doświadczeń Małopolskiej Jednostki Ewaluacyjnej .....	181
---	-----

### MONIKA LEMKE

Kryteria wyboru wykonawcy jako czynnik jakości ewaluacji.....	193
---	-----

### JOLANTA RUDNICKA

Wykorzystanie doświadczeń w zakresie ewaluacji Regionalnego Programu Operacyjnego Województwa Kujawsko-Pomorskiego na lata 2007-2013 w nowej perspektywie finansowej UE.....	205
--	-----



### **Szanowni Państwo!**

*Mam przyjemność zaprezentowania publikacji będącej podsumowaniem III Międzyregionalnej Konferencji Ewaluacyjnej „Nowe wyzwania dla ewaluacji programów operacyjnych w perspektywie 2014-2020”, która odbyła się w czerwcu 2016 r. pod honorowym patronatem Ministerstwa Rozwoju i Polskiego Towarzystwa Ewaluacyjnego.*

*Czas wybrano nieprzypadkowo. Początkowa faza wdrażania programów operacyjnych perspektywy finansowej 2014-2020 to dobry moment, by wykorzystując doświadczenia z ewaluacji programów operacyjnych lat 2007-2013 rozmawiać na temat wyzwań, przed jakimi stoi obecnie system ewaluacji polityki spójności.*

*Podczas dwudniowego spotkania przedstawiciele środowisk naukowych, samorządowych oraz doświadczeni ewaluatorzy dyskutowali na temat możliwie najskuteczniejszego wykorzystania wyników badań ewaluacyjnych w procesie programowania i skutecznego wdrażania unijnego wsparcia. Poruszono przede wszystkim tematykę związaną z podejściem teoretycznym w ewaluacji oraz metodologią zbierania danych i prezentowania wyników badań ewaluacyjnych, a także z rzucającymi nowe światło praktycznymi aspektami prowadzonych badań.*

*Z nadzieją, że zebrane artykuły dostarczą Państwu wiedzy i przydatnych inspiracji, zapraszam do lektury.*



*Piotr Czarnecki*

Marszałek  
Województwa Kujawsko-Pomorskiego



Niniejsza publikacja jest podsumowaniem III Międzyregionalnej Konferencji Ewaluacyjnej pn. „Nowe wyzwania dla ewaluacji programów operacyjnych w perspektywie 2014-2020”, która odbyła się w dniach 21-22 czerwca 2016 roku w Centrum Kulturalno-Kongresowym Jordanki w Toruniu pod patronatem Ministerstwa Rozwoju i Polskiego Towarzystwa Ewaluacyjnego. Była to kolejna udana konferencja, w której uczestniczyli przedstawiciele jednostek ewaluacyjnych, a także środowisk naukowych i badawczych z całego kraju.

Zainteresowanie konferencją wskazywało na rosnące potrzeby działań ewaluacyjnych i jednocześnie odzwierciedlało zwiększenie znaczenia ewaluacji w procesie analizy skuteczności, efektywności i wpływu realizacji programów operacyjnych na rozwój regionów. Pozytywnym zjawiskiem, stało się zastosowanie ewaluacji (a nie oceny) jako narzędzia planowania i projektowania programów realizowanych w perspektywie 2014-2020. Ujęcie tak prowadzonej ewaluacji w strategicznych i operacyjnych dokumentach wskazało na ogrom prac do jakich zobowiązane są i będą środowiska badawcze i ewaluacyjne. Wiele kwestii dotychczas pomijanych lub traktowanych pobieżnie, związanych z ewaluacją wpływu programów na rozwój regionalny, poprzez odpowiednie zapisy zostało sprowadzone do rangi zobowiązań. Stawia to poważne wyzwanie pod kątem przygotowania, a także przeprowadzenia ewaluacji dla programów realizowanych w latach 2014-2020.

Publikacja obejmuje jedynie niektóre koncepcje i cenne refleksje na temat roli ewaluacji podejmowane na konferencji w Toruniu. Książka składa się z 14 artykułów podzielonych na rozważania teoretyczne i metodologiczne oraz praktyczne.

W artykule **prof. dr hab. Pawła Churskiego** poruszona została tematyka wyzwań dla polityki spójności w kontekście nowego spojrzenia na rozwój regionalny. Przedstawia on wybrane wyniki badań realizowanych przez Zespół Zakładu Analizy Regionalnej Instytutu Geografii Społeczno-Ekonomicznej i Gospodarki Przestrzennej Uniwersytetu im. Adama Mickiewicza w Poznaniu w kontekście zróżnicowań rozwojowych występujących w polskiej przestrzeni oraz prawidłowości procesu absorpcji środków europejskich przeznaczanych zgodnie z założeniami polityki spójności na ograniczanie tych różnic<sup>1</sup>. Dzięki tej pracy można sobie wyobrazić skalę skuteczności dotychczasowych działań interwencyjnych oraz otworzyć się na wyzwania polityki spójności w Polsce i poszukiwanie nowych podejść wyjaśniających procesy rozwoju.

---

1 Churski P., *Polityka spójności a przemiany przestrzeni społeczno-ekonomicznej Polski w wynikach badań regionalnych*. (Artykuł z niniejszej publikacji).



W kolejnym artykule, **dr inż. Korneliusz Pylak** przeanalizował sposób prowadzenia badań ewaluacyjnych w Polsce. Zauważa on, że proces przygotowania i realizacji badania ewaluacyjnego pomija pewne elementy prowadzenia rzeczywistego procesu badawczego opartego o teorie i hipotezy naukowe dostępne w literaturze przedmiotu. Następnie przedstawia analizę możliwości zastosowania podejścia naukowego w takich badaniach.

Za rozszerzenie tych rozważań można uznać kolejne artykuły dotyczące ewaluacji opartej na teorii programu oraz kolejności metod kontrfaktycznych stosowanych w badaniach ewaluacyjnych. Teksty **prof. dr hab. Jarosława Górniaka** oraz **Rafała Trzcińskiego** oscylują wokół konieczności stosowania szerszego spojrzenia na informacje zbierane w badaniach ewaluacyjnych przy pomocy najpopularniejszych metod. Prof. Górniak podkreśla, że *„ilościowy obraz programu i jego efektów jest raczej powierzchowną rejestracją wartości założonych wskaźników, a nie modelowym ujęciem mechanizmu przyczynowego prowadzącego od interwencji do pożądanej zmiany”*<sup>2</sup>. Wskazuje przez to ograniczenia metod kontrfaktycznych pozbawionych mechanizmów przyczynowych tego typu. Zauważa, że *„zintegrowanie badania opartego na teorii programu z szacowaniem efektu przyczynowego interwencji przy pomocy fazy modelowania z użyciem kierunkowych grafów acyklicznych”*<sup>3</sup> będzie bardziej efektywnym sposobem realizacji badań ewaluacyjnych.

**Rafał Trzciński** skupia się na niedoskonałościach i ograniczeniach metod kontrfaktycznych w badaniach ewaluacyjnych i zaznacza konieczność stosowania jednocześnie ewaluacji opartej na teorii w praktyce badawczej, by w pełni wskazać ewentualne przyczyny lub porażki danej interwencji publicznej. Artykuł ten stanowi znakomity materiał do zaadaptowania w przyszłości tego sposobu prowadzenia badań ewaluacyjnych.

W kolejnym artykule **dr Monika Kwiecińska-Zdrenka** zauważa, iż niezbędne staje się poszukiwanie takich form ewaluacji, które pozwolą na zebranie bardziej satysfakcjonujących wyjaśnień źródeł zmiany oraz na ujawnienie efektów niezamierzonych działań. Przedstawia walory ewaluacji rozwojowej, opartej o zasady odwołujące się do teorii, która wykorzystując mieszane metody badawcze pozwoli zrozumieć zachodzące zmiany już w trakcie trwania tego procesu. Przedstawia również możliwości stosowania ewaluacji rozwojowej w instytucjach publicznych oraz wyznaczniki gotowości stosowania w nich tego rodzaju ewaluacji.

Tekst **dr Dariusza Majerka** przedstawia metodologiczne podstawy stosowania metody Propensity Score Matching w badaniach ewaluacyjnych. Prezentuje rozmaite algorytmy dopasowania prób w ramach tej metody, jej zalety i wady, jak również aplikacyjny charakter w zakresie badań ewaluacyjnych. Ilustruje konieczność wykorzystywania technik kontrfaktycznych

2 Górniak J., *Szacowanie przyczynowego wpływu interwencji: funkcje, walory, ograniczenia oraz relacje z innymi podejściami w ewaluacji*. (Artykuł z niniejszej publikacji).

3 Ibidem.

w badaniach ewaluacyjnych oraz ich rolę w szacowaniu wielkości efektów przyczynowo-skutkowych. Autor precyzyjnie przedstawia, jak efektywnie przeprowadzić ewaluację w oparciu o metodę Propensity Score Matching.

Kolejne teksty przedstawiają nowatorskie metody wykorzystywane do badań ewaluacyjnych, bazujące na socjologii wizualnej oraz różnorodnych zbiorach danych tzw. Big Data. **Dr Bogdan Jankowski i Magdalena Błaszczyk** wskazują na mnogość danych, które mogą zostać wykorzystane do badań ewaluacyjnych zwłaszcza do odkrywania nowych zjawisk i tworzenia korelacji. Przedstawiają źródła i narzędzia pozyskiwania takich danych, a także przykłady możliwości ich wykorzystania.

Z kolei tekst **Barbary Leszczyńskiej** mówi o możliwości wykorzystania nowatorskiej metody Design Thinking w pracy ewaluatora. Jej rozważania dotyczą praktycznej strony powstawania rekomendacji z przeprowadzonych ewaluacji. Z perspektywy ewaluatora tworzącego rekomendacje pokazuje, jak metoda Design Thinking umożliwi zrozumienie potrzeb, oczekiwań, ale też zwyczajów klienta/zamawiającego. Zaznacza, iż „skoncentrowanie na człowieku powinno być wartością szczególnie bliską administracji publicznej”<sup>4</sup>.

**Andrzej Reguński** w swoim artykule przedstawia refleksje związane z wykorzystaniem modelu partycypacyjnego w trakcie realizacji badań ewaluacyjnych prowadzonych w Polsce. Skupia się nad badaniami prowadzonymi dla perspektywy finansowej 2007-2013 i dokonuje pewnego rodzaju podziału badań, w których stosowano podejście partycypacyjne. Spostrzeżenia odnoszą się zarówno do wad, jak i zalet stosowania tego modelu w praktyce, jednak ich konkluzja wskazuje jasno na zalety, a nawet konieczność włączania ewaluatorów w prace na jak najwcześniejszym etapie programowania interwencji.

Kolejna część publikacji dotyczy rozważań nad praktycznymi aspektami ewaluacji. **Weronika Felcis** w swym artykule dzieli się cennymi refleksjami nad znaczeniem standardów ewaluacji w Polsce oraz nad tym jak powinien wyglądać system profesjonalizacji zbudowany na ich fundamencie. Artykuł przedstawia zarys sukcesów i porażek w procesie ewaluacyjnym w odniesieniu do standardów ewaluacji i zmusza do zastanowienia nad ich aktualnością i stosowaniem w pracy badawczej.

W artykule **Piotra Stronkowskiego** poruszone zostały kwestie dotyczące głównie doświadczeń z realizacji badań ewaluacyjnych w Polsce oraz wyzwań dla ewaluacji polityki spójności. Autor przedstawił szerokie spojrzenie na kwestię użyteczności badań ewaluacyjnych oraz główne problemy z wykorzystywaniem dostępnych badań i rekomendacji. Zaproponował możliwości rozwiązań, a także poprawy jakości badań.

---

4 Leszczyńska B., *W poszukiwaniu lepszych rekomendacji... Jak wykorzystać podejście Design Thinking w pracy ewaluatora.* (Artykuł z niniejszej publikacji).

**Barbara Porębska-Maciołek** i **Agnieszka Zawrzykraj** w swym artykule dzielą się swoimi doświadczeniami w dziedzinie ewaluacji regionalnego programu operacyjnego. Ich praktyka pracy pozwoliła wskazać szereg problemów z jakimi borykają się jednostki ewaluacyjne. Główne zastrzeżenia odnosiły się do wymaganych metod analizy danych, których znajomość na poziomie jednostki ewaluacyjnej nie jest wystarczająca i wymaga stałego dokształcania i rozwijania wiedzy, co niestety w przypadku uwarunkowań kadrowych i dostępnych zasobów nie zawsze jest możliwe. Ponadto wskazały na utrudnienia w dostępie do danych, które są niezbędne do realizacji badań, trudności na etapie tworzenia opisu zamówienia i wyboru wykonawcy, a także na problemy związane z niską jakością raportów końcowych.

Z kolei **Monika Lemke**, w swym artykule zobrazowała podejście do kryteriów wyboru wykonawcy, jakie dotychczas stosowano w zamawianych badaniach ewaluacyjnych, jak również wskazała na ich niedoskonałość i ograniczenia. Przedłożyła jednocześnie nowe rozwiązania, jakie wypracowała zachodniopomorska jednostka ewaluacyjna i doświadczenia w tym zakresie, wskazując na potrzebę stałego przeglądu kryteriów i modyfikacji ich do zmieniających się uwarunkowań.

Podsumowujący artykuł jest autorstwa **Jolanty Rudnickiej**, która relacjonuje zakres prac, jakie wykonuje jednostka ewaluacyjna, rolę jaką pełni w systemie programowania, zarządzania, a także wdrażania regionalnych programów operacyjnych województwa kujawsko-pomorskiego. Przedstawia również wymagania, które powinna spełnić każda komórka zajmująca się ewaluacją środków unijnych. Syntezą przytoczonych prac i zebranych doświadczeń w dziedzinie ewaluacji stały się wyzwania i zadania jakie stoją u progu realizacji nowej perspektywy finansowej 2014-2020.

W czasie konferencji zaprezentowano również referaty, które nie zostały uwiecznione w niniejszej publikacji<sup>5</sup>. O praktycznych aspektach badań ewaluacyjnych prowadzonych w Polsce wypowiedzieli się **Stanisław Bienias** z Fundacji IDEA Rozwoju, **Tomasz Kot** – Naczelnik Krajowej Jednostki Ewaluacji w Ministerstwie Rozwoju oraz **Przemysław Herman** – Naczelnik Wydziału Ewaluacji w Departamencie Europejskiego Funduszu Społecznego również z Ministerstwa Rozwoju. Kwestie przez nich poruszone dotyczyły głównie doświadczeń z realizacji badań w latach 2007-2013 oraz wyzwań dla polityki spójności z perspektywy Komisji Europejskiej i Polski. Przedstawione zostały dotychczasowe działania i osiągnięcia ewaluacyjne w Polsce, możliwości oraz plany badawcze na obecną perspektywę.

**Agnieszka Haber** z Polskiej Agencji Rozwoju Przedsiębiorczości, wypowiedziała się na temat możliwości wykorzystania podejścia *customer experience* w procesie ewaluacji. Wskazała, iż należy doprowadzić do zniwelowania różnic w postrzeganiu ewaluacji, które występują między ewaluatorami czy instytucjami zlecającymi z jednej a beneficjentami wsparcia z drugiej

<sup>5</sup> Prezentacje dostępne są na stronie <http://www.mojregion.eu/index.php/rpo/ewaluacja-szkolenia-konferencje>.

strony, co w konsekwencji powinno wpłynąć na bardziej świadome wykorzystywanie wyników badań ewaluacyjnych.

**Dr Małgorzata Wiśniewska** – Zastępca Dyrektora Departamentu Spraw Społecznych i Zdrowia w Urzędzie Marszałkowskim Województwa Kujawsko-Pomorskiego przedstawiła związki ewaluacji z obszarem ochrony zdrowia, wskazała na obligatoryjny charakter ewaluacji w tym zakresie. Przedstawiła płaszczyznę tych badań oraz możliwe utrudnienia w zakresie ich realizacji.

Publikacja z pewnością wprowadzi Państwa w zakres podejmowanych w Polsce działań ewaluacyjnych. Umożliwi również spojrzenie na aktualne problemy i wyzwania dla ewaluacji programów operacyjnych w perspektywie 2014-2020.

Jednostka Ewaluacyjna RPO  
Urząd Marszałkowski Województwa Kujawsko-Pomorskiego







PAWEŁ CHURSKI

## Wstęp

Procesy rozwoju społeczno-gospodarczego wykazują naturalną tendencję do polaryzacji w przestrzeni. Stanowi ona z jednej strony konsekwencję zróżnicowanych w przestrzeni zasobów i uwarunkowań rozwoju, a z drugiej strony oddziaływania czynników aglomeracji przestrzennej i koncentracji sektorowej oraz związanych z nimi efektów determinujących kształtowanie się trwałych przewag konkurencyjnych poszczególnych terytoriów. O ile przebieg spontanicznych procesów rozwojowych prowadzi zwykle do wzrostu polaryzacji i pogłębiania się różnic rozwojowych w przestrzeni<sup>1</sup>, to działania zmierzające do sterowania tymi procesami starają się ograniczać skalę tych różnic. W tej nieustannie towarzyszącej nam rywalizacji między procesami spontanicznymi a próbami ich sterowania i kontroli nie powinno chodzić o doprowadzenie do osiągnięcia mało realnego stanu równowagi, ale o dążenie do zmniejszenia poziomu różnic do stanu akceptowalnego społecznie, który nie będzie stanowił bariery rozwojowej<sup>2</sup>. Barierą taką jest występowanie zbyt dużych różnic w poziomie rozwoju między poszczególnymi obszarami. Zmniejszanie skali różnic rozwojowych jest zadaniem polityki spójności<sup>3</sup>. Podejmo-

- 1 Prawdopodobnie te znajdują swoje potwierdzenie w założeniach nurtu teorii ekonomii opartej na koncepcji ścieżki rozwoju nierównomiernego, rozwiniętej na gruncie krytyki paradygmatu równowagi gospodarczej i zrównoważonego wzrostu (Hansen N.M., *Criteria for growth center policy*, [w:] *Growth poles and growth centers in regional planning* (red.) A. Kukliński, Mouton, Paris 1972, s. 103-124; Todd D., *An appraisal of the development pole concept in regional analysis*, „Environment and Planning”, A, Vol. 6, no. 3/1974, s. 291-306). Obejmuje ono zarówno dorobek tradycyjnych koncepcji teoretycznych (m.in. Perroux F., *Note sur la nation de „pole de croissance”*, „Economie Appliquee”, vol. 8, no. 1-2/1955,, s. 307-320; Friedmann J., Alonso W., *Introduction*, [w:] *Regional development and planning*. (red.) J. Friedmann, W. Alonso, Cambridge, Mass: MIT Press, 1964, s. 1-13; Friedmann J., *A general theory of polarized development*, Santiago: Ford Foundation, Urban and Regional Development Advisory Program in Chile, Op. 81,1967; Friedmann J., *Ogólna teoria rozwoju spolaryzowanego*, [w:] *Przestrzeń krajów trzeciego świata – problemy metodologiczne*, (red.) M. Rościszewski, PZLG, IGIPIZ PAN, Warszawa 1974,), jak również ich współczesnych adaptacji oraz nowych podejść teoretyczno-metodologicznych (m.in. Romer P., *Increasing returns and long-run growth*, „Journal of Political Economy”, Vol. 94, no. 5/1986, s. 1002-1037; Romer P., *Endogenous technological change*, „Journal of Political Economy”, Vol. 98, no. 5, part II/1990, s. 71-102; Romer P., *The Origins of Endogenous Growth*, „The Journal of Economic Perspectives”, Vol. 8, no. (1)/1994; Lucas R.E., *On the mechanics of economic development*, „Journal of Monetary Economics”, vol. 22, no. 1/1988, s. 3-42; Krugman P., *Increasing returns and economic geography*, „Journal of Political Economy”, Vol. 99, no. 3/ 1991, s.483-499; Krugman P., *Development geography, and economic theory*, Ohlin Lectures, MIT Press, Cambridge-London 1995).
- 2 Faludi A., *From European Spatial Development to Territorial Cohesion Policy*, „Regional Studies”,Vol. 40,no. 6/2006, s. 667-678; Molle W., *European Cohesion Policy*, Routledge, London 2007; Gorzelak G., *Fakty i mity rozwoju regionalnego*, „Studia Regionalne i Lokalne”, nr 2 (36), Centrum Europejskich Studiów Regionalnych i Lokalnych Uniwersytetu Warszawskiego, Wydawnictwo Scholar, Warszawa 2009, s. 5-27.
- 3 *Growing unequal?: Income Distribution and Poverty in OECD Countries*, OECD Report Paris 2008; Barca F., *An Agenda for a Reformed Cohesion Policy*, A place-based approach to meeting European Union challenges and expectations, Independent Report, 2009; *Reshaping Economic Geography*, World Development Report 2009: The World Bank; *Regional Development Policies in OECD Countries*, OECD Report, Paris 2010.



wane w jej ramach działania mają na celu zwiększanie poziomu konwergencji, co w warunkach polityki spójności Unii Europejskiej odbywa się poprzez wzmacnianie spójności ekonomicznej, społecznej i terytorialnej<sup>4</sup>. Niestety relatywnie mała skuteczność dotychczasowych działań interwencyjnych zmusza do poszukiwań nowych podejść wyjaśniających procesy rozwoju na poziomie regionalnym i lokalnym. Skutkuje to modyfikacją tradycyjnych teorii wzrostu i rozwoju oraz włączeniem do mainstreamu badań ekonomicznych dorobku geografii ekonomicznej<sup>5</sup>. Odbywająca się w ten sposób swoista zmiana paradygmatu bazuje na wzroście znaczenia w wyjaśnianiu procesów rozwojowych kapitału ludzkiego i innowacji (*endogenous growth theory*), aglomeracji działalności oraz bliskości, która nie jest już tylko rozpatrywana w kategoriach bliskości fizycznej (*new economic geography*) oraz instytucji (*institutional economics*). Ponadto postępująca globalizacja zwraca uwagę na znaczenie lokalnych zasobów o charakterze materialnym i niematerialnym, które budują kapitał terytoriów warunkując ich przewagi konkurencyjne, uzasadniając upowszechnienie w praktyce polityki spójności podejście zorientowane terytorialnie (*place-based policy*)<sup>6</sup>. Zachodzące zmiany skutkują ewolucją sposobów postrzegania ukierunkowania interwencji w odniesieniu do wyraźnie zaznaczającej się w przestrzeni dychotomii rozwojowej, którą opisują kategorie obszarów polaryzacji rozwoju oraz peryferii rozwojowych. W przypadku tych pierwszych zmierza się z jednej strony do poprawy poziomu życia mieszkańców i warunków do prowadzenia działalności gospodarczej na obszarach o szczególnej aglomeracji ludności i podmiotów gospodarczych, a z drugiej strony dąży się do kształtowania siły i zasięgu powiązań funkcjonalnych tych obszarów z ich otoczeniem. W przypadku tych drugich z jednej strony zwraca się uwagę na tworzenie i wzmacnianie zasobów endogenicznych budujących potencjał własny tych obszarów poprawiający poziom życia ich mieszkańców i stopień wyposażenia infrastrukturalnego dla gospodarki, które przeciwdziałają odpływowi ludności oraz działalności gospodarczej, a z drugiej strony wzmacnia się czynniki sprzyjające przyjmowaniu przez te obszary pozytywnych oddziaływań z obszarów polaryzacji rozwoju. Prowadzi to w konsekwencji do przesunięcia szali podstawowego dylematu związanego z ukierunkowaniem polityki spójności z wyrównywania (ang. *equity*) na efektywność (ang. *efficiency*). W tym celu,

- 4 *Sixth Report on Economic, Social and Territorial Cohesion, Investment for jobs and growth: Promoting development and good governance in EU regions and cities*, Regional and Urban Policy, European Commission, July 2014, Brussels.
- 5 Krugman P., *Increasing returns...*, op. cit., s. 1; Krugman P., *Development geography...*, op. cit., s. 1; Barca. F., McCann P., Rodriguez-Pose A., *The case for regional development intervention: place-based versus place-neutral approaches*, „Journal of Regional Science”, s. 52, 1/2012, s. 134-152; Capello R., *Regional economics*, Routledge, New York 2014; Churski P., *Model polaryzacyjno-dyfuzyjny w przemianach polityki spójności – konsekwencje dla ukierunkowania polityki rozwoju*, [w:] (red.) P. Churski, „Biuletyn Instytutu Geografii Społeczno-Ekonomicznej i Gospodarki Przestrzennej UAM, seria Rozwój Regionalny i Polityka Regionalna”, 25/4. Bogucki Wydawnictwo Naukowe, Poznań 25/2014, s. 13-27.
- 6 Rodriguez-Pose A., Crescenzi R., *Research and Development, Spillovers, Innovation Systems, and the Genesis of Regional Growth in Europe*, „Regional Studies”, Vol. 42, no.1/2008, s. 51-67; Capello R., Nijkamp P., (Eds.), *Handbook of Regional Growth and Development Theories*, Edward Elgar, 2009; Zaucha J., Komornicki T., *The Concept of Territorial Cohesion as Understood and Practiced by Polish Regions*, Working Papers 2/2015, Institute for Development, Sopot 2015; Zaucha i inni, *Terytorialny wymiar wzrostu i rozwoju*, Wydawnictwo DIFIN, Warszawa 2015,.

przyjmując triadę: spójność-konkurencyjność-sprawność, zmierza się do reorientacji podstaw teoretycznych, podejścia rozwojowego, strategii interwencji, jej koncentracji merytorycznej i przestrzennej, wykorzystywanych narzędzi oraz głównych podmiotów ją realizujących<sup>7</sup>.

Celem analizy jest identyfikacja wielkości i charakteru zróżnicowań rozwojowych występujących w polskiej przestrzeni oraz prawidłowości procesu absorpcji środków europejskich przeznaczanych zgodnie z założeniami polityki spójności na ograniczanie skali tych różnic. Przedmiotem analizy są wybrane wyniki badań realizowanych z tego zakresu przez Zespół Zakładu Analizy Regionalnej Instytutu Geografii Społeczno-Ekonomicznej i Gospodarki Przestrzennej Uniwersytetu im. Adama Mickiewicza w Poznaniu, które wpisują się w szeroki nurt podobnych analiz realizowanych w Polsce<sup>8</sup>. Dotyczą one zróżnicowań rozwojowych na poziomie lokalnym, poziomu i struktury wykorzystania środków europejskich w obszarach polaryzacji rozwoju i peryferii rozwojowych oraz obserwowanej szczególnej pozycji i specyfiki ośrodków regionalnych w procesie absorpcji środków polityki spójności, jak również opinii przedstawicieli samorządu terytorialnego na temat tego procesu. Prezentowane analizy odnoszą się do stanów z okresu 2000-2015 i prowadzone są na poziomie powiatów (LAU1). Stanowią one rezultat badań realizowanych w ramach projektu *Rozwój społeczno-gospodarczy a kształtowanie się obszarów wzrostu i obszarów stagnacji gospodarczej*, który był finansowany przez Narodowe Centrum Nauki (N N306 791940), projektu własnego *Rola środków europejskich dla rozwoju miast Polski – 10 lat członkostwa w Unii Europejskiej* oraz indywidualnych analiz prowadzonych przez Zespół Pracowników Zakładu Analizy Regionalnej Instytutu Geografii

7 *Sixth Report...*, op. cit., s. 2.

8 Gawlikowska-Hueckel K., *Procesy rozwoju regionalnego w Unii Europejskiej, Konwergencja czy polaryzacja?* Wydawnictwo Uniwersytetu Gdańskiego, Gdańsk 2003; Nowińska-Łażniewska E., *Relacje przestrzenne w Polsce w okresie transformacji w świetle teorii rozwoju regionalnego*, Prace habilitacyjne nr 13, Wydawnictwo Akademii Ekonomicznej w Poznaniu, Poznań 2004; Szlachta J., *Polska ścieżka rozwoju regionalnego w poszerzonej Unii Europejskiej*, [w:] *Region Bałtycki w nowej Europie* (red.) T. Parteka, J. Szlachta, W. Szydarowski, „Biuletyn KPZK PAN”, z. 217, Warszawa 2005, s. 7-42; Wojnicka E., Tarkowski M., Klimczak P., *Przestrzenne i regionalne zróżnicowanie ośrodków wzrostu. Polaryzacja a wyrównywanie szans rozwojowych. Przesłanki dla kształtowania polityki regionalnej państwa*, Ekspertyza dla Ministerstwa Gospodarki i Pracy, styczeń 2005, Gdynia – Rzeszów; Pietrzyk I., *Regionalne implikacje europejskiej integracji gospodarczej dla Polski*, „Zeszyty Naukowe”, nr 721, Akademia Ekonomiczna w Krakowie, Kraków 2006; Michałek J. J., (red.) *Polska w Unii Europejskiej, Dynamika konwergencji ekonomicznej*, Wydawnictwo Naukowe PWN, Warszawa 2007; Tarkowski M., *Centra i peryferie rozwoju społeczno – gospodarczego Polski w okresie transformacji ustrojowej*, Wydawnictwo Bernardinum, Gdynia-Pelplin 2007; Tarajkowski J., Wojtasiewicz L. (red.), *Przestrzeń w polityce gospodarczej*, Katedra Polityki Gospodarczej i Planowania Rozwoju, Akademia Ekonomiczna w Poznaniu, Poznań 2008; Wójcik P., *Dywersjencja czy konwergencja: dynamika rozwoju polskich regionów*, „Studia Regionalne i Lokalne”, (2) 32, Warszawa 2008, s. 41-60; Gorzelak G., *Fakty i mity...*, op. cit.1; Jastrzębska M., *Znaczenie bezwrotnych środków zagranicznych jako źródeł finansowania działalności jednostek samorządu terytorialnego w latach 2004-2010*, „Finanse Komunalne”, nr 10, 2011, s.18-31; Swianiewicz P., *Środki unijne w samorządach – kto korzysta najwięcej?*, „Samorząd Terytorialny”, nr 5, 2012, s. 9-24; Misiąg J., Misiąg W., Tomalak M., 2013, *Ocena efektywności wykorzystania pomocy finansowej Unii Europejskiej jako instrumentu polityki spójności społeczno-gospodarczej oraz poprawy warunków życia*, Rzeszów; Smętkowski M., *Rozwój regionów i polityka regionalna w krajach Europy Środkowo-Wschodniej w okresie transformacji i globalizacji*, Wydawnictwo Naukowe Scholar, Warszawa 2013; Gorzelak G., 2014, *Wykorzystanie środków Unii Europejskiej dla rozwoju kraju – wstępne analizy*, „Studia Regionalne i Lokalne”, nr 3 (57), s. 5-25; Wojarska M., Zabielska M., *Samorząd terytorialny jako beneficjent funduszy unijnych (na przykładzie gmin województwa warmińsko-mazurskiego)*, „Prace Naukowe Uniwersytetu Ekonomicznego we Wrocławiu”, Wrocław 2015, nr 380, s. 266-275.

Społeczno-Ekonomicznej i Gospodarki Przestrzennej UAM w Poznaniu. Ich podstawą są zarówno publikowane i niepublikowane dane statystyczne Głównego Urzędu Statystycznego i Ministerstwa Rozwoju, jak również informacje pozyskane w drodze badań bezpośrednich techniką ankiety.

### **Obszary polaryzacji rozwoju i peryferii rozwojowych w polskiej przestrzeni społeczno-ekonomicznej<sup>9</sup>**

Podstawą identyfikacji rozkładu obszarów polaryzacji rozwoju i peryferii rozwojowych w układzie powiatów w Polsce w latach 2000-2010 była analiza stanu i zmian procesu rozwoju społeczno-gospodarczego, która została przeprowadzona w ujęciu częściowym obejmującym pięć aspektów: (1)*ludność i osadnictwo*, (2)*rynek pracy i strukturę gospodarki*, (3)*infrastruktura techniczna i dostępność przestrzenna*, (4)*sytuacja finansowa i poziom zamożności*, (5)*innowacyjna gospodarka i otoczenie biznesu*, oraz w ujęciu całościowym traktującym powyższe aspekty łącznie. W postępowaniu badawczym wykorzystano metodę wskaźnika syntetycznego J. Perkala (z-score index) oraz analizę skupień metodą k-średnich<sup>10</sup>. Liczba cech źródłowych opisujących analizowane aspekty procesu rozwoju społeczno-gospodarczego wynosiła 49 i wahała się od trzech (*innowacyjna gospodarka i otoczenie biznesu*) do 16 (*sytuacja finansowa i poziom zamożności*). Przeprowadzona procedura doprowadziła do klasyfikacji powiatów, ze względu na ich stan poziomu rozwoju społeczno-gospodarczego<sup>11</sup>. Analiza zmienności uzyskanych rozkładów przestrzennych typów rozwojowych powiatów w ujęciu całościowym pozwoliła dokonać ich klasyfikacji według kryterium zmienności przynależności do danej klasy. W tym celu przeprowadzono klasyfikację jednostek bazującą na wzorcu przedstawionym na rycinie 1. Podstawowym kryterium tej klasyfikacji była długość okresu pozostawania danej jednostki w jednej z trzech klas: obszary polaryzacji rozwoju – obszary przejściowe – obszary peryferii rozwoju, w analizowanym okresie lat 2000-2010 w ujęciu całościowym procesu rozwoju społeczno-gospodarczego. Założono, że grupę obszarów polary-

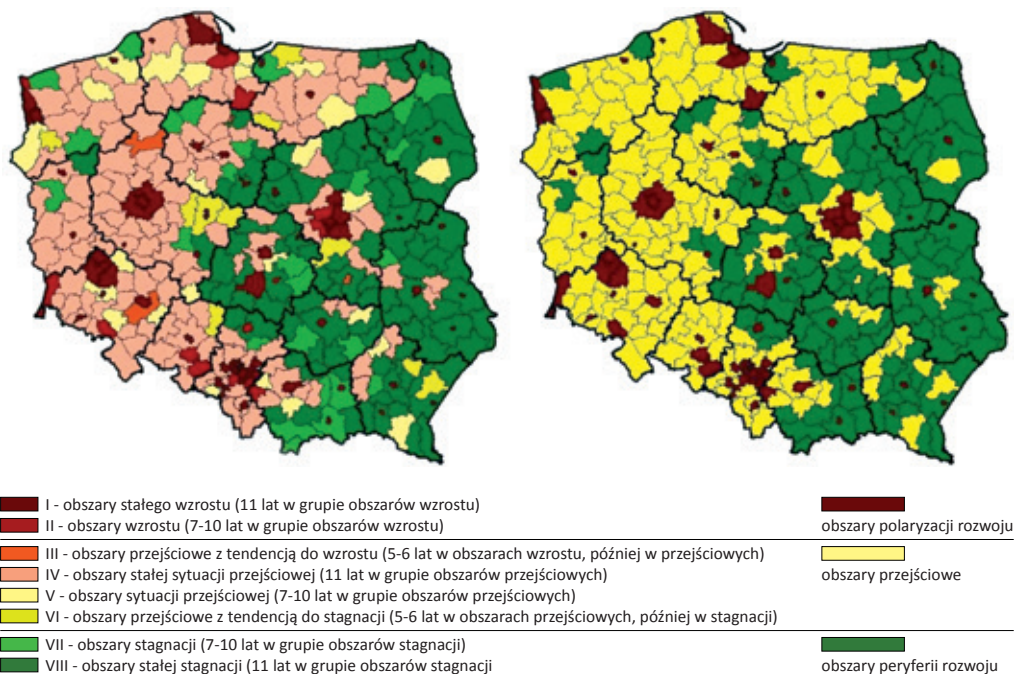
9 Patrz: Churski P., 2014b (red.), *Rozwój społeczno-gospodarczy a kształtowanie się obszarów wzrostu i obszarów stagnacji gospodarczej w Polsce*. Biuletyn Instytutu Geografii Społeczno-Ekonomicznej i Gospodarki Przestrzennej UAM, seria Rozwój Regionalny i Polityka Regionalna, 25/2014. Bogucki Wydawnictwo Naukowe, Poznań; Churski P., 2014c (red.), *The social and economic growth vs. the emergence of economic growth and stagnation areas*. Bogucki Wydawnictwo Naukowe, Poznań. Całościowe wyniki projektu *Rozwój społeczno-gospodarczy a kształtowanie się obszarów wzrostu i obszarów stagnacji gospodarczej*, który był finansowany przez Narodowe Centrum Nauki (N N306 791940), są udostępniane w serwisie internetowym <http://www.owsg.pl>.

10 Tryon R. C., *Cluster Analysis*, Edwards Bros, Oxford 1939; Smith D., *Geography and social indicators*. „South African Geographical Journal”, no. 54/1972 s. 43-57; Morrison D.F., *Wielowymiarowa analiza statystyczna*, PWN, Warszawa 1990; Kronhalter F., *A Study of the Competitiveness of Regional Based on a Cluster Analysis. The Example of East Germany*, Halle Institute for Economic Research, Halle/Saale 2003; Dymnicki A.B., Henry D.B., *Use of Clustering Methods to Understand More about the Case*, „Methodological Innovations Online”, Vol. 6 no. (2)/2011, s. 6-26. <http://www.pbs.plym.ac.uk/mi/pdf/31-08-11/2.%20Dymnicki%20-%20pp6-26.pdf>.

11 Dolata M., Borowczak A. *Distribution of economic growth and stagnation areas in Poland, 2000-2010*, [w:] *The social and economic growth vs. the emergence of economic growth and stagnation areas*, (red.) P., 2014. Churski, Bogucki Wydawnictwo Naukowe, Poznań, s. 29-42. Całościowe wyniki projektu są udostępniane w serwisie internetowym <http://www.owsg.pl>.

zacji rozwoju stanowią powiaty zaliczane do klasy I – stałego wzrostu (11 lat w grupie obszarów wzrostu) oraz do klasy II – obszary wzrostu (7-10 lat w grupie obszarów wzrostu), natomiast grupę obszarów peryferii rozwojowych stanowią powiaty zaliczane do klasy VIII – stałej stagnacji (11 lat w grupie obszarów stagnacji) oraz do klasy VII – obszary stagnacji (7-10 lat w grupie obszarów stagnacji) (ryc. 1). Obszary polaryzacji rozwoju stanowią miejsca o relatywnie najwyższym poziomie rozwoju. Charakteryzują się one pozytywnymi tendencjami demograficznymi i wysoką jakością kapitału ludzkiego, wzmocnioną jeszcze możliwością korzystania z usług edukacyjnych najwyższego poziomu. Rynek pracy tych obszarów cechuje duża różnorodność miejsc pracy i wysoki poziom przedsiębiorczości skutkujący najwyższymi wskaźnikami samozatrudnienia. Uzyskiwane przychody zapewniają dobrą sytuację finansową podmiotom gospodarczym oraz pozwalają na zapewnienie wysokiego poziomu i warunków życia mieszkańcom. Pod pojęciem obszarów peryferii rozwoju rozumie się natomiast powiaty o relatywnie najniższym poziomie rozwoju, który potwierdzony jest niskimi wartościami wskaźników rozwojowych. Wykazują one negatywne tendencje demograficzne przy jednocześnie relatywnie niskim poziomie kapitału ludzkiego, który bardzo często podlega „drenażowi” przez obszary polaryzacji rozwoju.

**Rycina 1. Rozkład przestrzenny klas obszarów polaryzacji rozwoju oraz peryferii rozwojowych w ujęciu całościowym w układzie powiatów**



Źródło: Opracowanie własne.

Rynek pracy tych obszarów bardzo często wykazuje cechy rynku monofunkcyjnego i charakteryzuje się wysokim poziomem bezrobocia. Ze względu na relatywnie niską jakość kapitału ludzkiego oraz powszechnie występujące braki infrastrukturalne gospodarka obszarów peryferii rozwojowych cechuje się niską zdolnością do generowania innowacji oraz niskim poziomem elastyczności, co znajduje potwierdzenie w jej strukturze. Sytuacja ta prowadzi do przegrywania rywalizacji konkurencyjnej i stanowi przyczynę złej sytuacji finansowej podmiotów gospodarczych oraz niskiego poziomu i warunków życia mieszkańców.

Zbiór powiatów należących do tak zdefiniowanej klasy obszarów polaryzacji rozwoju obejmuje 83 powiaty, co stanowi 22% ogólnej liczby jednostek. W strukturze tego zbioru znajduje się 69 powiatów (18,2%) zaklasyfikowanych w ujęciu całościowym rozwoju społeczno-gospodarczego jako obszary stałego wzrostu (klasa I) oraz 14 powiatów (3,7%), które co najmniej siedmiokrotnie klasyfikowano jako obszary wzrostu (klasa II) (ryc. 1). W układzie funkcjonalno-przestrzennym wyróżnione obszary obejmują: 18 powiatów grodzkich – ośrodków miejskich o randze krajowej; 7 powiatów ziemskich, stanowiących obszary metropolitalne tych ośrodków; 44 powiaty grodzkie lub ziemskie, w granicach których zlokalizowane są ośrodki miejskie o znaczeniu regionalnym lub subregionalnym; 14 powiatów grodzkich i ziemskich, w granicach których zlokalizowane są znaczące zakłady przemysłowe, których monofunkcyjna gospodarka w wielu przypadkach związana jest z przemysłem wydobywczym bazującym na lokalnych zasobach naturalnych. Uzyskany rozkład przestrzenny obszarów polaryzacji rozwoju oraz ich struktura funkcjonalno-przestrzenna pozwalają sądzić, że mimo względnie najwyższego poziomu rozwoju społeczno-gospodarczego w ujęciu całościowym, nie stanowią one jednorodnej grupy<sup>12</sup>. Obszary polaryzacji rozwoju występują w każdym regionie, a ich rozkład determinowany jest występowaniem powiatów o wskazanych powyżej funkcjach. Tym samym w regionach Polski Wschodniej, pozbawionych powiatów surowcowych obszary polaryzacji rozwoju ograniczają się do powiatów ze stolicami w ośrodkach regionalnych i subregionalnych, które ze względu na swoją relatywną słabość gospodarczą nie oddziałują na swoje bezpośrednie otoczenie i nie tworzą ich większych skupisk przestrzennych. Inaczej sytuacja przedstawia się w pozostałej części kraju. W tym przypadku obszary polaryzacji rozwoju rozwijające się wokół największych miast Polski: Warszawy, Łodzi, Krakowa, Wrocławia, Poznania, Gdańska, itd. obejmują zarówno powiaty grodzkie największych metropolii, jak również otaczające je powiaty aglomeracyjne. Dodatkowo rozkład ten uzupełniają obszary polaryzacji rozwoju kształtujące się na bazie powiatów przemysłowych i surowcowych, które w większości przypadków, poza obszarem Konurbacji Górnośląskiej oraz Zagłębia Lubiąsko-Głogowskiego występują pojedynczo nie tworząc

12 Perdał R., Hauke J., *Areas of economic growth and stagnation in Poland – growth factors*. [w:] *The social and economic growth vs. the emergence of economic growth and stagnation areas*, (red.) P. Churski, Bogucki Wydawnictwo Naukowe, Poznań 2014, s. 43-62. Całościowe wyniki projektu są udostępniane w serwisie internetowym <http://www.owsq.pl>.



skupisk przestrzennych, np.: powiat zgorzelecki – Kopalnia w Bogatyni i Elektrownia Turów, powiat policki – Zakłady Chemiczne POLICE S.A – Grupa Azoty, powiat kwidziński – International Paper Kwidzyn, powiat kozienicki – ENEA Wytwarzanie S.A., itd.

Zbiór powiatów należących do klasy obszarów peryferii rozwojowych obejmuje 126 powiatów, co stanowi 33,2% ogólnej liczby jednostek, a jednocześnie miejsce życia 24% populacji Polski. W strukturze tego zbioru znajduje się 99 powiatów (26,1%) zaklasyfikowanych w ujęciu całościowym rozwoju społeczno-gospodarczego jako obszary stałej stagnacji (Klasa VIII) oraz 27 powiatów (7,1%), które co najmniej siedmiokrotnie sklasyfikowano jako obszary stagnacji (Klasa VII) (ryc. 1). Uzyskany rozkład przestrzenny obszarów peryferii rozwojowych oraz ich struktura funkcjonalno-przestrzenna pozwalają sądzić, że mimo względnie najniższego poziomu rozwoju społeczno-gospodarczego w ujęciu całościowym, podobnie jak w przypadku obszarów polaryzacji rozwoju nie stanowią one jednorodnej grupy<sup>13</sup>. Znaczna koncentracja obszarów peryferii rozwojowych występuje głównie we wschodniej i środkowej części kraju, nawiązując swoim zasięgiem w dużym stopniu do granic reliktowych z okresu po III rozbiórce Polski w 1795r., które trwale podzieliły przestrzeń społeczno-gospodarczą naszego kraju. Obok tej generalnej prawidłowości należy zwrócić uwagę na wewnętrzne zróżnicowania obserwowane w układzie poszczególnych regionów. W województwach podlaskim, lubelskim i świętokrzyskim obszary peryferii rozwojowych zajmują powierzchnię niemal całych województw (bez ośrodków wojewódzkich i subregionalnych), w województwach mazowieckim, kujawsko-pomorskim, łódzkim i małopolskim – choć relatywnie liczne – obszary peryferii rozwojowych rozmieszczone są daleko od gospodarczego rdzenia regionów, najczęściej wzdłuż ich granic. Podobnie „peryferyjna” (w sensie przestrzennym) jest lokalizacja nielicznych obszarów peryferii rozwojowych w województwach zachodnich: zachodniopomorskim, lubuskim, śląskim i wielkopolskim. W trzech województwach obszary peryferii rozwojowych nie występują wcale. Są to województwa pomorskie, dolnośląskie i opolskie, których powiaty cechują się co najmniej przeciętną sytuacją społeczno-gospodarczą na tle innych jednostek w kraju. Powiaty zaliczone do obszarów peryferii rozwojowych charakteryzują się z reguły przewagą funkcji rolniczych, choć są zróżnicowane pod względem struktury funkcjonalnej. W szczególnie niekorzystnej sytuacji znajdują się obszary o przewadze tzw. rolnictwa socjalnego o przeważnie peryferyjnej lokalizacji w układach wewnątrzregionalnych. Sporadycznie w kraju jako obszary peryferii rozwojowych zaklasyfikowano obszary cechujące się dominacją funkcji leśnej (np. powiaty sulęciński, augustowski, hajnowski), a także obszary o szczególnych walorach środowiska przyrodniczego, wykazujące przewagę funkcji turystycznych i rekreacyjnych (np. tatrzański, bieszczadzki, suwalski sejneński). W relatywnie najlepszej sytuacji spośród wyróżnionych obszarów peryferii rozwojowych znajdują się powiaty, które pod względem typologii funkcjonalnej można zaliczyć do obszarów urbanizowanych lub wielofunkcyjnych obszarów przejściowych, stanowiących odpo-

13 Perdał R., Hauke J., *Areas of economic...*, op. cit., s. 7.

wiednio zewnętrzne i wewnętrzne części stref podmiejskich o mieszanych funkcjach i znacznym udziałem funkcji mieszkaniowej. Obszary te zlokalizowane są przeważnie w Polsce Wschodniej, w relatywnej bliskości do regionalnych centrów rozwoju lub większych ośrodków miejskich (np. rzeszowski, łąncucki, lubelski, wrocławski, płocki, częstochowski). Należy jednak podkreślić, że pomimo ich bezpośredniego sąsiedztwa z obszarami polaryzacji rozwoju, w tym byłymi i obecnymi miastami wojewódzkimi, pozostają one w klasie obszarów peryferii rozwojowych.

### **Poziom i struktura wykorzystania funduszy UE na obszarach polaryzacji rozwoju i peryferii rozwojowych w ujęciu całościowym i częściowym<sup>14</sup>**

Analiza wielkości i struktury inwestycji finansowanych ze środków unijnych w kontekście ich zgodności z potrzebami obszarów polaryzacji rozwoju i peryferii rozwojowych w Polsce wynikającymi z tworzenia i wzmacniania czynników rozwoju składała się z następujących kroków: przeprowadzenie kwerendy projektów realizowanych w okresie objętym analizą tj. latach 2004-2010, ustalenie rozkładu przestrzennego projektów, odpowiadającego przyjętemu układowi analizy tj. poziomowi LAU1, ustalenie struktury inwestycji w podziale na analizowanych pięć badanych aspektów procesu rozwoju społeczno-gospodarczego, oraz wyodrębnienie i systematyzacja inwestycji realizowanych w wyróżnionych obszarach polaryzacji rozwoju tj. 83 powiatach oraz obszarach peryferii rozwojowych tj. 126 powiatach. U podstaw kwerendy projektów realizowanych w ramach dwóch perspektyw finansowych na lata 2004-2006 oraz 2007-2013 leżały następujące założenia. W zbiorze projektów podlegających analizie nie uwzględniono inwestycji realizowanych przy wsparciu Funduszu Spójności, który nie jest funduszem strukturalnym. W analizie pominięto inwestycje realizowane w latach 2004-2006 ze środków programów SPO Restrukturyzacja i modernizacja sektora żywnościowego i rozwój obszarów wiejskich i SPO Rybołówstwo i przetwórstwo ryb (o łącznym udziale ok. 10% w NPR na lata 2004-2006). Rezygnacja z analizy inwestycji przeprowadzonych w ramach tych programów pozwoliła zachować homogeniczny układ finansowania polityki regionalnej w obu perspektywach finansowych. Ze zbioru projektów podlegających analizie wykluczono także inwestycje realizowane ze środków pomocy technicznej (zarówno w kontekście indywidualnych programów operacyjnych, jak i priorytetów poszczególnych programów). Przyjęto, że równoważą one koszty związane z obciążeniem administracji rządowej i regionalnej związane z obsługą programów unijnych.

14 Patrz: Churski P., Borowczak A., Perdał R., *Struktura inwestycji finansowanych ze środków unijnych a czynniki rozwoju w obszarach wzrostu w Polsce*. [w:] (red.) E. Małuszyńska, G. Mazur, I. Musiałkowska. Polska 10 lat członkostwa w Unii Europejskiej. Wydawnictwo Uniwersytetu Ekonomicznego w Poznaniu. Poznań 2015, s. 183-199; Churski P., Borowczak A., Perdał R., Czynniki rozwoju obszarów stagnacji w Polsce a ukierunkowanie interwencji środków unijnych. *Studia Obszarów Wiejskich*. Tom 37. Wiejskie obszary funkcjonalne. Komisja Obszarów Wiejskich PTG. Instytut Geografii i Przestrzennego Zagospodarowania PAN. Warszawa 2015, s. 115-130. Całościowe wyniki projektu *Rozwój społeczno-gospodarczy a kształtowanie się obszarów wzrostu i obszarów stagnacji gospodarczej*, który był finansowany przez Narodowe Centrum Nauki (N N306 791940), są udostępniane w serwisie internetowym <http://www.owsg.pl>.

W konsekwencji przyjętych założeń analizie podlegał zbiór 88,5 tys. projektów o łącznej wartości 21,9 mld € (tab. 1). W ogólnej liczbie projektów realizowanych do końca 2010 r. wyraźnie zaznacza się przewaga inwestycji realizowanych w perspektywie 2007-2013, jednakże pod względem ich wartości, udział projektów tej perspektywy wynosi jedynie ok. 30%. Przyczyną tego stanu był relatywnie niski stopień zaawansowania realizacji NSRO 2007-2013 pod koniec 2010 r.

**Tabela 1. Uwzględniony w analizie zakres interwencji strukturalnych polityki spójności w Polsce w układzie programów operacyjnych w latach 2004-2010<sup>15</sup>**

Programy operacyjne	Wartość ogółem (w €)	Liczba projektów (w szt.)
Infrastruktura i Środowisko	245 495 838	282
Innowacyjna Gospodarka	772 452 423	3 512
Kapitał Ludzki	1 086 176 808	28 903
Rozwój Polski Wschodniej	41 016 212	10
Regionalne Programy Operacyjne	4 741 416 465	11 149
PERSPEKTYWA FINANSOWA 2007-2013 <sup>15</sup>	6 886 557 746	48 856
Rozwój Zasobów Ludzkich	2 032 456 948	8 997
Transport	2 447 962 111	1 872
Wzrost Konkurencyjności Przedsiębiorstw	4 993 264 024	17 930
Zintegrowany Program Operacyjny Rozwoju Regionalnego	5 553 316 101	15 927
PERSPEKTYWA FINANSOWA 2004-2006	15 026 999 183	44 726
<b>RAZEM</b>	<b>21 913 556 929</b>	<b>88 582</b>

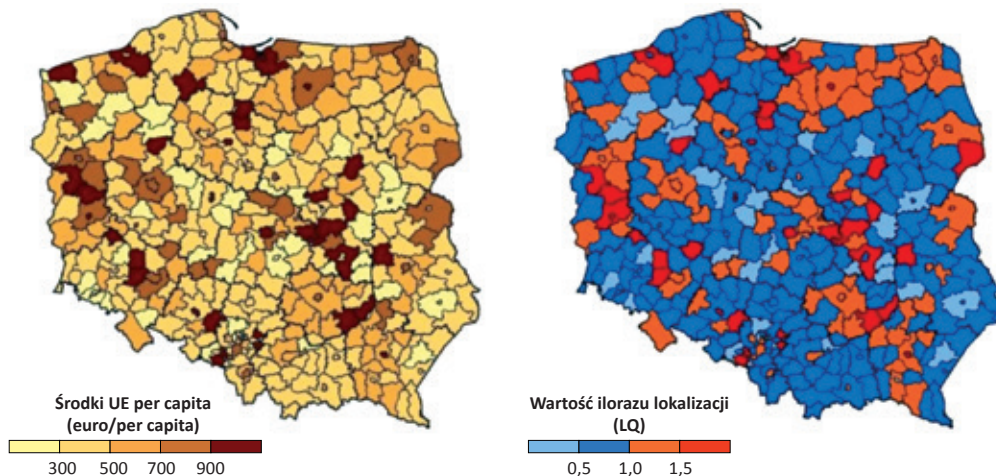
Źródło: Obliczenia własne na podstawie danych Ministerstwa Rozwoju Regionalnego, stan na 31.12.2010 r.

Identyfikacja przestrzennego podziału środków w układzie jednostek LAU1 wymagała ustalenia lokalizacji i zasięgu każdej z uwzględnionych w badaniu inwestycji. Przyjęto, że w przypadku, gdy inwestycję realizowano na szczeblu centralnym lub regionalnym podział środków w układzie odpowiednich powiatów (tj. krajowym lub regionalnym) powinien odpowiadać proporcji wynikającej z ich liczby mieszkańców. W ten sposób uzyskano rozkład wartości środków europejskich pozyskanych w ramach podjętej interwencji analizowanych programów operacyjnych, którego zróżnicowanie *per capita* przedstawia rycina 2. Dodatkowo można zauważyć, że jedynie w 35% przypadków występuje „nadmiar” pozyskanych środków w stosunku do liczby ludności ( $LQ > 1$ ). Względna nadwyżką w tym zakresie cechuje się 54% powiatów tworzących obszary polaryzacji rozwoju oraz zaledwie 31% powiatów tworzących peryferia rozwoju.

<sup>15</sup> Dla porównania poziom kontraktacji środków w programach operacyjnych NSRO 2007-2013 wynosił w 06.2011 r.: POiŚ 14%, POIG 19%, POKL 33%, RPW 21%, RPO 33% (*Sprawozdanie z realizacji...2011*).



**Rycina 2. Rozkład przestrzenny skumulowanej wartości pozyskanych środków UE *per capita* oraz wartości ilorazu lokalizacji (LQ) w układzie powiatów w latach 2004-2010**



Źródło: Obliczenia własne na podstawie danych Ministerstwa Rozwoju Regionalnego, stan na 31.12.2010 r.

Ustalenie tematycznej struktury nakładów w podziale na uwzględnione w niniejszej analizie aspekty rozwoju społeczno-gospodarczego wymagało szczegółowego przeglądu dokumentów programowych. Podziału inwestycji dokonano na poziomie działań lub poddziałań poszczególnych programów operacyjnych. Ogólne zasady tej systematyzacji można określić następująco. Do aspektu *ludność i osadnictwo* zakwalifikowano inwestycje dotyczące infrastruktury społecznej, poprawy warunków życia na obszarach zagrożonych marginalizacją (obszary wiejskie, obszary restrukturyzacji przemysłu, działania związane z rewitalizacją), inwestycje wspierające dziedzictwo kulturowe. Do aspektu *rynek pracy i struktura gospodarki* zakwalifikowano działania służące zmianie struktury zatrudnienia (np. reorientacja zawodowa osób odchodzących z rolnictwa), zwalczaniu bezrobocia w różnych grupach społecznych oraz w kontekście systemowym, adaptacja zasobów pracy do wyzwań nowoczesnej gospodarki, a także inwestycje w konkretne branże gospodarki (np. turystykę). Do aspektu *infrastruktura techniczna i dostępność przestrzenna* zakwalifikowano inwestycje związane z modernizacją i rozbudową systemów transportowych oraz infrastruktury ochrony środowiska. Do aspektu *sytuacja finansowa i poziom zamożności* zakwalifikowano inwestycje służące poprawie dostępu do kapitału dla przedsiębiorstw. Do aspektu *innowacyjna gospodarka i otoczenie biznesu* zakwalifikowano działania o charakterze proinnowacyjnym: wsparcie instytucjonalne otoczenia biznesu, tworzenie regionalnych strategii innowacyjnych, rozwój specjalistycznych usług o wysokim nasyceniu wiedzą (np. e-usługi), transfer technologii oraz działania związane z tworzeniem społeczeństwa informacyjnego. W rezultacie zastosowania przedstawionych zasad struktura wydatków polityki spójności w skali kraju, w badanym okresie, w układzie analizowanych aspektów rozwoju

społeczno-gospodarczego przedstawiała się następująco: 3,7 mld €, tj. 17% środków przeznaczono na *ludność i osadnictwo*; 2,2 mld €, tj. 10% środków na *wsparcie rynku pracy i struktury gospodarki*; 6,7 mld €, tj. 30% środków na *infrastrukturę techniczną i dostępność przestrzenną*; 7,0 mld €, tj. 32% środków na *poprawę sytuacji finansowej i poziomu zamożności*; natomiast 2,3 mld €, czyli 11% środków na *innowacyjną gospodarkę i otoczenie biznesu*. Uzyskane wyniki wskazują, że największa koncentracja nakładów środków strukturalnych udostępnionych w ramach interwencji programów operacyjnych krajowej polityki spójności w latach 2004-2010 dotyczyła na poziomie powiatowym przedsięwzięć kształtujących i wzmacniających czynniki rozwoju związane z *sytuacją finansową i poziomem zamożności oraz infrastrukturą techniczną i dostępnością przestrzenną*.

Na tym tle 83 wydzielone obszary polaryzacji rozwoju, stanowiące 22% wszystkich powiatów w Polsce, zrealizowały ok. 45% ogółu inwestycji, których wartość wyniosła ok. 9,9 mld €. Pomimo faktu, że powiaty te stanowią miejsca znacznej koncentracji ludności, charakteryzują się one wysokimi, przekraczającymi średnie krajowe, wartościami wskaźników *per capita*. Poziom inwestycji na obszarach polaryzacji rozwoju był wyższy o 19% niż przeciętny w kraju, co pozwala sądzić, że faktyczna aktywność beneficjentów polityki spójności, realizowanej zgodnie z podejściem wyrównawczym, odbiega od jej zakładanej koncentracji na obszarach słabszych ekonomicznie. Struktura środków pozyskanych przez obszary polaryzacji rozwoju jest podobna do struktury krajowej z dominującą jednak pozycją nakładów na *infrastrukturę techniczną i dostępność przestrzenną*, na którą na tych obszarach przeznaczono aż 3,3 mld €, tj. 49% całkowitej puli środków wydatkowanych na ten cel w kraju (tab. 2). Zajmują one także pierwsze miejsce w strukturze inwestycji zrealizowanych wyłącznie na obszarach polaryzacji rozwoju, stanowiąc 33% wartości zainwestowanych tam środków. Istotnymi kierunkami interwencji programów operacyjnych w badanym okresie na obszarach polaryzacji rozwoju w Polsce okazały się także działania wspierające *sytuację finansową i poziom zamożności* (tj. *de facto* dotacje dla przedsiębiorstw) oraz inwestycje w *innowacyjną gospodarkę i otoczenie biznesu*, które można powiązać z wyższą od przeciętnej koncentracją tego rodzaju działalności gospodarczej oraz lepszą kondycją finansową podmiotów gospodarczych na tych obszarach. Nie bez znaczenia jest także prawdopodobnie większa motywacja podmiotów zlokalizowanych w granicach obszarów polaryzacji rozwoju do uruchamiania procesów inwestycyjnych i aplikowania o środki europejskie w procedurach konkursowych, która z jednej strony stanowi efekt większego poziomu konkurencji, a z drugiej strony może stanowić rezultat efektów aglomeracji. W ujęciu *per capita* w strukturze inwestycji zrealizowanych na obszarach polaryzacji rozwoju pierwsze miejsce zajmują inwestycje prowadzone w aspekcie *innowacyjna gospodarka i otoczenie biznesu*. Biorąc jednak pod uwagę fakt, że jest to aspekt identyfikowany jedynie przy pomocy trzech dostępnych w statystyce publicznej wskaźników uzyskany wynik należy interpretować z dużą ostrożnością, zwracając uwagę na ko-

Tabela 2. Wielkość i struktura interwencji polityki spójności na obszarach polaryzacji rozwoju i peryferii rozwojowych w Polsce w latach 2004-2010

Aspekty	Wielkość środków UE ogółem				Wielkość środków UE per capita					
	Polska - mld €	Obszary polaryzacji rozwoju		Obszary peryferii rozwojowych		Polska - €	Obszary polaryzacji rozwoju		Obszary peryferii rozwojowych	
		mld €	PL=100%	mld €	PL=100%		€	PL=100%	€	PL=100%
<i>ludność</i>	3,7	1,4	37,2	0,9	25,0	96,0	92,9	96,8	102,5	106,8
<i>rynek pracy</i>	2,2	0,9	39,7	0,5	22,8	56,6	59,6	105,3	55,3	97,8
<i>infrastruktura techniczna</i>	6,7	3,3	48,9	1,4	21,0	175,8	227,3	129,2	157,6	89,7
<i>sytuacja finansowa</i>	7,0	3,1	43,5	1,2	16,9	184,1	209,3	113,7	133,1	72,4
<i>innowacyjna gospodarka</i>	2,3	1,4	59,5	0,3	14,9	61,1	94,7	154,9	38,9	63,6
<b>RAZEM</b>	21,9	9,9	45,1	4,4	19,9	573,7	683,8	119,2	487,3	85,0

Źródło: Obliczenia własne na podstawie danych Ministerstwa Rozwoju Regionalnego, stan na 31.12.2010 r.

lejne pod względem pozyskanych nakładów na mieszkańca aspekty rozwojowe, którymi są podobnie jak w ujęciu bezwzględny *infrastruktura techniczna i otoczenie biznesu oraz sytuacja finansowa i poziom zamożności*.

Z kolei 126 powiatów zidentyfikowanych w badaniu jako obszary peryferii rozwojowych, stanowiąc 33% wszystkich powiatów w Polsce, zrealizowało ok. 20% ogółu inwestycji, których wartość wyniosła ok. 4,4 mld €. Powiaty te stanowiąc miejsca potencjalnej koncentracji problemów strukturalnych rozwoju społeczno-gospodarczego, charakteryzują się niskimi w porównaniu do średnich w kraju wartościami wskaźników interwencji polityki spójności *per capita*. Poziom inwestycji na obszarach peryferii rozwojowych był niższy o 15% niż przeciętny w kraju, co potwierdza wskazywaną już prawidłowość braku zgodności faktycznej aktywności beneficjentów polityki spójności, realizowanej zgodnie z podejściem wyrównawczym z zakładaną jej koncentracją na obszarach słabszych ekonomicznie. Struktura środków pozyskanych przez obszary peryferii rozwojowych w ujęciu bezwzględny, podobnie jak w przypadku obszarów polaryzacji rozwoju, nie odbiega zasadniczo od struktury krajowej z dominującą pozycją nakładów na *infrastrukturę techniczną i dostępność przestrzenną*, na którą na tych obszarach przeznaczono 1,4 mld €, stanowiących jednak jedynie 21% całkowitej puli środków wydatkowanych na ten cel w kraju (tab. 2). Nie zmienia to faktu, że nakłady te zajmują pierwsze miejsce w strukturze pozyskanych środków polityki spójności zainwestowanych na obszarach peryferii rozwojowych (32,1%). Ważnymi kierunkami interwencji programów operacyjnych w badanym okresie na obszarach peryferii rozwojowych w Polsce są także działania wspierające *ludność i osadnictwo oraz rynek pracy i strukturę gospodarki*. W ujęciu względnym pierwsze miejsce w strukturze inwestycji zrealizowanych na obszarach peryferii rozwojowych zajmują inwestycje prowadzone w aspekcie *ludność i osadnictwo*, które jako jedyne przewyższają przeciętne nakłady na 1 mieszkańca ponoszone w kraju.

### **Rola ośrodków regionalnych w procesie absorpcji środków europejskich<sup>16</sup>**

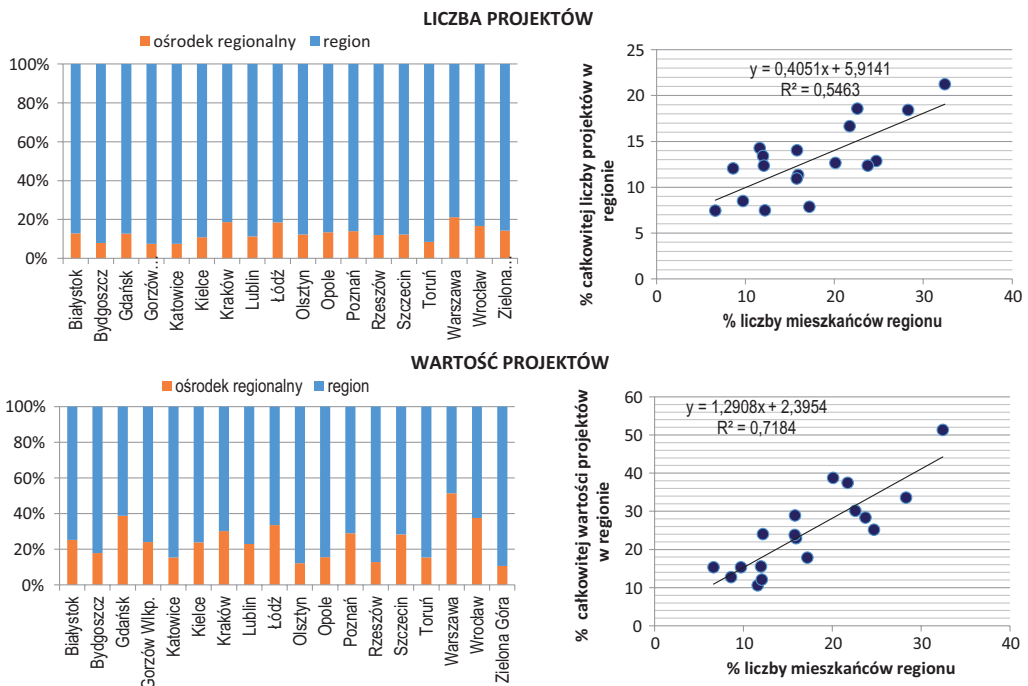
Specyficzną grupę obszarów w strukturze polskiej przestrzeni społeczno-ekonomicznej stanowią ośrodki regionalne pełniące funkcję stolic wojewódzkich. Niezależnie od ich pozycji w kraju, są one obszarami o najwyższym poziomie rozwoju społeczno-gospodarczego w systemie wewnątrzregionalnym, należąc bez wyjątków do obszarów polaryzacji rozwoju. Wyniki przeprowadzonych badań wskazują, że w okresie 2004-2013 ośrodki te utrzymały swoją wysoką pozycję na skali poziomu rozwoju społeczno-gospodarczego, przy czym wyższe wskaź-

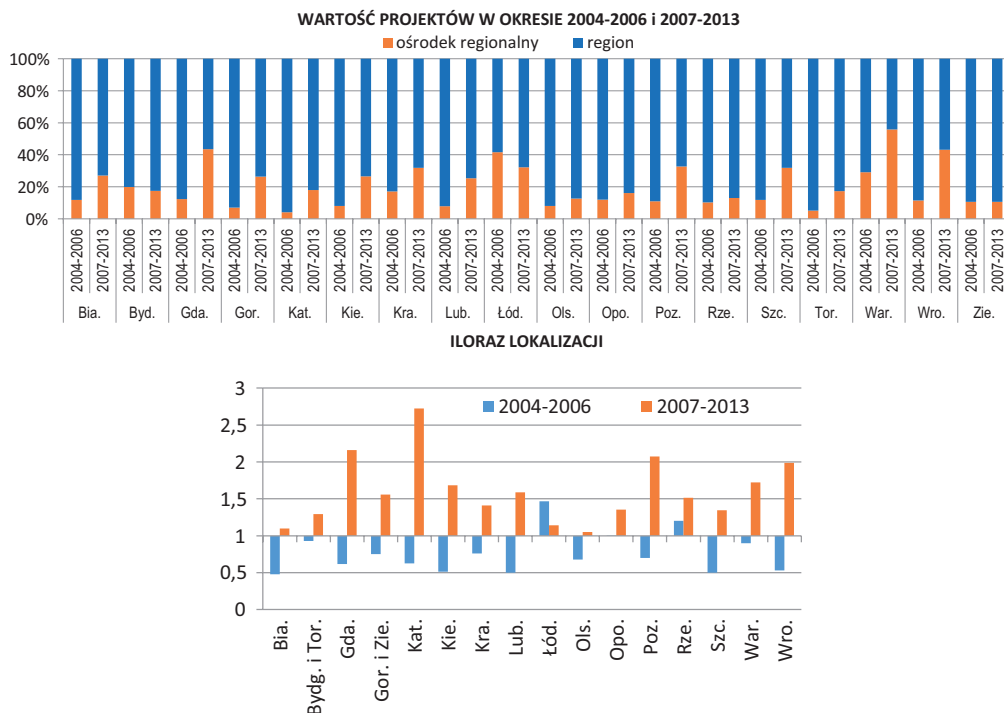
16 Patrz: Churski, P., Perdał R., Herodowicz T., *Can a biggie get more? The place of regional centres in process of absorption of European funds in Polish voivodeships*, Europa XXI, t. 30, Instytut Geografii i Przestrzennego Zagospodarowania PAN, Warszawa (w druku) 2016a. Artykuł stanowi podsumowanie częściowych wyników realizacji projektu własnego Zakładu Analizy Regionalnej Instytutu Geografii Społeczno-Ekonomicznej i Gospodarki Przestrzennej UAM *Rola środków europejskich dla rozwoju miast Polski – 10 lat członkostwa w Unii Europejskiej*.

niki dynamiki rozwoju obserwowane były w największych miastach. Specyfika tych obszarów nie dotyczy jedynie ich poziomu rozwoju społeczno-gospodarczego, ale wiąże się także z ich udziałem w procesie absorpcji środków UE. Określenie roli ośrodków regionalnych w procesie absorpcji środków UE na tle regionów przeprowadzono w układzie, który obejmował liczbę, wartość i strukturę projektów.

Udział ośrodków regionalnych w ogólnej liczbie realizowanych projektów w regionie w całym analizowanym okresie waha się od 7% w Gorzowie Wlkp. i Katowicach do 22% w Warszawie (ryc. 3). Należy podkreślić, że pozycja Gorzowa Wlkp. i Katowic nie jest w pełni porównywalna. Pierwsze miasto wraz z Zieloną Górą tworzy bipolarny układ regionalny województwa lubuskiego. Natomiast Katowice są miastem głównym województwa śląskiego, wchodzącym jednak w skład konurbacji górnośląskiej, przez co nie stanowią miejsca tak dużej koncentracji projektów jak w przypadku innych regionów. Największy udział projektów w liczbie wszystkich realizowanych w regionie występuje w Warszawie. Potwierdza to ponadprzeciętną pozycję tego miasta w układzie społeczno-ekonomicznym kraju, związaną z silną koncentracją podmiotów gospodarczych, a tym samym beneficjentów środków UE. Jest to szczególnie widoczne na tle bardzo zróżnicowanego rozwojowo Mazowsza. W badanych ośrodkach regionalnych stosunek liczby projektów do liczby ludności wskazuje na wysoki poziom współwystępowania tych zjawisk ( $r=0,739$ , przy  $p=0,05$ ).

**Rycina 3. Udział ośrodków regionalnych w absorpcji środków UE w regionie – liczba i wartość**





Źródło: Obliczenia własne na podstawie [www.mapadotacji.gov.pl](http://www.mapadotacji.gov.pl) (pobrano 30.09.2015 r.).

Jeszcze wyższy poziom koncentracji zaznacza się w przypadku udziału ośrodków regionalnych w wartości realizowanych projektów w regionie. Jego poziom waha się od 12% w Olsztynie i Rzeszowie (stolice województw w najsłabszym ekonomicznie obszarze Polski Wschodniej), do 51% w przypadku Warszawy, której dominacja na Mazowszu i w Polsce jest zasadniczo większa niż w zakresie liczby projektów (ryc. 3). Wartość projektów wykazuje wysoki poziom współwystępowania z liczbą ludności ( $r=0,848$ , przy  $p=0,05$ ). Prowadzi to do wniosku, że poziom koncentracji projektów UE w układzie ośrodek regionalny – region w istotnym stopniu pokrywa się z gęstością zaludnienia oraz zróżnicowaniem społeczno-ekonomicznym. Ośrodki regionalne charakteryzujące się najwyższym poziomem gęstości zaludnienia oraz stanowiące główne ośrodki rozwoju społeczno-gospodarczego są na tle regionu miejscem koncentracji relatywnie najbardziej kosztochłonnych projektów europejskich.

Udział ośrodków regionalnych w wartości absorpcji środków UE w regionach zasadniczo wzrósł w drugiej perspektywie budżetowej UE (ryc. 3). Wyjątek stanowią Bydgoszcz i Zielona Góra, które notowały minimalny spadek tego udziału. Wartość ilorazu lokalizacji określającego stopień koncentracji wartości projektów w ośrodkach regionalnych potwierdza te zależności. Większość ośrodków regionalnych notuje „niedobór” wartości pozyskanych projektów w perspektywie 2004-2006 w stosunku do ich udziału w liczbie ludności regionu. Jedynymi miastami, w których nie obser-

wuje się takiej zależności są Łódź i Rzeszów. W przypadku drugiej perspektywy (2007-2013) wszystkie ośrodki regionalne notują „nadwyżki” wartości projektów europejskich w stosunku do stopnia koncentracji ludności. Potwierdza to ich dominację w procesie absorpcji tych środków w regionach.

Udział ośrodków regionalnych w realizacji projektów europejskich na tle regionów wykazuje wyraźne zróżnicowanie w układzie funduszy i programów stanowiących źródła ich finansowania. Analizując dwa fundusze strukturalne: EFRR i EFS oraz uzupełniający je FS uwidacznia się ich specyfika wynikająca z ukierunkowania interwencji. Ośrodki regionalne charakteryzują się większym stopniem koncentracji projektów finansowanych w ramach EFRR i FS niż projektów finansowanych z EFS. Dwa pierwsze fundusze stanowią źródło wsparcia m.in. wysokonakładowych przedsięwzięć inwestycyjnych, których skala i wartość determinują ich realizację na terenach o najlepszej sytuacji finansowej, a jednocześnie o największej niewydolności układów infrastrukturalnych, do których zalicza się stolice województw. Wiąże się to z koniecznością zapewnienia wkładu własnego, finansowanego z budżetu lokalnego, którego wartość przy największych projektach przekracza możliwości mniejszych samorządów lokalnych. W rezultacie stopień koncentracji środków EFRR w ośrodkach regionalnych waha się od 10% dla Gorzowa Wlkp. do 45% dla Warszawy (ryc. 4). Udziały dla FS są jeszcze wyższe, ale zamykają się w szerszym przedziale od 1% w Olsztynie do 63% w Warszawie (średnio 26%). Inaczej przedstawia się sytuacja w zakresie EFS, który stanowi źródło finansowania przede wszystkim projektów miękkich, ukierunkowanych na poprawę sytuacji na rynku pracy, charakteryzujących się relatywnie niższymi budżetami. W tym przypadku ich dostępność dla obszarów peryferyjnych jest wyższa, co skutkuje niższym poziomem koncentracji wartości projektów w stolicach województw. W konsekwencji poziom koncentracji wartości projektów finansowanych z EFS w Katowicach wynosi jedynie 5%, a w Łodzi 18%, przy średniej utrzymującej się na poziomie 9% (ryc. 4). Zgodnie ze wskazanymi prawidłowościami wartości ilorazu lokalizacji dla ośrodków regionalnych wykazują najczęściej „nadwyżki” koncentracji wartości projektów europejskich finansowanych z EFRR i FS oraz „niedobory” w przypadku EFS. Wskazane tendencje potwierdza również analiza struktury wartości projektów europejskich w układzie programów operacyjnych. Najwyższymi poziomami koncentracji w ośrodkach regionalnych cechują się projekty z programów wspierających wysokonakładowe przedsięwzięcia inwestycyjne, a najniższymi - programy operacyjne ukierunkowane na interwencję w kapitał ludzki i społeczny. W latach 2004-2006 najwyższym poziomem koncentracji i „nadwyżką” charakteryzowały się projekty realizowane w ramach Sektorowego Programu Operacyjnego Transport (SPO T) – np. w Rzeszowie  $LQ=10,3$ . Najniższy poziom koncentracji odpowiadał projektom realizowanym w ramach SPO RZL – np. w Białymstoku  $LQ=0,02$ . W latach 2007-2013 wskazana tendencja nadal się utrzymywała. Największą koncentrację i „nadwyżki” notowano w przypadku projektów realizowanych z PO IŚ – przykładowo w Katowicach  $LQ=3,8$ . Natomiast najniższe wartości koncentracji i „niedobory” występują dla projektów wykonanych przy wsparciu PO KL - dla Wrocławia  $LQ=0,3$ .



Podsumowaniem przedmiotowej analizy jest ocena pozycji ośrodków regionalnych na skali koncentracji liczby i wartości projektów w układzie kategorii interwencji (ryc. 4). Uzyskany rozkład wskazuje na wyraźnie zaznaczające się prawidłowości. Pierwsza z nich wynika z istotnego zróżnicowania realizacji projektów, zwłaszcza inwestycyjnych, w zależności od możliwości finansowych beneficjentów. Pozwala to wskazać na ośrodki regionalne, jako na obszary uprzywilejowane w możliwości pozyskiwania największych pod względem poziomu dofinansowania projektów europejskich. Druga prawidłowość wiąże się ze zróżnicowaniem w przestrzeni regionu lokalizacji poszczególnych rodzajów aktywności społeczno-gospodarczych. Przykładowo usługi administracyjne oraz działalności zaspakajające potrzeby wyższego rzędu w dużym stopniu koncentrują się w ośrodkach regionalnych. W konsekwencji najwyższe poziomy koncentracji obserwuje się w przypadku projektów europejskich związanych z administracją, kulturą i sztuką oraz telekomunikacją i e-usługami. Najniższe zaś odpowiadają projektom wspierającym pracę i integrację społeczną, rewitalizację oraz rozwój firm (ryc. 4).

### **Ocena dostępu do środków europejskich w opiniach samorządów terytorialnych ośrodków regionalnych<sup>17</sup>**

Ocena dostępu do środków europejskich określająca opinie władz samorządowych ośrodków regionalnych w Polsce, przeprowadzona w niniejszym badaniu, obejmowała m.in. analizę opinii administracji samorządowej tych ośrodków na temat szans i zagrożeń wynikających z dotychczasowego dostępu do europejskich środków publicznych przeznaczanych na działania prorozwojowe. Opinie te zostały uporządkowane w układzie analizy SWOT. Wśród silnych stron i szans wynikających z dostępu do środków europejskich samorządy lokalne badanych ośrodków regionalnych wskazały najczęściej cztery korzyści (tab. 3). Najwięcej ośrodków, 14 z 16 badanych<sup>18</sup>, za najważniejszą korzyść uznało możliwość poprawy poziomu i warunków życia mieszkańców poprzez przede wszystkim realizację działań z zakresu polityki społecznej oraz poprawę warunków komunikacyjnych.

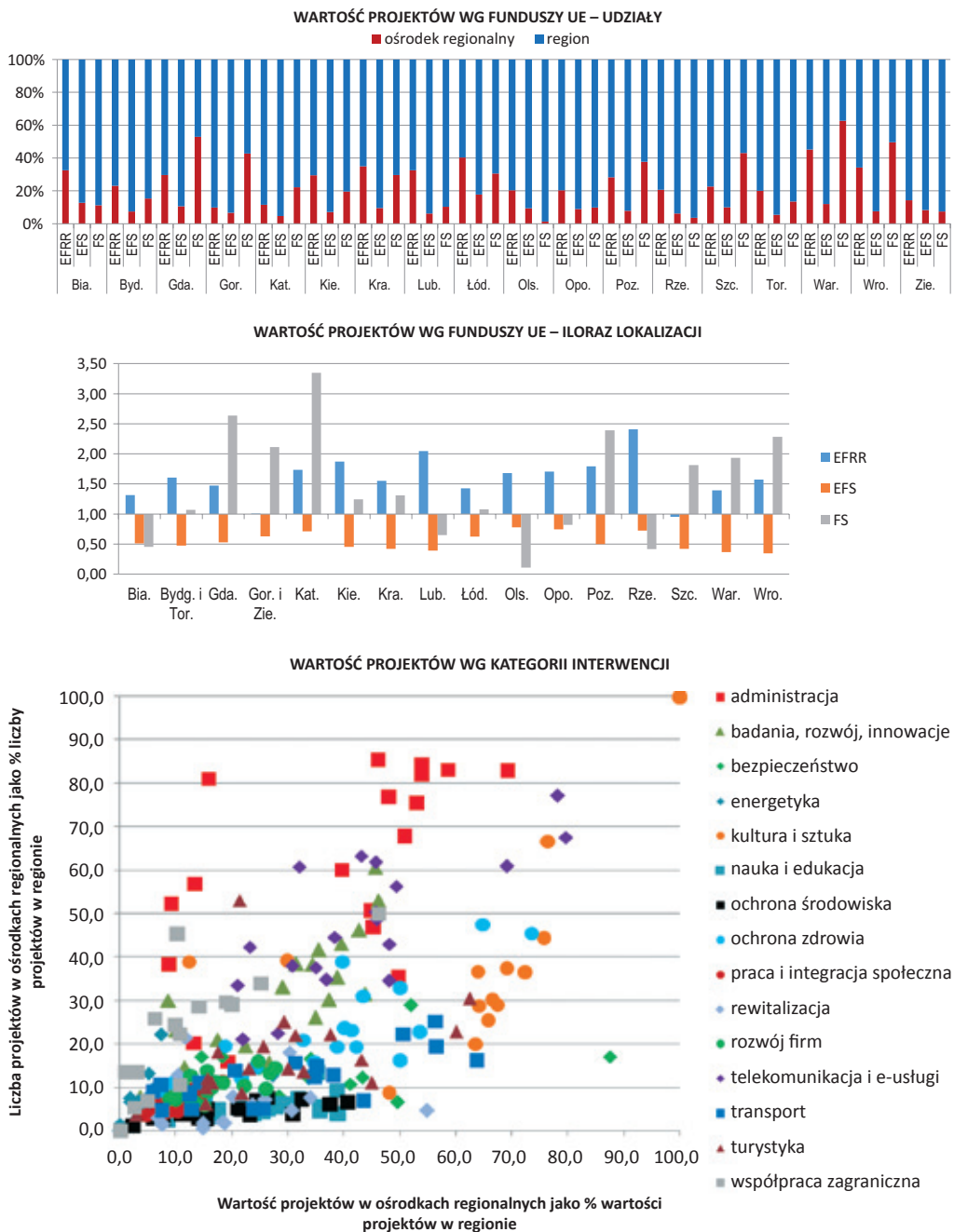
Fakt ten potwierdza obserwowaną w praktyce polskiej polityki rozwoju koncentrację działań władz na wewnętrznym wymiarze polityki miejskiej, której głównym zadaniem jest poprawa jakości życia w mieście oraz eliminacja barier rozwojowych wynikających przede wszystkim z nadmiernej koncentracji mieszkańców w jego granicach.

17 Patrz: Churski P., Perdał R., Herodowicz T., *Rola środków polityki spójności pozyskiwanych przez samorząd terytorialny w rozwoju społeczno-gospodarczym ośrodków regionalnych w Polsce*, Samorząd Terytorialny, t. 7/2016, Wolters Kluwer S.A., Warszawa (w druku) 2016b. Artykuł stanowi podsumowanie częściowych wyników realizacji projektu własnego Zakładu Analizy Regionalnej Instytutu Geografii Społeczno-Ekonomicznej i Gospodarki Przestrzennej UAM *Rola środków europejskich dla rozwoju miast Polski – 10 lat członkostwa w Unii Europejskiej*.

18 Przedstawiciele władz samorządowych Białegostoku nie udzieli odpowiedzi w analizowanym zakresie.



Rycina 4. Udział ośrodków regionalnych w absorpcji środków UE w regionie – fundusze i kierunki interwencji



Źródło: Obliczenia własne na podstawie [www.mapadotacji.gov.pl](http://www.mapadotacji.gov.pl) (pobrano 30.09.2015 r.).

Za równie istotną korzyść respondenci uznali (12 z 16 miast) możliwość wykorzystania środków europejskich do poprawy możliwości inwestycyjnych miasta i atrakcyjności inwestycyjnej jego obszaru. Wskazanie to sugeruje istotne znaczenie środków europejskich w kształtowaniu sytuacji budżetowej samorządu terytorialnego ośrodków regionalnych w Polsce. Dotyczy to przede wszystkim oddziaływania na możliwości inwestycyjne miasta, które, co należy uznać za zjawisko pozytywne, wykorzystywane są w dużym stopniu do wzmacniania atrakcyjności inwestycyjnej tych ośrodków w celu trwałej poprawy ich bazy ekonomicznej.

**Tabela 3. Analiza SWOT wpływu dostępu do środków europejskich na rozwój ośrodków regionalnych w świetle opinii samorządu terytorialnego**

SILNE STRONY I SZANSE	MIASTO	SŁABE STRONY I ZAGROŻENIA	MIASTO
poprawa poziomu i warunków życia mieszkańców poprzez przede wszystkim realizację działań z zakresu polityki społecznej oraz poprawę warunków komunikacyjnych	BYDGOSZCZ GDAŃSK GORZÓW WLKP. KIELCE KRAKÓW LUBLIN OLSZTYN POZNAŃ RZESZÓW SZCZECIN TORUŃ WARSZAWA WROCŁAW ZIELONA GÓRA	wzrost zadłużenia budżetów lokalnych	BYDGOSZCZ GDAŃSK GORZÓW WLKP. KRAKÓW LUBLIN ŁÓDŹ OLSZTYN POZNAŃ RZESZÓW WARSZAWA WROCŁAW ZIELONA GÓRA
poprawa możliwości inwestycyjnych miasta i atrakcyjności inwestycyjnej jego obszaru	BYDGOSZCZ GDAŃSK KIELCE KRAKÓW LUBLIN ŁÓDŹ OLSZTYN RZESZÓW SZCZECIN TORUŃ WROCŁAW WARSZAWA	nieefektywne i zmienne procedury administracyjne, w tym opóźnienia w rozpatrywaniu wniosków o płatność	BYDGOSZCZ KRAKÓW LUBLIN ŁÓDŹ OPOLE POZNAŃ RZESZÓW TORUŃ WROCŁAW
wzrost konkurencyjności	GDAŃSK GORZÓW WLKP. KIELCE LUBLIN OLSZTYN OPOLE POZNAŃ SZCZECIN WARSZAWA WROCŁAW ZIELONA GÓRA	uzależnienie od środków europejskich i ukierunkowania ich interwencji	GDAŃSK GORZÓW WLKP. LUBLIN OLSZTYN WROCŁAW SZCZECIN ZIELONA GÓRA

SILNE STRONY I SZANSE	MIASTO	SŁABE STRONY I ZAGROŻENIA	MIASTO
poszerzenie możliwości wzmacniania bazy ekonomicznej	GDAŃSK KRAKÓW LUBLIN ŁÓDŹ OPOLE POZNAŃ RZESZÓW ZIELONA GÓRA	wzrost bezrobocia	GDAŃSK
poprawa atrakcyjności turystycznej	GDAŃSK SZCZECIN ZIELONA GÓRA	obniżenie konkurencyjności gospodarki	POZNAŃ
promocja i współpraca międzynarodowa	BYDGOSZCZ KRAKÓW LUBLIN	trudności kooperacyjne z partnerami	RZESZÓW
rozwój kapitału ludzkiego	KRAKÓW GORZÓW WLKP. RZESZÓW	brak	KIELCE LUBLIN
poprawa funkcjonowania administracji	OPOLE RZESZÓW SZCZECIN		
poprawa warunków środowiskowych	KRAKÓW LUBLIN		
tworzenie nowych miejsc pracy	LUBLIN RZESZÓW		
wzrost innowacyjności gospodarki	GORZÓW WLKP.		
wzrost spójności wewnętrznej i zewnętrznej	TORUŃ		
poprawa trafności formułowania potrzeb rozwojowych	KATOWICE		
b/d	BIAŁYSTOK		

Źródło: Opracowanie własne na podstawie wyników badań ankietowych zrealizowanych w ramach projektu Rola środków europejskich dla rozwoju miejskich ośrodków regionalnych Polski – 10 lat członkostwa w Unii Europejskiej.

Wśród silnych stron i szans wynikających z dostępu do środków europejskich, które zauważone zostały przez co najmniej połowę badanych samorządów znajduje się jeszcze możliwość poprawy pozycji konkurencyjnej miasta. Dostrzeganie potrzeby kształtowania i wzmacniania pozycji konkurencyjnej miasta w warunkach postępującej globalizacji i integracji międzynarodowej uznać należy za bardzo pożądane. Szkoda jednak, że nie jest to powszechna korzyść wynikająca z dostępu do środków europejskich, którą wskazują przedstawiciele administracji samorządowej ośrodków regionalnych. Może to sugerować tendencje do wykorzystywania tych środków w zbyt dużym stopniu dla realizacji działań prorozwojowych ukierunkowanych na mieszkańców w perspektywie krótkookresowej, w wielu przypadkach ograniczających się do efektów popytowych i przynoszą-

cych wzrost obciążeń budżetowych wynikających z renty użytkowania nowych obiektów, w stosunku do działań prorozwojowych ukierunkowanych na inwestorów, co mogłoby przynosić trwałe i długookresowe efekty podażowe. Wśród pozostałych korzyści, wskazywanych mniej powszechnie w grupie badanych ośrodków, wskazać należy na: możliwość poprawy atrakcyjności turystycznej, promocję i współpracę międzynarodową, rozwój kapitału ludzkiego, poprawę funkcjonowania administracji, poprawę warunków środowiskowych oraz tworzenie nowych miejsc pracy. Pozostałe silne strony i szanse wskazywane przez respondentów należą do postrzeganych indywidualnie. Szkoda, że w tej grupie znalazły się korzyści związane z kształtowaniem zewnętrznego wymiaru polityki miejskiej, na które uwagę zwrócili jedynie przedstawiciele Torunia. Mają one bowiem podstawowe znaczenie w poprawie spójności wewnętrznej regionów poprzez zwiększanie zasięgu oddziaływania rozwojowego ich stolic w stosunku do bliższego i dalszego ich otoczenia. Świadczy to o nieskutecznej realizacji założeń polskiej polityki rozwoju, która zakłada wykorzystanie ośrodków regionalnych i subregionalnych w zmniejszaniu różnic rozwojowych poprzez wzmacnianie ich relacji funkcjonalnych, zgodnie z założeniami modelu polaryzacyjno-dyfuzyjnego. Opinie respondentów na temat słabych stron i zagrożeń wynikających z dostępu do środków europejskich są bardziej zbieżne<sup>19</sup>. Powszechnie wskazywanym zagrożeniem przez reprezentantów samorządów ośrodków regionalnych jest wzrost zadłużenia budżetów lokalnych wynikający z angażowania środków inwestycyjnych do zapewnienia wkładów własnych dla potrzeb realizacji projektów europejskich. Dostęp do środków europejskich traktowany jest jako czasowa szansa na realizację przedsięwzięć niemożliwych do wykonania w warunkach budżetów lokalnych jakimi dysponuje samorząd terytorialny w Polsce. Dotyczy to nawet ośrodków regionalnych, które posiadając status miast na prawach powiatu, charakteryzują się relatywnie najlepszą sytuacją finansową. W konsekwencji aktywne w procesie inwestycyjnym wykorzystujące europejskie środki publiczne samorządy, przy niewystarczających środkach własnych, zwiększają poziom zadłużenia budżetowego, dążąc do wykorzystania tej szansy. Skutkuje to w wielu przypadkach niebezpiecznym poziomem zadłużenia przekraczającym realnie dopuszczalny próg określony w ustawie o finansach publicznych. Ponadto przedstawiciele samorządów terytorialnych ośrodków regionalnych zwracają uwagę na nieefektywne i zmienne procedury administracyjne obowiązujące w realizacji oraz rozliczaniu projektów, wśród których szczególnie dotkliwy jest fakt powtarzających się opóźnień w rozpatrywaniu wniosków o płatność, pogłębiający wskazane już problemy finansowe. Bardzo istotnym zagrożeniem, równie powszechnie wskazanym przez respondentów, jest zwiększające się uzależnienie od środków europejskich i ukierunkowania interwencji finansowanej z ich środków. Z jednej strony zwrócenie uwagi na ten fakt ocenić należy pozytywnie, z drugiej jednak strony powstaje pytanie dlaczego samorząd terytorialny widząc to zagrożenie swoimi decyzjami pogłębia jego skalę. Wydaje się, że wynika to jednak z uwarunkowań zewnętrznych, niezależnych od samorządów terytorialnych, a wynikających z zasad ich finansowania

19 Przedstawiciele władz samorządowych Kielc i Lublina nie udzieli odpowiedzi w analizowanym zakresie.

określonych ustawą oraz niebezpiecznej praktyki programowania polityki rozwoju w Polsce w zbyt dużym stopniu podporządkowanej priorytetom europejskim, a w zbyt małym stopniu uwzględniającym specyficzne potrzeby poszczególnych terytoriów, zgodnie z podejściem *place-based policy*. Pozostałe zagrożenia wskazane przez respondentów mają charakter pojedynczych wskazań. Wśród nich znajdują się takie jak: wzrost bezrobocia, obniżenie konkurencyjności gospodarki, trudności kooperacyjne z partnerami. Dwa pierwsze są zgodne z dobrze rozpoznanymi prawidłowościami związanymi ze zmniejszaniem się udziału mechanizmu rynkowego w warunkach pogłębiającej się interwencji publicznej. Trzecie wskazanie może świadczyć o występujących w praktyce polskich uwarunkowań prawnych i administracyjnych barierach wdrażania zasady *multilevel governance*, zwłaszcza w realiach zasad regulujących wykorzystanie środków europejskich.

### Zakończenie

Podsumowanie przeprowadzonego przeglądu wybranych wyników badań nad zróżnicowaniami rozwojowymi w Polsce oraz wpływem interwencji polityki spójności na ich zmiany zrealizowanych przez Zespół Zakładu Analizy Regionalnej Instytutu Geografii Społeczno-Ekonomicznej i Gospodarki Przestrzennej Uniwersytetu im. Adama Mickiewicza w Poznaniu prowadzi do sformułowania najważniejszych wniosków i uwag polemicznych. Należy podkreślić, że ich zakres odnosi się do całościowych wyników przedmiotowych analiz, które ze względu na ograniczoną objętość niniejszego opracowania zostały w nim jedynie zasygnalizowane przy jednoczesnym odniesieniu do tekstów źródłowych zawierających pełne informacje na temat założeń, postępowania badawczego oraz całościowych rezultatów poszczególnych badań.

Uzyskane wyniki potwierdziły, podkreślany w literaturze przedmiotu, wpływ kapitału terytorialnego na zróżnicowanie procesów rozwojowych oraz wielkość absorpcji środków polityki spójności wspierających konwergencję społeczno-ekonomiczną. Ich wartością dodaną jest zwrócenie uwagi na specyfikę tych zależności, które obok ogólnych trendów wskazywanych przez innych autorów<sup>20</sup> dotyczą w szczególności sposobu regionów nowych państw członkowskich, których gospodarka i społeczeństwo przez dziesięciolecia funkcjonowały w realiach gospodarki centralnie sterowanej. Uzyskane rezultaty dla Polski, zgodnie z założeniem Bockman and Eyal<sup>21</sup> dotyczącymi potrzeby testowania wiedzy ekonomicznej w warunkach państw postsocjalistycznych, stanowią podstawę dla formułowania ważnych poznawczo rekomendacji dotyczących programowania i realizacji polityki spójności.

20 Rodríguez-Pose A., *Do Institutions Matter for Regional Development?* „Regional Studies”, Vol. 47, no. 7/2013, s. 1034-1047; McCann P., Varga A., *The Reforms to the Regional and Urban Policy of the European Union: EU Cohesion Policy*, „Regional Studies”, Vol. 49, no. 8/2015, s. 1255-1257; Partridge M. D., Rickman D. S., Olfert M. R., Tan Y., *When Spatial Equilibrium Fails: Is Place-Based Policy Second Best?*, „Regional Studies”, Vol.49, no. 8/2015, s. 1303-1325.

21 Bockman J. K., Eyal G., *Eastern Europe as a laboratory for economic knowledge: The transnational roots of neoliberalism*, „American Journal of Sociology”, 108 (2), 2002, s. 310-352.

Rozkład obszarów polaryzacji rozwoju i peryferii rozwojowych oraz jego zmienność w czasie potwierdzają występowanie istotnego zróżnicowania rozwojowego przestrzeni ekonomicznej Polski w układzie lokalnym (jednostek terytorialnych szczebla powiatowego). Podstawowym ujawnionym podziałem tej przestrzeni jest podział na obszary stagnujące (o niskim poziomie rozwoju), zlokalizowane głównie w Polsce Centralnej i Wschodniej, a obszary przejściowe (o przeciętnym poziomie rozwoju), skoncentrowane w Polsce Zachodniej. Taki podział może być podstawą do percepcji obszaru kraju w kontekście stereotypowej dychotomii: Polska A i Polska B, w znacznym stopniu nawiązującej do przebiegu granic reliktowych (społeczno-gospodarczych podziałów tzw. długiego trwania). Jednak na ten obraz nakłada się astrefowy układ rozproszonych obszarów polaryzacji rozwoju, stanowiących swego rodzaju „super-kategorię” jednostek charakteryzujących się jego wysokim poziomem, którymi są duże i średnie ośrodki miejskie oraz obszary silnej aglomeracji przestrzennej i koncentracji sektorowej przemysłu. Ten fakt sprawia, że najistotniejszym rzeczywistym podziałem polskiej przestrzeni ekonomicznej jest podział na ośrodki miejskie/przemysłowe i peryferie. Istnieje jednak dostrzegalny związek pomiędzy dychotomią występującą w układzie wschód-zachód a dychotomią wynikającą z układu wielośrodkowej aglomeracji przestrzennej potencjału ludnościowego i gospodarczego (ośrodki miejskie/przemysłowe - obszary peryferyjne). Związek ten polega na tym, że obszary polaryzacji rozwoju w zachodnich regionach kraju cechują się mniejszym gradientem poziomu rozwoju w relacji do swojego otoczenia (ich bezpośrednie sąsiedztwo stanowią jednostki relatywnie nieznacznie odbiegające poziomem rozwoju), niż ich odpowiedniki w Polsce Centralnej i Wschodniej, gdzie w bezpośrednim sąsiedztwie relatywnie niewielkich obszarów wzrostu występują często jednostki permanentnej stagnacji. Pamiętając, że proces dyfuzji rozwoju jest uwarunkowany poziomem gospodarczym sąsiadujących ze sobą jednostek i zachodzi z większą intensywnością pomiędzy jednostkami o bardziej zbliżonym poziomie rozwoju społeczno-gospodarczego interwencja polityki spójności w Polsce musi ulec zmianie. Powinna ona w większym stopniu wspierać z jednej strony wzmocnienie właściwie identyfikowanych potencjałów endogenicznych „Polski lokalnej”, a z drugiej strony wzmocniać siłę i zasięg oddziaływania obszarów polaryzacji rozwoju. Niestety dotychczasowe efekty interwencji polityki spójności w Polsce są niesatysfakcjonujące i wbrew założeniom prowadzą do pogłębiania skali zróżnicowań rozwojowych na poziomie lokalnym. Świadczy to o ograniczonych możliwościach absorpcji impulsów rozwojowych przez obszary zmarginalizowane pod względem społeczno-ekonomicznym, z nierozpoznanymi lub zidentyfikowanymi niewłaściwie endogennymi zasobami rozwojowymi.

Na tle tych ogólnych prawidłowości można sformułować kilka uwag szczegółowych odnoszących się do wybranych wyzwań polityki spójności w Polsce:

- Absorpcja środków europejskich w Polsce koncentruje się na obszarach polaryzacji rozwoju, w tym w największych miastach stanowiących ośrodki regionalne, co stoi w sprzeczności z założeniami modelu wyrównawczego polityki rozwoju. Powstaje zatem pytanie: Czy należy przeciwdziałać tym prawidłowościom, wiedząc że koncentracja środków na obszarach polaryzacji rozwoju zapewnia większą stopę zwrotu podjętych inwestycji niż w przypadku ich realizacji na obszarach słabszych ekonomicznie? Wydaje się, że działanie takie będzie bezzasadne, a co ważniejsze bezskuteczne. Bardziej rozsądna wydaje się modyfikacja ukierunkowania, poprzez zmianę kryteriów dostępu w postępowaniach konkursowych, pozyskiwanych przez obszary polaryzacji rozwoju środków na tworzenie warunków do silniejszej i charakteryzującej się większym zasięgiem dyfuzji procesów rozwojowych. Właściwym wydaje się również zwiększenie aktywności ośrodków subregionalnych i lokalnych, których oddziaływanie na otoczenie pozwoliłoby bardziej skutecznie pokryć „białe plamy” rozwoju w polskiej przestrzeni społeczno-gospodarczej,
- Wzrost utrzymania koncentracji interwencji środków UE na finansowaniu projektów związanych z rozbudową i modernizacją infrastruktury technicznej przy dużo niższym poziomie interwencji w pozostałych aspektach rozwojowych oddala w czasie możliwość osiągnięcia bezpośrednich efektów interwencji o charakterze podażowym, a ponadto zwiększa zagrożenia związane z wystąpieniem zjawiska zbyt dużego zainwestowania infrastrukturalnego (renta użytkownika). Jak zatem należy skutecznie dywersyfikować strukturę absorpcji nie obniżając jednocześnie efektywności podjętej interwencji poprzez realizację projektów miękkich finansujących bieżące zadania społeczne lub stanowiących jednostkowe, w wielu wypadkach wysokonakładowe ale nieuzasadnione działania? Wymaga to szybkiego zweryfikowania stosowanych obecnie instrumentów i wdrożenia zintegrowanych programów interwencji ukierunkowanej na rozwój kapitału ludzkiego i społecznego, które byłyby realizowane zarówno na poziomie krajowym, jak i regionalnym.
- Brak satysfakcjonującej skuteczności rozwojowej dotychczasowej interwencji środków europejskich w Polsce może być konsekwencją przesunięcia jej rezultatów w czasie, złego jej ukierunkowania lub też braku możliwości ich identyfikacji. Ten ostatni wariant stawia przed ewaluacją zadanie pełnienia funkcji skutecznego narzędzia identyfikacji efektywności udzielanej interwencji. Narzędzia skutkującego nie tylko oceną ale, co jest istotą ewaluacji, korektą podejmowanych działań, bazującego na jednoznacznych podstawach teoretycznych<sup>22</sup>. Wydaje się, że pomimo bardzo dynamicznego rozwoju tego rodzaju aktywności, nie wszystkie, niezależnie od ich nazwy, można dzisiaj określić mianem pełnej ewaluacji.

22 Borowczak A., *Stan i kierunki badań ewaluacyjnych rozwoju regionalnego w Polsce*, „Studia KPZK PAN”, CXL (II), 2012, s. 247-259; Borowczak A., *Zastosowanie wybranych koncepcji badawczych w ewaluacji programów rozwoju regionalnego*, Praca Doktorska, AMUR, <https://repozytorium.amu.edu.pl/handle/10593/13277> (pobrano 20 maja 2015).

Biorąc pod uwagę postępujący proces rozwoju kraju i regionów skutkujący zmianą uwarunkowań realizacji polityki spójności po 2023 roku należy na koniec zadać pytanie: Czy jesteśmy właściwie przygotowani do koordynacji procesów rozwojowych po 2020 roku? Wydaje się, że nie, a skala zagrożeń wynikająca z tego stanu powinna determinować podjęcie pilnych prac nad stworzeniem zintegrowanego systemu realizacji polityki rozwoju, którego większościowym źródłem finansowania przestaną być europejskie środki publiczne.



## Bibliografia

- Barca F.**, *An Agenda for a Reformed Cohesion Policy*, A place-based approach to meeting European Union challenges and expectations, Independent Report, 2009
- Barca. F., McCann P., Rodriguez-Pose A.**, *The case for regional development intervention: place-based versus place-neutral approaches*, „Journal of Regional Science”, 52, 1/2012, s. 134-152
- Bockman J. K., Eyal G.**, *Eastern Europe as a laboratory for economic knowledge: The transnational roots of neoliberalism*, „American Journal of Sociology”, 108, 2/2002, s. 310-352
- Borowczak A.**, *Stan i kierunki badań ewaluacyjnych rozwoju regionalnego w Polsce*, „Studia KPZK PAN”, CXL, II/2012, s. 247-259
- Borowczak A.**, *Zastosowanie wybranych koncepcji badawczych w ewaluacji programów rozwoju regionalnego*, Praca Doktorska, AMUR, <https://repozytorium.amu.edu.pl/handle/10593/13277>, 2015, (pobrano 20 maja 2015)
- Capello R.**, *Regional economics*, Routledge, New York 2014
- Capello R., Nijkamp P.**, (red.), *Handbook of Regional Growth and Development Theories*, Edward Elgar, Cheltenham, Northampton 2009
- Churski P.**, *Model polaryzacyjno-dyfuzyjny w przemianach polityki spójności – konsekwencje dla ukierunkowania polityki rozwoju*, [w:] (red.) P.Churski, „Biuletyn Instytutu Geografii Społeczno-Ekonomicznej i Gospodarki Przestrzennej UAM, seria Rozwój Regionalny i Polityka Regionalna”, Bogucki Wydawnictwo Naukowe, 25/2014, s. 13-27
- Churski P.**, (red.), *Rozwój społeczno-gospodarczy a kształtowanie się obszarów wzrostu i obszarów stagnacji gospodarczej w Polsce*, Biuletyn Instytutu Geografii Społeczno-Ekonomicznej i Gospodarki Przestrzennej UAM, seria Rozwój Regionalny i Polityka Regionalna”, Bogucki Wydawnictwo Naukowe, 25/2014
- Churski P.**, (red.), *The social and economic growth vs. the emergence of economic growth and stagnation areas*, Bogucki Wydawnictwo Naukowe, Poznań 2014
- Churski P., Borowczak A., Perdał R.**, *Struktura inwestycji finansowanych ze środków unijnych a czynniki rozwoju w obszarach wzrostu w Polsce*, [w:] *Polska 10 lat członkostwa w Unii Europejskiej*, (red.) E.Małuszyńska, G.Mazur, I.Musiałkowska, Wydawnictwo Uniwersytetu Ekonomicznego w Poznaniu, Poznań 2015
- Churski P., Borowczak A., Perdał R.**, *Czynniki rozwoju obszarów stagnacji w Polsce a ukierunkowanie interwencji środków unijnych*, *Studia Obszarów Wiejskich*”, Tom 37, Wiejskie obszary funkcjonalne, Komisja Obszarów Wiejskich PTG, Instytut Geografii i Przestrzennego Zagospodarowania PAN, Warszawa 2015
- Churski P., Perdał R., Herodowicz T.**, *Can a biggie get more? The place of regional centres in process of absorption of European funds in Polish voivodeships*, *Europa XXI*, t. 30, Instytut Geografii i Przestrzennego Zagospodarowania PAN, Warszawa 2016 (w druku)

- Churski, P., Perdał R., Herodowicz T.**, *Rola środków polityki spójności pozyskiwanych przez samorząd terytorialny w rozwoju społeczno-gospodarczym ośrodków regionalnych w Polsce*, Samorząd Terytorialny, t. 7/2016, Wolters Kluwer S.A., Warszawa 2016 (w druku)
- Dolata M., Borowczak A.**, *Distribution of economic growth and stagnation areas in Poland, 2000-2010*, [w:] *The social and economic growth vs. the emergence of economic growth and stagnation areas*, (red.) P. Churski, Bogucki Wydawnictwo Naukowe, Poznań 2014, s. 29-42
- Dymnicki A.B., Henry D.B.**, *Use of Clustering Methods to Understand More about the Case*, „Methodological Innovations Online”, Vol. 6 no. (2)/2011, s. 6-26.  
<http://www.pbs.plym.ac.uk/mi/pdf/31-08-11/2.%20Dymnicki%20-%20pp6-26.pdf>
- Faludi A.**, *From European Spatial Development to Territorial Cohesion Policy*, „Regional Studies”, Vol. 40, no. 6/2006, s. 667-678
- Friedmann J., Alonso W.**, *Introduction*, [w:] *Regional development and planning*, (red.) J. Friedmann, W. Alonso, Cambridge, Mass: MIT Press, 1964, s. 1-13
- Friedmann J.**, *A general theory of polarized development*, Santiago: Ford Foundation, Urban and Regional Development Advisory Program in Chile, Op. 81, 1967
- Friedmann J.**, *Ogólna teoria rozwoju spolaryzowanego*, [w:] *Przestrzeń krajów trzeciego świata – problemy metodologiczne*, (red.) Rościszewski M., PZLG, IGiPZ PAN, Warszawa 1974
- Gawlikowska-Hueckel K.**, *Procesy rozwoju regionalnego w Unii Europejskiej, Konwergencja czy polaryzacja?* Wydawnictwo Uniwersytetu Gdańskiego, Gdańsk 2003
- Gorzela G.**, *Fakty i mity rozwoju regionalnego*, „Studia Regionalne i Lokalne”, nr 2 (36), Centrum Europejskich Studiów Regionalnych i Lokalnych Uniwersytetu Warszawskiego, Wydawnictwo Scholar, Warszawa 2009, s. 5-27
- Gorzela G.**, *Wykorzystanie środków Unii Europejskiej dla rozwoju kraju – wstępne analizy*, „Studia Regionalne i Lokalne”, nr 3 (57)/2014, s. 5-25
- Growing unequal?: Income Distribution and Poverty in OECD Countries*, OECD Report Paris 2008
- Hansen N.M.**, *Criteria for growth center policy*, [w:] *Growth poles and growth centers in regional planning* (red.) A. Kukliński, Mouton, Paris 1972, s. 103-124
- Jastrzębska M.**, *Znaczenie bezwrotnych środków zagranicznych jako źródeł finansowania działalności jednostek samorządu terytorialnego w latach 2004-2010*, „Finanse Komunalne”, nr 10/2011, s. 18-31
- Kronhalter F.**, *A Study of the Competitiveness of Regional Based on a Cluster Analysis. The Example of East Germany*, Halle Institute for Economic Research, Halle/Saale 2003
- Krugman P.**, *Increasing returns and economic geography*, *Journal of Political Economy*, Vol. 99, no. 3/1991, s. 483-499
- Krugman P.**, *Development geography, and economic theory*, Ohlin Lectures, MIT Press, Cambridge-London 1995

- Lucas R.E.**, *On the mechanics of economic development*, „Journal of Monetary Economics”, vol. 22, no. 1/1988, s. 3-42
- McCann P., Varga A.**, *The Reforms to the Regional and Urban Policy of the European Union: EU Cohesion Policy*, „Regional Studies”, Vol. 49 no. 8/2015, s. 1255-1257
- Michałek J. J.** (red.), *Polska w Unii Europejskiej, Dynamika konwergencji ekonomicznej*, Wydawnictwo Naukowe PWN, Warszawa 2007
- Misiąg J., Misiąg W., Tomalak M.**, *Ocena efektywności wykorzystania pomocy finansowej Unii Europejskiej jako instrumentu polityki spójności społeczno-gospodarczej oraz poprawy warunków życia*, Rzeszów 2013
- Molle W.**, *European Cohesion Policy*, Routledge, London 2007
- Morrison D. F.**, *Wielowymiarowa analiza statystyczna*, PWN, Warszawa 1990
- Nowińska-Łaźniewska E.**, *Relacje przestrzenne w Polsce w okresie transformacji w świetle teorii rozwoju regionalnego*, Prace habilitacyjne nr 13, Wydawnictwo Akademii Ekonomicznej w Poznaniu, Poznań 2004
- Perdał R., Hauke J.**, *Areas of economic growth and stagnation in Poland – growth factors*, [w:] *The social and economic growth vs. the emergence of economic growth and stagnation areas*, (red.) Churski P., Bogucki Wydawnictwo Naukowe, Poznań 2014, s. 43-62
- Partridge M. D., Rickman D. S., Olfert M. R., Tan Y.**, *When Spatial Equilibrium Fails: Is Place-Based Policy Second Best?*, „Regional Studies”, Vol. 49, no. 8/2015, s. 1303-1325
- Perroux F.**, *Note sur la nation de „pole de croissance”*, „Economie Appliquee”, vol. 8, no. 1-2/1955, s. 307-320
- Pietrzyk I.**, *Regionalne implikacje europejskiej integracji gospodarczej dla Polski*, „Zeszyty Naukowe”, nr 721, Akademia Ekonomiczna w Krakowie, Kraków 2006
- Regional Development Policies in OECD Countries*, OECD Report, Paris 2010
- Reshaping Economic Geography*, World Development Report 2009: The World Bank 2009
- Rodríguez-Pose A., Crescenzi R.**, *Research and Development, Spillovers, Innovation Systems, and the Genesis of Regional Growth in Europe*, Regional Studies, Vol. 42, no. 1/2008, s. 51-67.
- Rodríguez-Pose A.**, *Do Institutions Matter for Regional Development?*, Regional Studies, Vol. 47, no. 7/2013, s. 1034-1047
- Romer P.**, *Increasing returns and long-run growth*, Journal of Political Economy, Vol. 94, no. 5/1986, s. 1002-1037
- Romer P.**, *Endogenous technological change*, „Journal of Political Economy”, Vol. 98, no. 5, part II/1990, s. 71-102
- Romer P.**, *The Origins of Endogenous Growth*, „The Journal of Economic Perspective”s, Vol 8 no. (1)/1994
- Tarajkowski J., Wojtasiewicz L.** (red.), *Przestrzeń w polityce gospodarczej*, Katedra Polityki Gospodarczej i Planowania Rozwoju, Akademia Ekonomiczna w Poznaniu, Poznań 2008

- Tarkowski M.**, *Centra i peryferie rozwoju społeczno – gospodarczego Polski w okresie transformacji ustrojowej*, Wydawnictwo Bernardinum, Gdynia-Pelplin 2007
- Todd D.**, *An appraisal of the development pole concept in regional analysis*, „Environment and Planning”, A, Vol. 6, no. 3/1974, s. 291-306
- Tryon R. C.**, *Cluster Analysis*, Edwards Bros, Oxford 1939
- Sixth Report on Economic, Social and Territorial Cohesion, Investment for jobs and growth: Promoting development and good governance in EU regions and cities*, Regional and Urban Policy, European Commission, Brussels, July 2014
- Smętkowski M.**, *Rozwój regionów i polityka regionalna w krajach Europy Środkowo-Wschodniej w okresie transformacji i globalizacji*, Wydawnictwo Naukowe Scholar, Warszawa 2013
- Smith D.**, *Geography and social indicators*, South African Geographical Journal, no. 54/1972, s. 43-57
- Swianiewicz P.**, *Środki unijne w samorządach – kto korzysta najwięcej*, „Samorząd Terytorialny”, nr 5/2012, s. 9-24
- Szlachta J.**, *Polska ścieżka rozwoju regionalnego w poszerzonej Unii Europejskiej*, [w:] *Region Bałtycki w nowej Europie* (red.) T. Parteka, J. Szlachta, W. Szydarowski, „Biuletyn KPZK PAN”, z. 217/2015, Warszawa, s. 7-42
- Wojarska M., Zabielska I.**, *Samorząd terytorialny jako beneficjent funduszy unijnych (na przykładzie gmin województwa warmińsko-mazurskiego)*, Prace Naukowe Uniwersytetu Ekonomicznego we Wrocławiu, nr 380/2015, s. 266-275
- Wojnicka E., Tarkowski M., Klimczak P.**, *Przestrzenne i regionalne zróżnicowanie ośrodków wzrostu. Polaryzacja a wyrównywanie szans rozwojowych. Przesłanki dla kształtowania polityki regionalnej państwa*, Ekspertyza dla Ministerstwa Gospodarki i Pracy, Gdynia – Rzeszów 2005
- Wójcik P.**, *Dywergencja czy konwergencja: dynamika rozwoju polskich regionów*, Studia Regionalne i Lokalne, Vol. (2) no. 32/2008, s. 41-60
- Zaucha J., Komornicki T.**, *The Concept of Territorial Cohesion as Understood and Practiced by Polish Regions*, Working Papers 2/2015, Institute for Development, Sopot
- Zaucha i inni**, *Terytorialny wymiar wzrostu i rozwoju*, Wydawnictwo DIFIN, Warszawa 2015



### **PAWEŁ CHURSKI**

Absolwent Wydziału Nauk Geograficznych i Geologicznych Uniwersytetu Adama Mickiewicza w Poznaniu. Doktor habilitowany nauk o Ziemi w zakresie geografii o specjalnościach: geografia społeczno-ekonomiczna, gospodarka przestrzenna, rozwój regionalny, polityka regionalna. Po studiach podjął pracę w Zakładzie Analizy Regionalnej Instytutu Geografii Społeczno-Ekonomicznej i Gospodarki Przestrzennej UAM, którym obecnie kieruje. Zainteresowania badawcze: problemy rozwoju lokalnego i regionalnego Polski, polityka regionalna Polski i Unii Europejskiej, stan i zmiany rynku pracy w okresie transformacji społeczno-gospodarczej w Polsce i w Europie, poziom i warunki życia, zastosowania geograficznych systemów informacji oraz metod ekonomicznych do analiz przestrzennych. Publikacje i badania naukowe: 8 monografii, 19 redakcji publikacji zwartych, ponad 100 artykułów naukowych publikowanych zarówno w Polsce, jak i za granicą. Uczestnik wielu konferencji krajowych i zagranicznych. Kierownik oraz członek zespołów badawczych, w tym projektów finansowanych przez Ministerstwo Nauki i Szkolnictwa Wyższego, Narodowe Centrum Nauki i Narodowe Centrum Badań i Rozwoju. Ekspert Ministerstwa Rozwoju Regionalnego oraz administracji samorządowej różnych szczebli. Wiceprzewodniczący Komitetu Przestrzennego Zagospodarowania Kraju Polskiej Akademii Nauk.

# CZY BADANIA EWALUACYJNE MOGĄ BYĆ BARDZIEJ NAUKOWE A NAUKA BARDZIEJ UŻYTECZNA DLA BADAŃ EWALUACYJNYCH?

## KORNELIUSZ PYLAK

### Wstęp

Czy badania ewaluacyjne są pracami naukowymi? Czy badania ewaluacyjne w pełni obejmują przygotowanie i realizację procesu badawczego? Czy tworzą i rozwijają teorię badań ewaluacyjnych? Czy badania ewaluacyjne wykorzystują możliwości, jakie daje podejście naukowe? Czy podejście naukowe może przynieść korzyści badaniom ewaluacyjnym, a przede wszystkim umożliwić uzyskanie wyników bardziej wiarygodnych i użytecznych dla odbiorców?

Wiele lat doświadczeń w prowadzeniu badań ewaluacyjnych, ogromny dorobek publikacyjny rynku ewaluacji w Polsce, a także pierwsze refleksje dotyczące „naukowości” i przydatności tych badań do rozwoju systemu wdrażania Europejskich Funduszy Strukturalnych i Inwestycyjnych (które *de facto* wymusiły stosowanie ewaluacji w codziennej praktyce instytucji) pojawiające się w publikacjach naukowych (poza raportami ewaluacyjnymi) – leżały u podstaw zainteresowania się tą tematyką. K. Olejniczak wskazuje na bardzo ubogi zakres i metodykę badań ewaluacyjnych, których trzy główne słabości można określić jako:<sup>1</sup>

1. Brak spójnej, osadzonej w literaturze tematu całościowej koncepcji badawczej.
2. Ograniczanie pakietu metod i technik badawczych do kilku podstawowych elementów.
3. Brak analizy relacji przyczynowo-skutkowych efektów interwencji.

Z reguły koncepcja badawcza prezentowana w raportach odzwierciedla koncepcję wynikającą z opisów zamówienia, chociaż często wykonawcy są proszeni o przedstawienie w ofertach swojej oryginalnej koncepcji badania. W wielu przypadkach sprowadza się ona jednak do uzyskania odpowiedzi na postawione przez Zamawiającego pytania badawcze. W ramach oceny ofert wykonawcy są często proszeni o dodanie dodatkowych pytań badawczych i technik badawczych, co sprowadza się do zaproponowania takich pytań i technik, które nie będą stwarzały dodatkowych problemów podczas badania i nie będą zwiększały budżetu badania. Nie są proponowane natomiast problemy i techniki je rozwiązujące, które mogą wnieść coś nowego do wyników badania – co bierze się z prostego faktu, że na etapie składania oferty żaden z wykonawców nie zna dogłębnie przedmiotu badania, chyba, że zdążył zdobyć takie doświadczenie przed przystąpieniem do pisania oferty.

---

1 Olejniczak K. (red.), *Praktyka ewaluacji efektów programów rozwoju regionalnego – studium porównawcze*, Raporty i Analizy EUROREG 2/2009, s. 2. Pobrano ze strony: [http://www.euroreg.uw.edu.pl/dane/web\\_euroreg\\_publications\\_files/611/olejniczak\\_2009\\_praktyka\\_ewaluacji\\_efektow.pdf](http://www.euroreg.uw.edu.pl/dane/web_euroreg_publications_files/611/olejniczak_2009_praktyka_ewaluacji_efektow.pdf).

Takie podejście do przygotowywania koncepcji powoduje duże zagrożenie dla jakości prowadzonego później badania. Siłą rzeczy bowiem metody, techniki i narzędzia badawcze muszą być znane w momencie pisania oferty i podpisywania umowy na realizację badania, a więc dużo wcześniej niż analiza literaturowa. Brak analizy literatury na etapie tworzenia koncepcji badania może powodować, że nie jest ona twórcza i nie zawsze prowadzi do pełnego rozwiązania problemów badawczych. To analiza literatury może wskazać na pewne mechanizmy powstawania efektów lub jakichś zjawisk, które warto byłoby w badaniu zweryfikować. Bez analizy literatury, badanie może doprowadzić jedynie do ponownego odkrycia znanego w literaturze mechanizmu, który może wówczas zostać opisany, ale nie wyjaśniony (bo narzędzia badawcze zostały zaprojektowane „na ślepo” i siłą rzeczy nie mogą zbierać informacji o czynnikach sprawczych (zmiennych wyjaśniających dane zjawisko), tylko dowiadywać się o ich istnieniu). Dlatego, jak wskazuje K. Olejniczak, wyniki badań ewaluacyjnych sprowadzają się do opisu zaplanowanych w koncepcji produktów badania, a nie na odkryciu i opisanu rzeczywistych zmian, wywołanych programami i ewolucji mechanizmów powodujących te zmiany – nie spełniają zatem podstawowej funkcji, dla jakiej zostały powołane do życia – służeniu decyden- tom i budowaniu ich wiedzy o mechanizmach rozwoju regionów i kraju<sup>2</sup>. J. Szlachta również wskazuje na techniczny i odtwórczy charakter badań ewaluacyjnych, bazujący na wytycznych europejskich i polskich<sup>3</sup> (w tym kontekście opisów przedmiotu zamówienia). Jako jedną z przyczyn wskazuje się stosowanie podejścia konsultacyjnego, doradczego, zamiast podejścia naukowego<sup>4</sup>.

Dlatego **celem badawczym** niniejszego artykułu jest *ocena stosowanego w badaniach ewaluacyjnych podejścia w kontekście możliwości implementacji podejścia naukowego*. Podejście naukowe ma przede wszystkim wymiar koncepcyjny tzn. tworzy obraz całego systemu i mechanizmów w nim występujących, a więc tworzy pewien abstrakcyjny model danego systemu, który w toku badania właściwego zostanie zweryfikowany pozytywnie lub negatywnie. W każdym przypadku powstanie potwierdzony w badaniu model rzeczywisty obejmujący związki przyczynowo-skutkowe między zmiennymi opisującymi ten model.

Artykuł ma następującą strukturę. Po określeniu przyczyn podjęcia problemu badawczego i postawieniu celu badania, w kolejnym rozdziale przedstawiono analizę literatury dotyczącą prowadzenia procesu naukowego i na tej bazie wyprowadzono hipotezy badawcze. Dodatkowo określono przedmiot i zakres badania mający na celu weryfikację postawionych hipotez. W na-

2 Olejniczak K. (red.), *Praktyka ewaluacji...*, op. cit., s. 80.

3 Szlachta, J., *Polskie doświadczenia w zakresie ewaluacji*, II Konferencja Ewaluacyjna – Ewaluacja ex-ante jako narzędzie doskonalenia zarządzania strategicznego w administracji publicznej, Ministerstwo Rozwoju Regionalnego & Polska Agencja Rozwoju Przedsiębiorczości, Warszawa 2006.

4 Olejniczak K., *The shortcomings of ERDF evaluation practice in the Member States [w:] Evaluating the effects of regional interventions. A look beyond current Structural Funds' practice*, (red.) K. Olejniczak, M. Kozak, S. Bienias, Ministry of Regional Development, Warsaw 2011, s. 133–134.



stępnym rozdziale zaprezentowano wyniki badania i dyskusję. Wnioski i zalecenia wynikające z przeprowadzonych badań przedstawiono w zakończeniu.

## Metodyka badania

Cel badania dotyczy prowadzenia procesu naukowego. Proces naukowy rozpoczyna się od uświadomienia sobie braku wiedzy odnośnie pewnych zjawisk nurtujących badacza<sup>5</sup>. Wówczas należy określić **przedmiot badania**, a więc podmioty, w tym jednostki i instytucje, procesy i zjawiska zachodzące między nimi, tworzące pewną rzeczywistość<sup>6</sup>. Znajomość przedmiotu badania jest niezbędna do odpowiedniego zdefiniowania **problemów badawczych**. Problemy badawcze, określane w formie pytań<sup>7</sup>, muszą wyczerpać zakres naszej niewiedzy, a więc określają obszar badawczych poszukiwań. W tym kontekście problemy powinny już definiować zależności między zmiennymi opisującymi przedmiot badania<sup>8</sup>. Na przykład błędnie określony problem badawczy: *Jakie są najważniejsze przyczyny nie podejmowania próby aplikowania o środki z funduszy strukturalnych?* powinien zostać przedstawiony w postaci zależności dwóch zmiennych: *W jakim stopniu samorządy zaspokajały swoje potrzeby za pomocą własnych środków lub innych instrumentów finansowych?* (są tu dwie zmienne: poziom wykorzystania własnych środków w finansowaniu inwestycji i poziom aplikowania o środki z funduszy strukturalnych), albo kolejne przykłady dobrze określonych problemów badawczych: *Na ile samorządy, które nie złożyły ani jednego wniosku o dofinansowanie rozważały taką możliwość i dlaczego z niej zrezygnowały?*<sup>9</sup> Takie postawienie problemu badawczego wymaga już wiedzy odnośnie potencjalnych przyczyn nieaplikowania, ale dzięki temu dowiemy się już o skali zjawiska – będzie można na przykład przeanalizować i porównać poziom środków własnych na inwestycje u podmiotów, które nie aplikowały i u tych, które realizowały inwestycje ze środków unijnych.

W dalszej kolejności należy sformułować **hipotezy robocze**. Hipotezy są konieczne, aby właściwie zaplanować metody i narzędzia badawcze oraz poprawnie ukierunkować cały proces badawczy. Hipoteza jest twierdzeniem, domysłem częściowo tylko uzasadnionym na podstawie analizy literatury, za pomocą którego należy wyjaśniać prawdziwe zjawiska<sup>10</sup> zachodzące w rzeczywistości określonej przez przedmiot badania. Kluczowe jest, aby podczas tworzenia hipotez nie wykorzystywać wyłącznie własnych domysłów<sup>11</sup>. Hipotezy są niezwykle pomocne przy roz-

5 Pilch, T., *Zasady badań pedagogicznych*. Zakład Narodowy im. Ossolińskich, Wrocław 1977, s. 101.

6 Sztumski J., *Wstęp do metodologii i techniki badań społecznych*, Wyd. Śląsk, Katowice 1995, s. 19–20.

7 Łobocki M., *Wprowadzenie do metodologii badań pedagogicznych*, Wyd. Impuls, Kraków 2007, s. 110.

8 Pilch, T., *Zasady badań...* op. cit., s. 24–25.

9 Grupa Gomułka, Raport końcowy z badania ewaluacyjnego: „*Analiza przyczyn braku aktywności lub niskiej aktywności niektórych samorządów gminnych w sięganiu po wsparcie z Funduszy Unijnych*”, na zlecenie Ministerstwa Rozwoju Regionalnego, Warszawa 2009, s. 8. Pobrano ze strony: [http://www.archiwum.ewaluacja.gov.pl/Wyniki/Documents/rrit\\_062.pdf](http://www.archiwum.ewaluacja.gov.pl/Wyniki/Documents/rrit_062.pdf).

10 Kotarbiński, T., *Kurs logiki dla prawników*, Państwowe Wydawnictwa Naukowe, Warszawa 1960, s. 181.

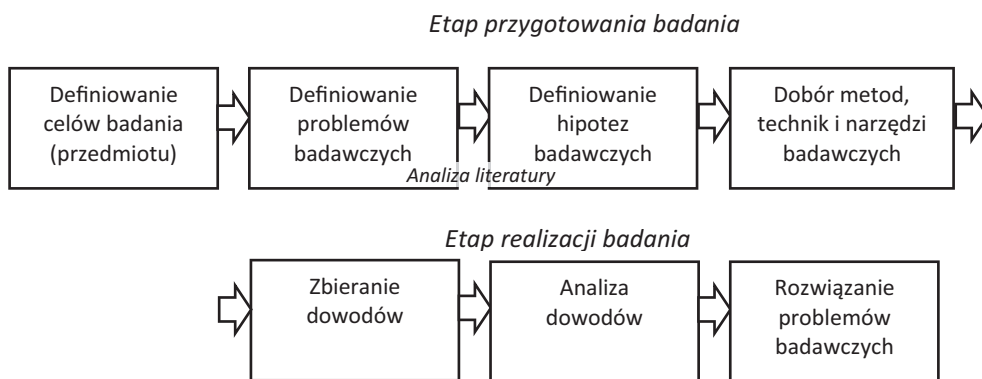
11 Adamkiewicz-Drwiłło H.G., *Współczesna metodologia nauk ekonomicznych*, Dom Organizatora TNOiK, Toruń 2008, s. 71.



wiązywaniu problemów badawczych, bowiem wymagają wytypowania zmiennych zależnych i niezależnych opisujących dany problem badawczy (co już częściowo należy zrobić podczas formułowania problemów badawczych), a następnie określenia wszelkich zależności, które mogą pojawić się między zmiennymi. Zależności te powinny być już zweryfikowane przez uznane prace badawcze, ale ich kontekst, charakter, wreszcie czas – mogą nie sprawdzić się w nowej sytuacji badawczej. Co więcej, hipotezy nie mogą tłumaczyć wszystkich zjawisk zachodzących w danym środowisku za pomocą wcześniejszych teorii – przynajmniej jedno zjawisko powinno być dotąd niewyjaśnione<sup>12</sup>, w przeciwnym wypadku sens prowadzenia ponownego procesu badawczego jest bliski zeru. Ma szczególne znaczenie w raportach ewaluacyjnych o podobnym przedmiocie badania. Generalnie, **każda hipoteza podlega weryfikacji** i może zostać **udowodniona** poprzez zebranie danych popierających zakładaną zależność lub **obalona** w momencie niemożności zebrania takich danych lub uzyskania danych świadczących o fałszywości założenia<sup>13</sup>. Najczęściej hipotezy są budowane w oparciu o precyzyjne wskazanie zależności między dwiema zmiennymi.

Dopiero po zdefiniowaniu hipotez badawczych istnieje możliwość dopasowania odpowiednich **metod i technik badawczych**, ponieważ dopiero wtedy wiadomo, jakie cechy u jakich podmiotów należy badać i w jaki sposób znaleźć i oszacować związki między nimi. Na tym kończy się **etap przygotowania** badania i pozostaje **etap realizacji**, na który składa się zbieranie dowodów, analiza dowodów i rozwiązanie problemów badawczych poprzez weryfikację hipotez. Cały proces naukowy można zatem przedstawić w sposób przedstawiony na rysunku 1:

**Rysunek 1. Przebieg procesu badawczego**



Źródło: Opracowanie własne.

<sup>12</sup> Ibidem, s.71.

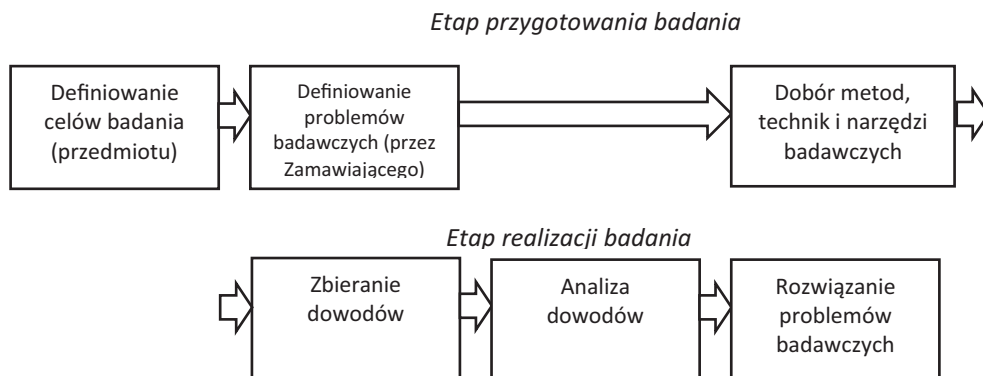
<sup>13</sup> Pilch, T., *Zasady badań...*, op. cit., s. 27.

Szczegółowo warto przyjrzeć się pierwszemu etapowi, w którym można odnaleźć kluczowy problem badawczy. Metody, techniki i narzędzia badawcze są określone i zatwierdzone przed weryfikacją i poprawnym określeniem problemów badawczych i postawieniu hipotez. Może się zatem okazać, że po postawieniu hipotez (jeżeli do tego dojdzie) zaplanowane metody i techniki nie doprowadzą do ich weryfikacji. Taka sytuacja jest wymuszona przez proces zamawiania badań ewaluacyjnych, które muszą być kontraktowane na określoną kwotę obejmującą realizację określonej koncepcji badawczej. Można zatem na bazie powyższych rozważań postawić następującą hipotezę:

### H.1. W badaniach ewaluacyjnych brakuje etapu formułowania hipotez badawczych wynikających z przeprowadzonej analizy literaturowej.

Jeżeli hipoteza zostanie potwierdzona, oznaczać to będzie, że następuje skrócenie procesu o bardzo ważne etapy badania (częściowo definiowanie problemów badawczych oraz formułowanie hipotez), przez co dobór metod, technik i narzędzi badawczych odbywa się na bazie niepełnych informacji o przedmiocie badania:

**Rysunek 2. Hipotetyczny przebieg badań ewaluacyjnych**



Źródło: Opracowanie własne.

Hipoteza zostanie potwierdzona, jeżeli większość badań nie formułuje hipotez badawczych lub większość badań prowadzi w niewłaściwy sposób proces ich formułowania. Dlatego, aby zweryfikować hipotezę przeanalizowane zostaną następujące elementy procesu przygotowania badania ewaluacyjnego (wynikające z analizy literatury przedstawionej na początku tego rozdziału):

1. Obecność analizy literatury, z której wyłaniają się mechanizmy funkcjonowania opisywanej w badaniu rzeczywistości.
2. Wskazanie czynników zależnych i niezależnych (wyjaśniających i wyjaśnianych) z literatury / badań wstępnych.

3. Wydzielenie i sformułowanie hipotez w tekście.
4. Przedstawienie sposobów weryfikacji hipotez.
5. Przygotowanie narzędzi badawczych w oparciu o hipotezy badawcze (w sposób niezbędny do ich weryfikacji).
6. Przeprowadzenie weryfikacji hipotez – wskazanie, które hipotezy zostały zweryfikowane pozytywnie, które negatywnie.

Mimo, że teoretycznie wykonawcy mogliby na własną rękę weryfikować i zmieniać problemy badawcze, a także formułować hipotezy robocze, trudno jest się spodziewać realizacji dodatkowych obowiązków, wymagających dodatkowych nakładów pracy (analiza literaturowa jest wbrew pozorom kosztowna i długotrwała), jak również trudno jest wyobrazić sobie negocjacje nad zmianą koncepcji badania (metod i technik, populacji badanych itd.) po podpisaniu umowy. Stąd można przewidywać, że powyższa hipoteza zostanie zweryfikowana pozytywnie. Jeżeli tak się stanie, to pozytywnie zweryfikowana zostanie również poniższa **hipoteza główna H.2:**

**H.2. Większość badań ewaluacyjnych nie prowadzi właściwego (zdefiniowanego wyżej) procesu naukowego.**

Wynika to z faktu, że etap tworzenia problemów i hipotez badawczych jest niezwykle ważnym elementem procesu naukowego, dzięki któremu powstają nowe teorie i twierdzenia dotyczące interwencji publicznej. Bez tego elementu, precyzyjnie rzecz ujmując – bez elementu nowości – badanie ewaluacyjne przyjmuje charakter opisowego procesu monitorowania i sprawozdawczości.

W celu weryfikacji postawionych hipotez przeprowadzono badanie 100 raportów z badań ewaluacyjnych. Z uwagi na przedmiot badania wzięto pod uwagę wyłącznie badania on-going oraz ex-post. Do badania wykorzystano bazę badań ewaluacyjnych znajdującą się na stronie [www.archiwum.ewaluacja.gov.pl](http://www.archiwum.ewaluacja.gov.pl). Raporty dobrano do badania poprzez losowanie.

### Wyniki badania i dyskusja

W niniejszym rozdziale przedstawiono wyniki badania wybranych raportów. Zaprezentowano przykłady badań, w których proces badawczy przeprowadzono poprawnie lub w których sformułowano hipotezy, ale zaobserwowano pewne mankamenty. Nie przedstawiono z nazwy badań, w których zabrakło w ogóle formułowania hipotez badawczych.

Analiza raportów z badań ewaluacyjnych wykazała, że w **czterech na pięć badań (78%) nie prowadzi się w ogóle procesu formułowania hipotez badawczych** i nie przeprowadza się ich weryfikacji, a więc nie zanotowano w nich żadnego z 6 wskazanych w poprzednim rozdziale elementów. Natomiast w przypadku badań ewaluacyjnych, które formułują hipotezy (22 raporty), **proces ten jest z reguły niepełny lub nawet szczątkowy**, co przedstawia Tabela 1.

**Tabela 1. Wyniki badania raportów końcowych – wystąpienie elementów procesu formułowania i weryfikacji hipotez**

	Analiza literatury	Wskazanie czynników z literatury / badań wstępnych	Wydzielenie hipotez	Sposoby weryfikacji	Przygotowanie narzędzi badawczych w oparciu o hipotezy	Przeprowadzenie weryfikacji
Brak elementu w raporcie	11	13	16	11	11	13
Element wystąpił w raporcie	11	9	6	11	11	9

Źródło: Opracowanie własne na podstawie badań własnych.

Można wskazać cztery badania ewaluacyjne, w których **przeprowadzono pełny proces badawczy**. Przykładem badania ewaluacyjnego, w którym w pełni wykorzystano podejście naukowe jest *analiza przyczyn braku aktywności niektórych samorządów województwa zachodniopomorskiego*<sup>14</sup>. Zdefiniowano w nim pięć obszarów problemowych, które grupują czynniki – zmienne wyjaśniające aktywność samorządów w ubieganiu się o wsparcie z programów. Przeprowadzono również badanie jakościowe (studia przypadku), w ramach których określono atuty, których brak może zmniejszać szanse niektórych samorządów gminnych na skuteczne sięganie po wsparcie z funduszy Unii Europejskiej. W trakcie badania zweryfikowano hipotezy o wpływie tych czynników na aktywność oraz skuteczność ubiegania się przez samorzady o wsparcie unijne w ramach programu, a także na ocenę wpływu poszczególnych czynników. Wykorzystano w tym celu badanie telefoniczne, dzięki któremu uzyskano reprezentatywne wyniki niezbędne do weryfikacji hipotez zidentyfikowanych podczas wcześniejszych (jakościowych) etapów ewaluacji.

Innym przykładem badania z poprawnie przeprowadzonym procesem badawczym jest *badanie wpływu spójności wewnętrznej na konkurencyjność regionu łódzkiego*<sup>15</sup>. Hipotezy zostały sformułowane w tekście (nie zostały wydzielone, co jest jedynym – drobnym mankamentem tego raportu) na bazie rozległej analizy literatury naukowej, w tym światowej (m.in. D. Ricardo, Camagni, Krugman), warsztat badawczy został rozbudowany, a weryfikacja hipotez została przeprowadzona w oparciu o zaawansowane techniki statystyczne i ekono-

14 Opinia – agencja badania rynku & INFOR Training, *Analiza przyczyn braku aktywności niektórych samorządów w sięganiu po wsparcie unijne w ramach programu ZPORR celem ujednoczenia Planu Komunikacji RPO WZ na lata 2007–2013 oraz Planu Działań Informacyjnych i Promocyjnych na rok 2009. Raport końcowy*, na zlecenie Urzędu Marszałkowskiego Województwa Zachodniopomorskiego, Warszawa 2008, s. 22–28.

15 Instytut Badań Strukturalnych, *Spójność wewnętrzna a konkurencyjność regionu łódzkiego. Diagnoza strategiczna*, na zlecenie Urzędu Marszałkowskiego Województwa Łódzkiego, Łódź 2009, s. 10–18, 24–52. Pobrano ze strony: [http://www.archiwum.ewaluacja.gov.pl/Wyniki/Documents/1\\_071.pdf](http://www.archiwum.ewaluacja.gov.pl/Wyniki/Documents/1_071.pdf).

metryczne (m.in. ścieżki rozwojowe). Weryfikacja hipotez została odniesiona do literatury naukowej („...formułowana w literaturze przedmiotu hipoteza dopuszczająca występowanie rosnących przychodach skali znajduje swoje potwierdzenie empiryczne w przypadku polskich województw”).

Podobnie, w przypadku badania *oceny wpływu polityki spójności na rozwój miast polskich*<sup>16</sup> również można wskazać na poprawnie przygotowany i przeprowadzony proces naukowy z wyraźnie zaznaczonymi hipotezami. Również w badaniu *oceny systemu realizacji polityki spójności w Polsce w ramach perspektywy 2004–2006*<sup>17</sup> określono na podstawie literatury hipotezę główną („wiodącą funkcją realizacji polityki spójności w Polsce w ramach perspektywy 2004–2006 była funkcja wdrożeniowo-absorpcyjna”) i cztery hipotezy wyjaśniające słabość systemu. Hipotezy znalazły się w oddzielnym rozdziale, podobnie jak i ich weryfikacja („Weryfikacja hipotez badawczych i całościowa ocena systemu”).

Analiza pozostałych raportów końcowych z badań ewaluacyjnych wykazała, że nagminną praktyką jest **umieszczanie spisu literatury, która rzekomo została przeanalizowana** (konieczność takich zapisów wynika ze szczegółowych opisów zamówień, które narzucają obowiązek jej przeprowadzenia). Rzadko można spotkać odwołania do literatury w treści raportów, a szczególnie trudno jest odnaleźć cytowania literatury innej niż akty prawne, w tym rozporządzenia unijne i poprzednie raporty z badań.

Dodatkowo występują badania, które wprawdzie wskazują, że postawiono hipotezy robocze i zrobiono analizę literatury, na podstawie której sformułowano hipotezy, niemniej jednak w dalszej części raportu trudno **odnaleźć nie tylko analizę literatury, ale również same hipotezy**. Przykładem takiego badania jest analiza *determinantów rozwoju współpracy międzynarodowej beneficjentów funduszy strukturalnych na poziomie Narodowej Strategii Spójności*<sup>18</sup>. Podobnie w badaniu<sup>19</sup> dotyczącym *oceny zintegrowanego systemu działań informacyjnych, promocyjnych oraz szkoleniowych w zakresie funduszy europejskich w Województwie Małopolskim* wskazano, że hipotezy były stawiane („...została przedstawiona prezentacja głównych wyników badania, wniosków, hipotez i pro-

16 Centrum Europejskich Studiów Regionalnych i Lokalnych EUROREG Uniwersytetu Warszawskiego, *Ocena wpływu polityki spójności na rozwój miast polskich (w ramach ewaluacji ex post NPR 2004-2006)*, na zlecenie Ministerstwa Rozwoju Regionalnego, Warszawa 2010, s. 52–53. Pobrano ze strony:

[http://www.archiwum.ewaluacja.gov.pl/Ewaluacja\\_ex\\_post\\_NPR/Documents/raport\\_koncowy\\_expost9.pdf](http://www.archiwum.ewaluacja.gov.pl/Ewaluacja_ex_post_NPR/Documents/raport_koncowy_expost9.pdf).

17 EGO s.c., *Ocena systemu realizacji polityki spójności w Polsce w ramach perspektywy 2004–2006*, na zlecenie Ministerstwa Rozwoju Regionalnego, Warszawa 2010, s. 8–13, 192–197. Pobrano ze strony:

[http://www.archiwum.ewaluacja.gov.pl/Ewaluacja\\_ex\\_post\\_NPR/Documents/expost\\_koncowy\\_4.pdf](http://www.archiwum.ewaluacja.gov.pl/Ewaluacja_ex_post_NPR/Documents/expost_koncowy_4.pdf).

18 Stowarzyszenie Doradców Gospodarczych Pro-Akademia, *Determinanty rozwoju współpracy międzynarodowej beneficjentów funduszy strukturalnych na poziomie Narodowej Strategii Spójności*, dla Ministerstwa Rozwoju Regionalnego, Warszawa 2008, s. 6. Pobrano: [http://www.archiwum.ewaluacja.gov.pl/Wyniki/Documents/rrit\\_060.pdf](http://www.archiwum.ewaluacja.gov.pl/Wyniki/Documents/rrit_060.pdf).

19 Agrotec Polska Sp. z o.o., *Raport z badania pt.: „Ocena zintegrowanego systemu działań informacyjnych, promocyjnych oraz szkoleniowych w zakresie funduszy europejskich w Województwie Małopolskim”*, na zlecenie Urzędu Marszałkowskiego Województwa Małopolskiego, Kraków 2009, s. 25. Pobrano ze strony: [http://www.archiwum.ewaluacja.gov.pl/Wyniki/Documents/1\\_081.pdf](http://www.archiwum.ewaluacja.gov.pl/Wyniki/Documents/1_081.pdf).

pozycje rekomendacji”), natomiast w tekście nie sformułowano hipotez, sposobów i procesu ich weryfikacji. W jednym z raportów przedstawiono rozdział zatytułowany „*Cele i hipotezy*”, w którym nie pojawiły się hipotezy (w kolejnych rozdziałach przedstawiono jedynie sytuację społeczno-gospodarczą badanego obszaru), a tym samym nie są one w dalszej części weryfikowane.

W badaniach, w których wystąpiły hipotezy, z reguły **nie były one wyraźnie zaprezentowane** (np. poprzez oddzielenie w tekście i ewentualne ponumerowanie), nie określano również **sposobów ich weryfikacji** (kiedy uznajemy, że dana hipoteza będzie zweryfikowana pozytywnie, kiedy negatywnie).

W badaniu *wpływu interwencji finansowanych z funduszy strukturalnych UE na przekształcenie obszarów wiejskich*<sup>20</sup> wykorzystano analizę studiów przypadku oraz 26 wywiadów pogłębionych do określenia hipotez dotyczących potencjalnych obszarów najsilniejszego oddziaływania. W raporcie napisano, że sformułowano kilka istotnych hipotez roboczych, które po zweryfikowaniu przedstawiono w postaci wniosków i rekomendacji. Nie wskazano jednak, jakie to są hipotezy, nie napisano, w jaki sposób zostaną one zweryfikowane, a także czy je zweryfikowano pozytywnie, czy negatywnie.

W przeciwieństwie do powyższego przypadku badania, bardzo często w raportach hipotezy lub jak je czasami autorzy określają – hipotetyczne przyczyny danego zjawiska – są **określane na podstawie przypuszczeń** wykonawcy lub zamawiającego i wynikają wyłącznie z doświadczenia i wiedzy osób przygotowujących lub realizujących badanie. Dla przykładu, w badaniu dotyczącym *metodologii badań dla działań informacyjnych i promocyjnych* określono hipotezy dotyczące nierównomiernej dystrybucji informacji w województwie, wskazano, w jaki sposób hipoteza będzie weryfikowana, niemniej czynniki wpływające na dystrybucję informacji wskazano intuicyjnie za Zamawiającym badanie (nie przeprowadzono analizy literatury w tym zakresie)<sup>21</sup>. Podobnie, nie przeprowadzono również analizy literatury niezbędnej do postawienia hipotezy dotyczącej odpływu ludności w badaniu *wpływu polityki spójności na stan zróżnicowań i rozwój polskich regionów*<sup>22</sup>. Hipotezy bazowały w tym przypadku na domysłach i prawdopodobnie doświadczeniu autorów raportu, o czym świadczy ten fragment raportu: „*W zasadzie należałoby się spodziewać, że wykorzystywanie środków funduszy unijnych w danym rejonie (powiecie), pobudzając rozwój lokalny, poprawiając infrastrukturę i sytuację na rynku pracy,*

20 Agrotec Polska Sp. z o.o., *Raport z badania pt.: „Wpływ interwencji finansowanych z funduszy strukturalnych UE na przekształcenie obszarów wiejskich”*, na zlecenie Ministerstwa Rozwoju Regionalnego, Warszawa 2007, s. 107, 215, 221. Pobrano ze strony: [http://www.archiwum.ewaluacja.gov.pl/Wyniki/Documents/rrit\\_001.pdf](http://www.archiwum.ewaluacja.gov.pl/Wyniki/Documents/rrit_001.pdf).

21 Proksen Sp. z o.o. & Pracownia Badań Społecznych i Marketingowych „Soma”, *Metodologia badań dla działań informacyjnych i promocyjnych realizowanych ze środków Unii Europejskiej w Województwie Kujawsko-Pomorskim*, na zlecenie Urzędu Marszałkowskiego Województwa Kujawsko-Pomorskiego, Toruń 2008, s. 25, 31, 40. Pobrano ze strony: [http://www.archiwum.ewaluacja.gov.pl/Wyniki/Documents/rrit\\_045.pdf](http://www.archiwum.ewaluacja.gov.pl/Wyniki/Documents/rrit_045.pdf).

22 CASE (Centrum Analiz Społeczno-Ekonomicznych), *Badanie ewaluacyjne nt. „Wpływ polityki spójności na stan zróżnicowań i rozwój polskich regionów”*. Raport końcowy, na zlecenie Ministerstwa Rozwoju Regionalnego, Warszawa 2008, s. 81. Pobrano ze strony: [http://www.archiwum.ewaluacja.gov.pl/Wyniki/Documents/rrit\\_051.pdf](http://www.archiwum.ewaluacja.gov.pl/Wyniki/Documents/rrit_051.pdf).

przyciąga nowych mieszkańców”. W innym badaniu<sup>23</sup> wykonawca przyjął kilka hipotetycznych przyczyn braku aktywności gmin w pozyskiwaniu środków z funduszy strukturalnych w okresie finansowania 2004–2006, również nie korzystając w tym zakresie z literatury przedmiotu.

Podsumowując można powiedzieć, że **obie hipotezy badawcze zostały potwierdzone**. W zdecydowanej większości badań (78%) nie jest prowadzony pełny proces badawczy – nie są formułowane hipotezy badawcze, a w przypadku tych badań, gdzie starano się stworzyć hipotezy – był to proces niepełny lub czasami szczątkowy. W przypadku połowy z tych raportów odnaleziono analizę literatury, przy czym była ona najczęściej oparta na dokumentach programowych, a rzadko – nawet na jednej lub kilku pozycjach naukowych. W połowie przypadków nie przedstawiono sposobów weryfikacji hipotez i nie przygotowywano narzędzi badawczych w oparciu o sformułowane hipotezy (najczęściej, ponieważ tego sformułowania nie było). W większości raportów nie wskazano czynników na bazie literatury lub badań wstępnych – czynniki były najczęściej dobierane na bazie własnych domysłów lub wskazań zamawiającego w opisach zamówienia. Jednak w największej liczbie raportów nie odnaleziono wydzielonych hipotez (precyzyjnego sformułowania zależności pomiędzy dwoma lub większą liczbą zmiennych) – por. Tabela 1. Jak wskazano na początku, jedynie w czterech przypadkach odnaleziono wszystkie elementy procesu badawczego dotyczącego formułowania i weryfikacji hipotez.

## Zakończenie

Badania ewaluacyjne powinny bazować na prowadzeniu rzeczywistego procesu badawczego, popartego pogłębioną analizą literatury nie tylko branżowej (aktów prawnych i raportów z badań), ale przede wszystkim literatury naukowej z danej dziedziny (polskiej i zagranicznej). Dzięki temu z biegiem czasu same raporty z badań ewaluacyjnych stałyby się pełnoprawnymi publikacjami naukowymi i mogłyby stanowić dorobek tworzącej się w Polsce dyscypliny naukowej. Jest to podejście zgodne z podejściem „ewaluacji wspieranej teorią” (ang. *theory-driven evaluations*)<sup>24</sup>, które popularyzuje w Polsce K. Olejniczak<sup>25</sup>. Pozwala ona przedstawić założenia programu na temat mechanizmu oczekiwanej zmiany społecznej, stworzyć solidną podstawę koncepcyjną badania oraz zbudować spójną narrację raportu końcowego. Jest realizowana w następujących krokach: opis kontekstu, odkrycie potencjalnych mechanizmów i zmiennych je opisujących, rekonstrukcja zakładanej zmiany, weryfikacja rzeczywistych zmian, wyjaśnienie zaobserwowanych zależności<sup>26</sup>. Zatem w sensie konceptualnym, jest to zastosowanie podejścia

23 Grupa Gomułka, *Raport końcowy...*, op. cit., s. 31–32.

24 Chen, H.-T., Rossi, P.H., *Evaluating with sense: The theory-driven approach*, Evaluation Review, No. 7/1983, s. 283–302.

25 Olejniczak K. (red.), *Praktyka ewaluacji...*, op. cit., s. 82.

26 Olejniczak K., *Tracing theories behind practice – a theory-driven approach* [w:] *Evaluating the effects of regional interventions. A look beyond current Structural Funds' practice*, (red.) K. Olejniczak M., Kozak, S. Bienias, Ministry of Regional Development, Warsaw 2011, s. 204.



naukowego w ewaluacji. Opiera się bowiem na hipotezach, które charakteryzują mechanizmy zachodzące w modelu opisującym określony wycinek rzeczywistości. Poniżej omówione zostały najważniejsze wnioski i zalecenia z przeprowadzonych badań, które pozwolą na zastosowanie podejścia naukowego w badaniach ewaluacyjnych.

Pierwszą kwestią jest analiza literatury. Mając na względzie poprawność prowadzonego procesu naukowego, trafność stawianych hipotez i poprawność ich weryfikacji, zamawiający powinni w większym stopniu wymagać od wykonawców analizy literatury nie tylko tej, którą zawarli w opisach zamówień, ale również naukowej, w której można znaleźć wyniki przydatnych badań zawierających na przykład czynniki wpływające na proces (zmiennie objaśniające), albo związki przyczynowo-skutkowe, mogące być podstawą formułowanych hipotez.

Analiza literatury powinna być prowadzona w określonym celu. Powinna ona doprowadzić do wskazania zmiennych opisujących problem badawczy (np. niski poziom aktywności beneficjentów w staraniu się o fundusze) i zależności między nimi. Na ich bazie należy sformułować hipotezy badawcze, które w największym stopniu powinny być oryginalne (niekoniecznie w sensie twórczym, wystarczy, aby były odkrywcze w kontekście czasu, miejsca lub przedmiotu badania). Kolejnym elementem jest przedstawienie sposobów weryfikacji hipotez – określenia, kiedy hipoteza będzie zweryfikowana pozytywnie, kiedy negatywnie. W oparciu o hipotezy należy przygotować metody, techniki i narzędzia badawcze tak, aby zdobyć informacje o wszystkich zawartych w nich zmiennych i potencjalnych zależnościach między nimi. W końcowej części (części głównej raportu) powinno się pojawić odniesienie do hipotez i ich weryfikacja oraz określenie, czy zostały one zweryfikowane pozytywnie, czy negatywnie.

Jak wskazano w poprzednim rozdziale, główną przyczyną mankamentów w prowadzeniu procesu badawczego jest sposób organizacji badań ewaluacyjnych – konieczność określania koncepcji badania i metod oraz technik badawczych wraz z wielkościami prób przed rozpoczęciem całego procesu. Wynika to ze specyfiki zamówień publicznych. Dlatego do rozwiązania tego problemu można zastosować trzy podejścia. Po pierwsze, zamawiający mogą przygotować odrębne badanie na przygotowanie szczegółowych opisów przedmiotów zamówień dla badań zaplanowanych na dany okres, w których zawarta byłaby analiza literatury wraz z postawionymi hipotezami. Wówczas w ofertach na wykonanie konkretnych badań potencjalni wykonawcy prezentowaliby swoją koncepcję weryfikacji już przygotowanych hipotez badawczych. Drugie rozwiązanie polegałoby na wypracowaniu wraz z kilkoma oferentami hipotez badawczych i przygotowania z ich udziałem szczegółowego opisu przedmiotu zamówienia. Można do tego wykorzystać tryb zamówienia negocjacji z ogłoszeniem lub bez ogłoszenia. Słabą stroną tego rozwiązania jest długi czas realizacji badania. Trzecie rozwiązanie polega na zastosowaniu w opisie zamówienia klauzuli, że pewne elementy, jak pytania (problemy) badawcze, narzędzia mogą ulec zmianie w momencie dopracowania przez wykonawcę koncepcji badania. Wadą tego rozwiązania jest mała elastyczność ewentualnych zmian (nie jest możliwe na przykład



zmniejszenie lub zwiększenie próby badawczej, dodanie kolejnych technik generujących koszty, bowiem zmiany muszą być dokonywane w obrębie zatwierzonego już budżetu. Dlatego to rozwiązanie jest skuteczne w przypadku dobrze dopracowanych opisów zamówienia, które wymagają jedynie niewielkiej korekty.

Dzięki wprowadzonym zmianom do procesu przygotowania i prowadzenia badań ewaluacyjnych można dojść do wniosku, że ewaluacja programów rozwoju regionalnego mogłaby stać się odrębną dyscypliną naukową w Polsce. Według Lilesa, Johnsona i Meade dyscypliny naukowe powinny spełniać sześć cech, którymi są: 1) posiadanie wiodącego problemu badawczego, na który dyscyplina próbuje znaleźć odpowiedź; 2) specyficzny sposób postrzegania świata – paradygmat stanowiący zbiór pojęć i teorii tworzących podstawy danej dyscypliny; 3) posiadanie dyscyplin referencyjnych, które ustanawiają daną dyscyplinę; 4) posiadanie zasad i praktyk powiązanych z daną dyscypliną; 5) posiadanie aktualnej agendy badawczej; 6) rozwój edukacji i promocji profesjonalistów<sup>27</sup>. Już obecnie ewaluacja posiada jasno sprecyzowany problem badawczy, zbiór pojęć, dyscypliny referencyjne powiązane przede wszystkim z rozwojem regionalnym, aktualną agendą badawczą i system rozwoju profesjonalistów. Brakuje jej paradygmatów stanowiących odrębne teorie wpływu interwencji na rozwój, a także zasad i praktyk powiązanych z daną dyscypliną, które mogłyby zostać zlikwidowane dzięki wprowadzeniu rozwiązań zaproponowanych w niniejszej publikacji.

---

27 Liles D. H., Johnson M. E., & Meade L., *Enterprise engineering discipline [w:] Proceedings of the Fifth Industrial Engineering Research Conference*, (red.) Askin R.g., Bidanda B., Jagdale S., Institute of Industrial Engineers, Minneapolis, Minnesota, USA 1996. s. 479–484.

## Bibliografia

- Adamkiewicz-Drwiłło H.G.**, *Współczesna metodologia nauk ekonomicznych*, Dom Organizatora TNOiK, Toruń 2008
- Agrotec Polska Sp. z o.o., *Raport z badania pt.: „Ocena zintegrowanego systemu działań informacyjnych, promocyjnych oraz szkoleniowych w zakresie funduszy europejskich w Województwie Małopolskim”*, na zlecenie Urzędu Marszałkowskiego Województwa Małopolskiego, Kraków 2009, pobrano ze strony: [http://www.archiwum.ewaluacja.gov.pl/Wyniki/Documents/1\\_081.pdf](http://www.archiwum.ewaluacja.gov.pl/Wyniki/Documents/1_081.pdf)
- Agrotec Polska Sp. z o.o., *Raport z badania pt.: „Wpływ interwencji finansowanych z funduszy strukturalnych UE na przekształcenie obszarów wiejskich”*, na zlecenie Ministerstwa Rozwoju Regionalnego, Warszawa 2007, pobrano ze strony: [http://www.archiwum.ewaluacja.gov.pl/Wyniki/Documents/rrit\\_001.pdf](http://www.archiwum.ewaluacja.gov.pl/Wyniki/Documents/rrit_001.pdf)
- CASE (Centrum Analiz Społeczno-Ekonomicznych), *Badanie ewaluacyjne nt. „Wpływ polityki spójności na stan zróżnicowań i rozwój polskich regionów”. Raport końcowy*, na zlecenie Ministerstwa Rozwoju Regionalnego, Warszawa 2008, pobrano ze strony: [http://www.archiwum.ewaluacja.gov.pl/Wyniki/Documents/rrit\\_051.pdf](http://www.archiwum.ewaluacja.gov.pl/Wyniki/Documents/rrit_051.pdf)
- Centrum Europejskich Studiów Regionalnych i Lokalnych EUROREG Uniwersytetu Warszawskiego, *Ocena wpływu polityki spójności na rozwój miast polskich (w ramach ewaluacji ex post NPR 2004-2006)*, na zlecenie Ministerstwa Rozwoju Regionalnego, Warszawa 2010, pobrano ze strony: [http://www.archiwum.ewaluacja.gov.pl/Ewaluacja\\_ex\\_post\\_NPR/Documents/raport\\_koncowy\\_expost9.pdf](http://www.archiwum.ewaluacja.gov.pl/Ewaluacja_ex_post_NPR/Documents/raport_koncowy_expost9.pdf)
- Chen, H.-T., Rossi, P.H.**, *Evaluating with sense: The theory-driven approach*, Evaluation Review, No. 7/1983, s. 283–302
- EGO s.c., *Ocena systemu realizacji polityki spójności w Polsce w ramach perspektywy 2004–2006*, na zlecenie Ministerstwa Rozwoju Regionalnego, Warszawa 2010, pobrano ze strony: [http://www.archiwum.ewaluacja.gov.pl/Ewaluacja\\_ex\\_post\\_NPR/Documents/expost\\_koncowy\\_4.pdf](http://www.archiwum.ewaluacja.gov.pl/Ewaluacja_ex_post_NPR/Documents/expost_koncowy_4.pdf)
- Grupa Gomułka, *Raport końcowy z badania ewaluacyjnego: „Analiza przyczyn braku aktywności lub niskiej aktywności niektórych samorządów gminnych w sięganiu po wsparcie z Funduszy Unijnych”*, na zlecenie Ministerstwa Rozwoju Regionalnego, Warszawa 2009, pobrano ze strony: [http://www.archiwum.ewaluacja.gov.pl/Wyniki/Documents/rrit\\_062.pdf](http://www.archiwum.ewaluacja.gov.pl/Wyniki/Documents/rrit_062.pdf)
- Instytut Badań Strukturalnych, *Spójność wewnętrzna a konkurencyjność regionu łódzkiego. Diagnoza strategiczna*, na zlecenie Urzędu Marszałkowskiego Województwa Łódzkiego, Łódź 2009, pobrano ze strony: [http://www.archiwum.ewaluacja.gov.pl/Wyniki/Documents/1\\_071.pdf](http://www.archiwum.ewaluacja.gov.pl/Wyniki/Documents/1_071.pdf)
- Kotarbiński, T.**, *Kurs logiki dla prawników*, Państwowe Wydawnictwa Naukowe, Warszawa 1960

- Liles D. H., Johnson M. E., Meade L.**, *Enterprise engineering discipline* [w:] *Proceedings of the Fifth Industrial Engineering Research Conference*, (red.) Askin R. G., Bidanda B., Jagdale S., Institute of Industrial Engineers, Minneapolis, Minnesota, 1996
- Łobocki M.**, *Wprowadzenie do metodologii badań pedagogicznych*, Wyd. Impuls, Kraków 2007
- Olejniczak K.** (red.), *Praktyka ewaluacji efektów programów rozwoju regionalnego – studium porównawcze*, Raporty i Analizy EUROREG 2/2009, pobrano ze strony: [http://www.euroreg.uw.edu.pl/dane/web\\_euroreg\\_publications\\_files/611/olejniczak\\_2009\\_praktyka\\_ewaluacji\\_efektow.pdf](http://www.euroreg.uw.edu.pl/dane/web_euroreg_publications_files/611/olejniczak_2009_praktyka_ewaluacji_efektow.pdf)
- Olejniczak K.**, *The shortcomings of ERDF evaluation practice in the Member States* [w:] *Evaluating the effects of regional interventions. A look beyond current Structural Funds' practice*, (red.) K. Olejniczak, M. Kozak, S. Bienias, Ministry of Regional Development, Warsaw 2011
- Olejniczak K.**, *Tracing theories behind practice – a theory-driven approach* [w:] *Evaluating the effects of regional interventions. A look beyond current Structural Funds' practice*, (red.) K. Olejniczak, M. Kozak, S. Bienias, Ministry of Regional Development, Warsaw 2011
- Opinia – agencja badania rynku & INFOR Training, *Analiza przyczyn braku aktywności niektórych samorządów w sięganiu po wsparcie unijne w ramach programu ZPORR celem ujednolicenia Planu Komunikacji RPO WZ na lata 2007-2013 oraz Planu Działań Informacyjnych i Promocyjnych na rok 2009. Raport końcowy*, na zlecenie Urzędu Marszałkowskiego Województwa Zachodniopomorskiego, Warszawa 2008
- Pilch, T.**, *Zasady badań pedagogicznych*, Zakład Narodowy im. Ossolińskich, Wrocław 1977
- Proksen Sp. z o.o. & Pracownia Badań Społecznych i Marketingowych „Soma”, *Metodologia badań dla działań informacyjnych i promocyjnych realizowanych ze środków Unii Europejskiej w Województwie Kujawsko-Pomorskim*, na zlecenie Urzędu Marszałkowskiego Województwa Kujawsko-Pomorskiego, Toruń 2008, pobrano ze strony: [http://www.archiwum.ewaluacja.gov.pl/Wyniki/Documents/rrit\\_045.pdf](http://www.archiwum.ewaluacja.gov.pl/Wyniki/Documents/rrit_045.pdf)
- Stowarzyszenie Doradców Gospodarczych Pro-Akademia, *Determinanty rozwoju współpracy międzynarodowej beneficjentów funduszy strukturalnych na poziomie Narodowej Strategii Spójności*, na zlecenie Ministerstwa Rozwoju Regionalnego, Warszawa 2008, pobrano ze strony: [http://www.archiwum.ewaluacja.gov.pl/Wyniki/Documents/rrit\\_060.pdf](http://www.archiwum.ewaluacja.gov.pl/Wyniki/Documents/rrit_060.pdf)
- Szlachta, J.**, *Polskie doświadczenia w zakresie ewaluacji*, II Konferencja Ewaluacyjna – Ewaluacja ex-ante jako narzędzie doskonalenia zarządzania strategicznego w administracji publicznej, Ministerstwo Rozwoju Regionalnego & Polska Agencja Rozwoju Przedsiębiorczości, Warszawa 2006
- Sztumski J.**, *Wstęp do metodologii i techniki badań społecznych*, Wyd. Śląsk, Katowice 1995



### KORNELIUSZ PYLAK

Doktor nauk ekonomicznych, pracownik naukowy Politechniki Lubelskiej i stały współpracownik firmy doradczej WYGPSDB Sp. o.o. Jest autorem wielu wytycznych do studiów wykonalności, wytycznych dla beneficjentów, podręczników oceny projektów, kwalifikowania wydatków, analizy efektów projektów, a także ponad 70 badań ewaluacyjnych. Specjalizuje się w analizach: systemów oceny i wyboru projektów, w tym systemu kryteriów oceny; systemu wskaźników i monitorowania realizacji programów; przedsiębiorczości i innowacyjności, w tym oceny wpływu programów na poziom konkurencyjności i innowacyjności, realizacji RSI i rozwoju organizacji inteligentnych; ogólnie systemu wdrażania funduszy europejskich, w tym komplementarności, linii demarkacyjnej, przepływów finansowych, barier, poziomu przygotowania instytucji itp.



## JAROSŁAW GÓRNIAK

### Od ewaluacji wpływu do teorii programu i z powrotem

Badania ewaluacyjne w sektorze publicznym służą praktyce. Koncentrują swoją uwagę na podejmowanych decyzjach dotyczących działań publicznych, ich konsekwencji oraz zrozumieniu tego, co, na kogo i w jakich okolicznościach działa, a w jakich nie, by przyczynić się do poprawy jakości przyszłych programów. Ewaluacja jest też oceną programu ze względu na odniesienie się do przyjętego systemu wartości. Teoria ewaluacji to zasadniczo zbiór dyrektyw dotyczących sposobu postępowania badawczego wynikających z prezentowanych założeń dotyczących funkcji ewaluacji i adekwatnej metodologii badań. Nie jest ona, oczywiście, oderwana od ogólniejszych założeń ontologicznych, epistemologicznych i metodologicznych konstytuujących teoretyczne podejście do życia społecznego. Dobrym tego przykładem są takie podejścia, jak np.:

- empirycystyczna koncepcja ewaluacji wyrastająca z epistemologii ewolucyjnej Karla Poppera i kładąca nacisk na badania eksperymentalne i quasi-eksperymentalne w kontekście koncepcji „społeczeństwa eksperymentującego” lansowanej przez Donalda Campbella<sup>1</sup>;
- koncepcja ewaluacji opartej na epistemologicznych podstawach brytyjskiego realizmu krytycznego, dążąca do budowania empirycznie adekwatnych modeli mechanizmu społecznego, leżącego u podstaw ewaluowanego programu, wyjaśniających w określonym kontekście uzyskane wyniki interwencji<sup>2</sup>;
- koncepcja ewaluacji czwartej generacji, wyrastająca z relatywistycznego konstruktywizmu epistemologicznego, podważająca znaczenie czy wręcz możliwość obiektywizmu pomiaru i obserwacji oraz orientacji na cele i efekty, koncentrująca się na negocjowaniu definicji sytuacji i wartości interesariuszy zaangażowanych w program, partycypacji wszystkich interesariuszy w ustalaniu potrzeb informacyjnych, kryteriów ewaluacji, warunków jej prowadzenia, interpretacji danych i wykorzystaniu wyników i zacierająca różnice pomiędzy ewaluacją jako przedsięwzięciem diagnostycznym a moderowaniem procesu zmiany społecznej<sup>3</sup>.

1 Campbell D. T., Russo M. J., *Social Experimentation (Sage Classics Series) (1 ed.)*. SAGE Publications, London 1998.

2 Pawson R., Tilley N., *Realistic Evaluation*. Sage Publications, London 1997.

3 Guba E. G., Lincoln Y. S., *Fourth Generation Evaluation (1 ed.)*. Sage Publications, London 1989.

Wśród praktyków dominuje raczej eklektyzm i otwartość na preferencje ze strony zlecających badania. Ewaluacja oparta na teorii programu jest w gruncie rzeczy pewną próbą elastycznej kodyfikacji takiego eklektyzmu łączącego w sobie dążenie do odtworzenia „teorii programu”, często sprowadzanej raczej do modelu logicznego programu rozumianego jako zbiór założeń dotyczących oddziaływania, rezultatów, produktów, działań i nakładów niż społecznego mechanizmu zmiany postulowanego przez Pawsona i Tilly’ego. Wokół takiego modelu logicznego konstruowane były kryteria ewaluacyjne. Proces ewaluacji, obok rekonstrukcji przyjmowanych założeń, co do związków między działaniami a oczekiwanymi efektami obejmował pomiar wartości wskaźników poszczególnych elementów modelu logicznego. Podejście to kładło także nacisk na analizę procesu wdrożenia, by uzyskać lepsze zrozumienie przestarek sukcesu lub porażki. Bezsporną zaletą ewaluacji opartej na teorii programu było powiązanie procesu ewaluacji z procesem planowania interwencji, co sprzyja procesowi uczenia się administracji i doskonaleniu procesów planowania. Zwolennicy tego nurtu zwracają uwagę na korzyści, które może przynieść uwzględnianie bardziej złożonych modeli logicznych, pośrednich efektów, zmian w teorii programu w toku jego realizacji i zróżnicowania efektów dla różnych grup beneficjentów. Słusznie zwraca się uwagę na to, że proste rozstrzygnięcia tego, czy program działa, czy też nie i czy należy go kontynuować, czy się z niego wycofać są mniej cenione przez zarządzających programami niż wiedza, jak można program ulepszyć, skorygować na poszczególnych etapach czy w stosunku do poszczególnych odbiorców<sup>4</sup>. Trzeba też podkreślić, że przykłady wzorcowych ewaluacji opartych o teorię programu zawierają generalnie nie tylko opracowanie ilościowe wyników, lecz także oparte są na planach quasi-eksperymentalnych<sup>5</sup>. Wynika to z zapotrzebowania na obiektywną wiedzę o tym, co faktycznie osiągnięto w programie. Poważny problem stanowi raczej to, że ilościowy obraz programu i jego efektów jest raczej powierzchowną rejestracją wartości założonych wskaźników, a nie modelowym ujęciem mechanizmu przyczynowego prowadzącego od interwencji do pożądanej zmiany.

W toku realizacji perspektywy 2007-2013 narastał krytycyzm wobec, z jednej strony, samych programów realizowanych w ramach Polityki spójności Unii Europejskiej, od których oczekiwano silniejszej orientacji na klarowne wyniki, oraz z drugiej strony wobec instytucji odpowiedzialnych za dostarczanie informacji na temat osiągniętych efektów. Zaniepokojenie budziło korzystanie z łatwo komunikowalnych, zagregowanych miar, częściej dotyczących środków, a nie celów. W związku z tym sugerowany był wzrost nacisku na prawidłowe ustalenie rzeczywistego wpływu realizowanych działań i ponoszonych nakładów na pożądane zmiany<sup>6</sup>. W efekcie Komisja Europejska zwiększyła nacisk na ustalenie tego wpływu w toku ewaluacji.

4 Birckmayer J. D., Weiss C. H., *Theory-based evaluation in practice. Evaluation review*, Vol. 24 no. 4/2000, s. 407-431.

5 Ibidem.

6 Barca F., McCann P., *Outcome Indicators and Targets: Towards a New System of Monitoring and Evaluation in EU Cohesion Policy*, 2011.

W Polsce zaowocowało to wprowadzeniem obowiązku prowadzenia ewaluacji wpływu w obrębie poszczególnych osi priorytetowych.

Nacisk na ewaluację wpływu jest uzasadniony wypełnianiem przez ewaluację funkcji zapewnienia rozliczalności polityki publicznej. Ewaluacja jest z założenia środkiem do wypełniania takiej funkcji: obywatele mają prawo do uzasadnienia potrzeby wprowadzania określonych interwencji, wiedzy na temat tego, w jakim stopniu zostały osiągnięte założone cele, czy uzyskano je w sposób gospodarny i czy, *last but not least*, czy (z)realizowany program jest użyteczny, czyli czy zaspokoił istotne potrzeby społeczne, rozwiązał problemy, miał pozytywne konsekwencje rozwojowe, które przeważały koszty i nieplanowane, negatywne oddziaływania. Kluczową kwestią jest przy tym odpowiedź, czy program w ogóle wywarł wpływ na uzyskanie pożądanej zmiany, nawet gdy taką zmianę zaobserwowano. Nawet jeśli okresowo słabnie zainteresowanie opinii publicznej treścią społeczno-gospodarczą działalności władz publicznych i realizującej ich decyzje administracji, to długofalowa jakość rządzenia wymaga wypełniania funkcji rozliczalności.

Ustalenie tego, czy interwencja publiczna oddziałuje na zaobserwowaną zmianę nie jest kwestią prostą. Jak dotąd wyzwanie to zostało zoperacjonalizowane w postaci zamówień na ewaluację realizowaną metodami kontrfaktycznymi, przy czym najczęściej rozumie się przez to zapewnienie trafności wewnętrznej w procesie analizowania związku przyczynowego między interwencją a opisaną przy pomocy wskaźników zmianą. Ze względu na ograniczenia możliwości zastosowania planów quasi-eksperymentalnych opartych na planie zrandomizowanym lub na zrandomizowanym planie blokowym stosuje się *Propensity Score Matching* (PSM) będące analogiem kompletnej randomizacji lub dopasowanie kompletne, które można traktować jako analog planu blokowego<sup>7</sup>. W Polsce szczególnie popularna jest metoda PSM, co wynika m.in. z wykorzystania jej w pierwszych ewaluacjach realizowanych metodami kontrfaktycznymi i zaprezentowania w powszechnie dostępnych publikacjach PARP<sup>8</sup>. W ostatnio realizowanych ewaluacjach pojawiły się także inne techniki szacowania wpływu interwencji, ale generalnie takie, które mają na celu ustalenie trafnego wewnętrznie oszacowania wielkości efektu przyczynowego bez wnikania w złożoność mechanizmu społecznego generującego zmianę. Bywa to uzupełniane analizami jakościowymi, które koncentrują swoją uwagę na procesie i kontekście wdrożenia, ale oba te elementy analizy nie pozostają ze sobą w ścisłym związku wynikającym z zintegrowanego planu badań.

7 King G., Nielsen R., *Why propensity score should not be used for matching*. Working Paper, 2016.

8 Trzciński R., *Wykorzystanie techniki propensity score matching w badaniach ewaluacyjnych*. PARP, Warszawa 2009.



## Ograniczenia eksperymentu randomizowanego

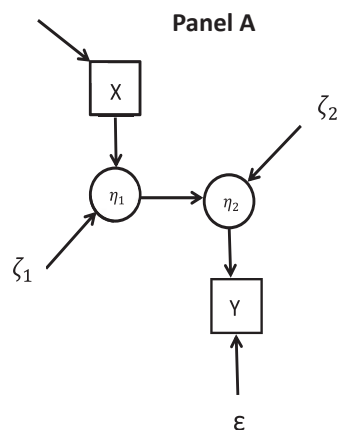
Eksperyment randomizowany jest „złotym standardem” badań przyczynowych. Wcześniej wskazane zostały już niektóre ograniczenia tego podejścia i opartych na jego fundamencie analogów (PSM i dopasowania dokładnego), które stosuje się ze względu na rzadką możliwość stosowania eksperymentu w ewaluacji. Jednym z bardziej bolesnych jest efekt „czarnej skrzynki” w której ukryty jest mechanizm stojący u podstaw procesu uzyskiwania efektu interwencji. Ale są też inne ograniczenia.

Wyniki badań opartych na planie zrandomizowanym, ale odnoszących się do tzw. eksperymentu terenowego, w którym badacz nie ma pełnej kontroli nad wszystkim aspektami przebiegu interwencji, mogą być zagrożone różnymi obciążeniami. Jednym z nich jest tzw. „obciążenie randomizacyjne” oszacowań – mierzony jest raczej efekt zamiaru uczestnictwa niż efekt uczestnictwa w programie: uczestnicy mogą odpaść, a nie-uczestnicy mogą znaleźć substytut<sup>9</sup>.

W przywołanym artykule autorzy wymieniają także inne ograniczenia eksperymentów społecznych, także o zbliżonym charakterze do omówionych w tym artykule. Tutaj zwrócimy uwagę – w oparciu o znakomity artykuł Kennetha A. Bollen i Judei Perla – na pewną szczególną wiązkę problemów, które pojawiają się w kontekście klasycznego, prostego eksperymentu randomizowanego wówczas, gdy zarówno faktyczny bodziec, jak i efekt są zmiennymi ukrytymi manifestującymi się przez obserwowalne wskaźniki – tak jest najczęściej w przypadku badań społecznych. Problemy te są dziedziczone przez badania quasi-eksperymentalne o takim samym planie. Diagramy przyczynowe A, B i C oraz wykaz problemów przytoczone są za wspomnianą publikacją<sup>10</sup>.

Na panelu A, X to bodziec użyty w warunkach randomizacji, Y to wynik pomiaru efektu.  $\eta_1$  i  $\eta_2$  to zmienne latentne obrazujące: (1) faktyczną zmienną niezależną – przyczynę, na którą działamy empirycznym bodźcem, by wywołać zmianę jej wartości (np. postawa) i (2) zmienną zależną – efekt, który jest mierzony przez określony wskaźnik. Problem, który pojawiają się już w tym ujęciu jest następujący:

- Jak silnie bodziec działa na  $\eta_1$ ?
- Jak dobry jest pomiar efektu?
- Jeśli te związki są słabe, wówczas wynik może być nieistotny nawet gdy wpływ na poziomie zmiennych latentnych jest silny.

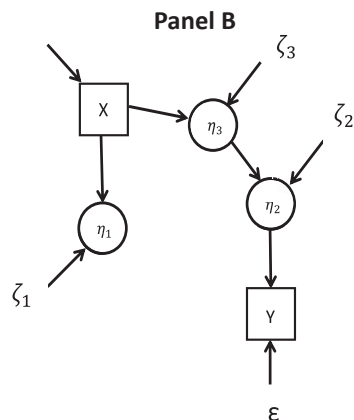


9 Heckman J. J., Smith, J. A., *Assessing the case for social experiments*, The Journal of Economic Perspectives, Vol. 9 no.(2)/1995, s. 85-110.

10 Bollen K. A., Pearl, J. *Eight myths about causality and structural equation models*, [w:] Morgan S. L., *Handbook of causal analysis for social research*, Springer 2013, dostępne jako [http://link.springer.com/chapter/10.1007/978-94-007-6094-3\\_15](http://link.springer.com/chapter/10.1007/978-94-007-6094-3_15).

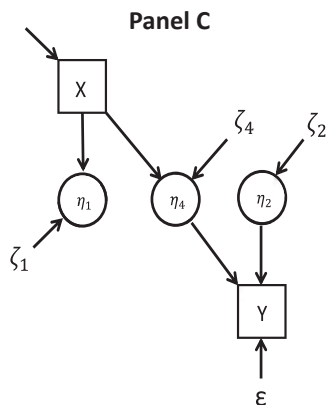
Problemem może być to, że odczyt na poziomie zmiennych jawnych może odzwierciedlać inny mechanizm przyczynowy, co ilustruje Panel B.

- X działa na inną zmienną, która wpływa na latentną zmienną zależną.
- X i Y pozostają w istotnej statystycznie zależności.



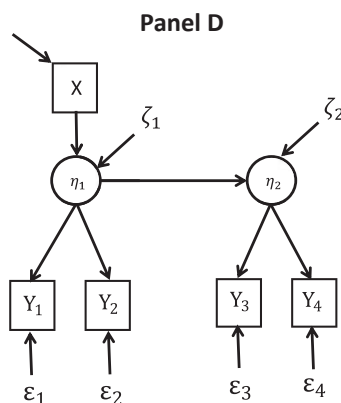
Możliwy jest jeszcze inny mechanizm przyczynowy, który ilustruje Panel C.

- Zależność między X a Y jest istotna.
- $\eta_1$  nie wpływa na  $\eta_2$ .
- Bodziec wpływa na inną zmienną -  $\eta_4$ , która wpływa na Y wraz z  $\eta_2$ .



Jak się można upewnić co do tego, czy nasz eksperyment potwierdza teorię o przyczynowym wpływie  $\eta_1$  na  $\eta_2$  (np. że frustracja wywołuje agresję)? Konieczne jest przygotowanie pomiaru pod kątem modelu strukturalnego. Jego postać zawiera Panel D.

1. Zmierzyć wartości:
  - a. min. 2 wskaźników  $\eta_1$
  - b. min. 2 wskaźników  $\eta_2$
2. Oszacować SEM:
  - a. Jeśli mamy sytuację z panelu B lub C – dopasowanie będzie słabe
  - b. Słaby wpływ X na  $\eta_1$  przełoży się na niski  $R^2$  dla  $\eta_1$
  - c. Słaby pomiar  $\eta_2$  odbije się w niskich  $R^2$  dla  $y_3$  i  $y_4$



Jak widać z tego przykładu, nawet randomizowany eksperyment może nie uchronić nas przed błędnymi wnioskami, jeśli nie otworzymy czarnej skrzynki zawierającej mechanizm generujący obserwowane relacje. Ten prosty przykład pokazuje, dlaczego warto szukać sposobu na powiązanie twardego pomiaru oddziaływań przyczynowych z uważną elaboracją modelu mechanizmu przyczynowego.

### W kierunku syntezy

Same ustalenia ilościowe nie są wystarczającym produktem badania ewaluacyjnego. Nawet jeśli dobrze oszacujemy efekt netto, może to mieć ograniczony wpływ na proces uczenia się decydentów, gdyż zdecydowanie wyższą wartość ma to samo oszacowanie wzbogacone o obraz mechanizmu przyczynowego. Wiedza dostarczana przez ewaluatorów powinna wspierać i stymulować proces uczenia się decydentów i zespołów wdrażających programy i prowadzić do wzrostu społecznej wartości programów realizowanych w przyszłości. To, jaki sposób podejścia najlepiej służy temu procesowi, jest przedmiotem debaty teoretyków. Ważne jest przy tym, by ewaluacja nie ograniczała się do ustalenia historycznego efektu interwencji. James Heckman wskazuje na trzy kluczowe wyzwania, przed którymi stoi ewaluacja:

- Ewaluacja wpływu (przyczynowego) przeszłych interwencji na obserwowane efekty, w tym ich wpływ na dobrobyt osób, na które interwencja jest ukierunkowana.
- Przewidywanie wpływu interwencji już wdrożonych w jednym środowisku w innych środowiskach, z uwzględnieniem dobrobytu.
- Przewidywanie efektów interwencji nigdy nie wprowadzanych w nowych warunkach środowiskowych, z uwzględnieniem skutków dla dobrobytu<sup>11</sup>.

Zdaniem Heckmana podejście „statystyczne” oparte na randomizacji i jej analogach, jak PSM, jest zainteresowane tylko pierwszym z tych wyzwań, natomiast nie jest w stanie sprostać dwóm pozostałym. Po to, by móc się z nimi zmierzyć, konieczne jest podejście „ekonometryczne” oparte na modelowaniu strukturalnym zależności zachodzących w toku interwencji. Oczywiście, kluczowym problemem jest zbudowanie adekwatnego modelu interwencji pozwalającego właściwie oszacować efekt przyczynowy, a jednocześnie dostarczyć wglądu w mechanizm, który może być podstawą do trafniejszego modelowania polityki w przyszłości. Jak się zdaje, dużą rolę w procesie dochodzenia do takiego modelu może odegrać jakościowo zorientowane badanie na gruncie ewaluacji opartej o teorię programu, zwłaszcza jeśli wykroczy poza uproszczoną matrycę logiczną w kierunku badania mechanizmu działającego w określonym kontekście, tak jak to proponuje podejście realistyczne Pawsona i Tilly’ego. Dobrze zweryfikowane teorie programu, czy lepiej: teorie mechanizmu leżącego u podstaw programu są ważną zdobyczą ewaluacji. Jeśli nawet nie stają się natychmiast elementem podręczne-

11 Heckman J. J., *The scientific model of causality*, Sociological methodology, Vol. 35(1)/2005, s. 1-97.

go zasobu wykorzystywanego przez decydentów, to trafiają do bazy wiedzy o rozwiązaniach wchodzącej w skład zbioru problemów do rozwiązania, koncepcji rozwiązań czekających na problemy, okazji umożliwiających wybór oraz ludzi motywowanych do działania na rzecz rozwiązań<sup>12</sup>.

W literaturze pojawiły się już koncepcje oparte na komplementarności ewaluacji opartej na teorii i ewaluacji oddziaływania, jak np. zaproponowana przez White'a *theory based impact evaluation*<sup>13</sup>. Proponuje on sześć kluczowych zasad ewaluacji wpływu opartej na teorii:

1. Nakreśl (*map out*) łańcuch przyczynowy (teorię programu).
2. Zrozum kontekst.
3. Antycypuj heterogeniczność.
4. Rygorystyczna ewaluacja wpływu z użyciem wiarygodnego podejścia kontrfaktycznego.
5. Rygorystyczna analiza stanu faktycznego.
6. Używaj metod mieszanych<sup>14</sup>.

W podobnym duchu postulują van der Knapp i in. na gruncie kryminologii i integracji campbellowskiego standardu badań na planie eksperymentalnym i opartego na nich przeglądu systematycznego z założeniami ewaluacji realistycznej, zmierzającej do analizy mechanizmu leżącego u podstaw programu, kontekstu, w którym program działa i wyników, zarówno zamierzonych, jak i niezamierzonych<sup>15</sup>.

### Jak to zrobić?

Przytoczone wyżej teksty White'a czy Knappa i in. trafnie wskazują kierunek, natomiast nie podpowiadają precyzyjnie ścieżki, którą można pójść. Rozsądne wydaje się skonstruowanie prostego schematu postępowania:

1. Najpierw działamy w ramach podejścia ewaluacji opartej na teorii programu, najlepiej w bardziej zaawansowanym wydaniu, zbliżonym do ewaluacji realistycznej.
2. W oparciu o ustalenia etapu 1. konstruujemy ukierunkowany graf acykliczny (DAG) reprezentujący ścieżkę działania interwencji na efekt i ustalamy minimalny zestaw zmiennych podlegających kontroli, by można było bez zakłócenia przez tzw. „tylne drzwi” oszacować zależność przyczynową<sup>16</sup>.

12 Cohen M. D., March J. G. Olsen J. P., *A Garbage Can Model of Organizational Choice*, [w:] *Administrative Science Quarterly*, Vol 17 no. (1)/1972, s. 1-25. dostępne jako: <http://www.jstor.org/stable/2392088>.

13 White H., *Theory-based impact evaluation: principles and practice*, *Journal of development effectiveness*, Vol. 1 no. (3)/2009, s. 271-284.

14 White H., *Theory-based impact...*, op. cit., s. 274.

15 van der Knaap L. M., Leeuw F. L., Bogaerts S., Nijssen L. T. J., *Combining Campbell Standards and the Realist Evaluation Approach The Best of Two Worlds*, *American journal of evaluation*, Vol. 29 no. (1)/2008, s. 48-57.

16 Morgan S. L., Winship C., *Counterfactuals and Causal Inference: Methods and Principles for Social Research (Analytical Methods for Social Research) (2 ed.)*. Cambridge University Press, Cambridge 2014.

3. Projektujemy plan gromadzenia danych zgodnie z wynikami analizy DAG z uwzględnieniem wymogów pomiaru, m.in. ilustrowanych wyżej.
4. Realizujemy starannie badanie.
5. Szacujemy parametry modelu przyczynowego.
6. Podsumowujemy wyniki ewaluacji opisując mechanizm działający u podstaw programu oraz przedstawiając wielkość efektu wywołanego przez interwencję.

Pewnego dodatkowego wyjaśnienia wymaga etap 2 i 3. Nie ma tu miejsca na szerokie referowanie języka graficznego opisu zależności przyczynowych opartego na grafach acyklicznych. Jest dostępna w tym zakresie obszerna już literatura, w tym zwłaszcza łatwo dostępne na stronie autora artykuły Judei Perla czy ostatnio wydany podręcznik stanowiący dobre wprowadzenie do tych zagadnień<sup>17</sup>. Wszechstronne omówienie zagadnień związanych z badaniami przyczynowymi zawiera też inna znana monografia<sup>18</sup>.

Zasadniczo, w oparciu o wyniki uzyskane z jakościowej fazy badań opartej na teorii programu należy wykonać następujące kroki:

1. Ustalić co jest interwencją, a co wynikiem.
2. Przeprowadzić główną ścieżkę przyczynowego oddziaływania interwencji na efekt uwzględniając wszystkie potencjalne mediatory.
3. Należy dodać do modelu wszystkie zmienne, które mogą przyczynowo wpływać na wynik interwencji.
4. Wskazać oddziaływania przyczynowe pomiędzy wszystkimi zmiennymi wpływającymi na wynik interwencji a interwencją.
5. Ustalić, czy występują jeszcze jakiegokolwiek zależności przyczynowe pomiędzy którymkolwiek zmiennymi uwzględnionymi w modelu i uwzględnić reprezentujące je strzałki w modelu.
6. Ustalić dla każdej pary zmiennych, czy istnieją jakiegokolwiek zmienne, które wpływają przyczynowo na parę<sup>19</sup>.

Graficzny proces budowania modelu przyczynowego można wykonać przy pomocy bezpłatnego programu DAGITTY: <http://www.dagitty.net>. Tam także można znaleźć bibliotekę przykładów oraz publikacji. Program automatycznie wskazuje minimalny zestaw zmiennych, które muszą zostać objęte pomiarem i które z nich mają być „warunkowane” w badaniu, czyli ustalone na stałym poziomie, w celu wykluczenia zakłócenia obrazu interesującej nas zależności przyczynowej. To bardzo ważna cecha DAG. Pozwala ona zredukować dużą liczbę zmiennych pierwotnie branych pod uwagę na podstawie eksploracji przeprowadzonej w ramach fazy jakościowej, opartej na teorii programu, realizowanej w duchu metody teorii ugruntowanej,

17 Pearl J., Glymour M., Jewell, N. P., *Causal Inference in Statistics: A Primer (1 ed.)*. Wiley 2016.

18 Morgan S. L., Winship C., *Counterfactuals and Causal...*, op. cit.

19 Ta lista kroków została sporządzona na podstawie znakomitego wykładu Scotta Vennersa umieszczonego w youtube.com: <https://www.youtube.com/watch?v=pJhU4fimHBQ>.

do listy tych, które należy objąć pomiarem, a do tego pozwala zaplanować oszczędny plan badawczy ukierunkowany na zablokowanie „tylnych drzwi” informacyjnych w modelu i poprawne oszacowanie efektu interwencji.

### Podsumowanie

Zintegrowanie badania opartego na teorii programu z szacowaniem efektu przyczynowego interwencji przy pomocy fazy modelowania z użyciem kierunkowych grafów acyklicznych zdaje się być efektywnym sposobem prowadzenia badań ewaluacyjnych, które łączą w sobie dwie kluczowe cechy: są użyteczne dla zarządzających programami i dają poprawny obraz mechanizmu przyczynowego leżącego u podstaw programu i jego efektu netto. Trzeba od razu dodać, że prace nad modelowaniem zależności przyczynowych w oparciu o podejście kontrfaktyczne z udziałem DAG jest stale rozwijane.

Obecnie w druku jest interesująca monografia dotycząca innowacji w tym obszarze, która nosi wdzięczną nazwę *Single World Intervention Graphs*, której wczesnym zarysem jest dostępny w sieci *working paper*<sup>20</sup>. Autorzy zaproponowali nowy sposób identyfikacji na diagramach wariantów kontrfaktycznych przez podział węzłów, co uprościło i uporządkowało reprezentację mechanizmów przyczynowych.

Dostępnych jest już całkiem sporo programów, które wspomagają modelowanie mechanizmów przyczynowych; część z nich to pakiety lub internetowe aplikacje komercyjne, ale są też aplikacje bezpłatne, wśród których jest wspomniany wyżej, przyjazny DAGITTY. To będzie sprzyjać upowszechnieniu zarysowanego tu podejścia.

---

20 Richardson T. S., Robins J. M., *Single World Intervention Graphs (SWIGs): A unification of the counterfactual and graphical approaches to causality*. Center for the Statistics and the Social Sciences, University of Washington Series. Working Paper, Vol.128 (30)/2013. dostępne jako <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.294.7647&rep=rep1&type=pdf>.

## Bibliografia

- Barca F., McCann P.**, *Outcome Indicators and Targets: Towards a New System of Monitoring and Evaluation in EU Cohesion Policy*, 2011
- Birkmayer J. D., Weiss C. H.**, *Theory-based evaluation in practice. Evaluation review*, Vol. 24 no. 4/2000, s. 407-431
- Bollen K. A., Pearl, J.** *Eight myths about causality and structural equation models*, [w:] Morgan S. L., *Handbook of causal analysis for social research*, Springer 2013, dostępne jako [http://link.springer.com/chapter/10.1007/978-94-007-6094-3\\_15](http://link.springer.com/chapter/10.1007/978-94-007-6094-3_15)
- Campbell D. T., Russo M. J.**, *Social Experimentation (Sage Classics Series)* (1 ed.). SAGE Publications, London 1998
- Cohen M. D., March J. G. Olsen J. P.**, *A Garbage Can Model of Organizational Choice*, [w:] *Administrative Science Quarterly*, Vol 17 no. (1)/1972, s. 1-25. dostępne jako: <http://www.jstor.org/stable/2392088>
- Górniak J.**, *Efekt przyczynowy w ewaluacji*, [w:] Haber A., Szataj M. (red.), *Ewaluacja wobec wyzwań stojących przed sektorem finansów publicznych* (pp. 189-200), PARP, Warszawa 2009, dostępne jako <http://www.parp.gov.pl/files/74/81/305/6635.pdf>
- Guba E. G., Lincoln Y. S.**, *Fourth Generation Evaluation* (1 ed.). Sage Publications, London 1989
- Heckman J. J., Smith, J. A.**, *Assessing the case for social experiments, The Journal of Economic Perspectives*, Vol. 9 no.(2)/1995, s. 85-110
- Heckman J. J.**, *The scientific model of causality, Sociological methodology*, Vol. 35(1)/2005, s. 1-97
- King G., Nielsen R.**, *Why propensity score should not be used for matchning. Working Paper*, 2016
- Morgan S. L., Winship C.**, *Counterfactuals and Causal Inference: Methods and Principles for Social Research (Analytical Methods for Social Research)* (2 ed.). Cambridge University Press, Cambridge 2014
- Pawson R., Tilley N.**, *Realistic Evaluation*. Sage Publications , London 1997
- Pearl J., Glymour M., Jewell, N. P.**, *Causal Inference in Statistics: A Primer* (1 ed.). Wiley 2016
- Richardson T. S., Robins J. M.**, *Single World Intervention Graphs (SWIGs): A unification of the counterfactual and graphical approaches to causality. Center for the Statistics and the Social Sciences, University of Washington Series. Working Paper*, Vol.128 (30)/2013. dostępne jako <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.294.7647&rep=rep1&type=pdf>
- Trzcíński R.**, *Wykorzystanie techniki propensity score matching w badaniach ewaluacyjnych*. PARP, Warszawa 2009
- van der Knaap L. M., Leeuw F. L., Bogaerts S., Nijssen L. T. J.**, *Combining Campbell Standards and the Realist Evaluation Approach The Best of Two Worlds, American journal of evaluation*, Vol. 29 no. (1)/2008, s. 48-57
- White H., *Theory-based impact evaluation: principles and practice, Journal of development effectiveness*, Vol. 1 no. (3)/2009, s. 271-284





### JAROSŁAW GÓRNIAK

Socjolog i ekonomista, profesor nauk społecznych specjalizujący się w zakresie metod badań społecznych, metodologii ewaluacji i analiz polityk publicznych, socjologii gospodarki, edukacji i organizacji oraz statystycznej analizy danych. Dziekan Wydziału Filozoficznego oraz kierownik Zakładu Socjologii Gospodarki, Edukacji i Metod Badań Społecznych w Instytucie Socjologii Uniwersytetu Jagiellońskiego. Twórca i do 2015 r. kierownik Centrum Ewaluacji i Analiz Polityk Publicznych UJ, a obecnie przewodniczący Rady Centrum. W ostatnich latach był kierownikiem naukowym kilku dużych programów badawczych, w tym wieloletniego, systemowego projektu Bilans Kapitału Ludzkiego. Kierował opracowaniem dla potrzeb MNiSW projektu „Programu rozwoju szkolnictwa wyższego na lata 2015-2030” oraz „Diagnozy szkolnictwa wyższego” w ramach prac pod auspicjami KRASP nad Programem Rozwoju Szkolnictwa Wyższego do 2020r. w Polsce. Był ekspertem UNDP i Banku Światowego, jest autorem licznych opracowań i ekspertyz dla centralnej i regionalnej administracji publicznej dotyczących strategicznego zarządzania rozwojem, ewaluacji i analiz polityk publicznych oraz polityki rynku pracy i szkolnictwa wyższego. Służy swą opinią w ciałach doradczych. Obecnie jest członkiem Komitetu Polityki Naukowej (któremu przewodniczy do końca bieżącej kadencji), Rady Naukowej Ośrodka Badań nad Przedsiębiorczością PARP i Rady Naukowej Instytutu Badań Edukacyjnych. Laureat nagrody Lider Ewaluacji 2014 za wkład w rozwój kultury ewaluacyjnej w Polsce.





## RAFAŁ TRZCIŃSKI

### Wprowadzenie

Ewaluacja może pełnić rozmaite funkcje. Jedną z częściej wymienianych jest funkcja rozwojowa, zwana również formatywną<sup>1</sup>. W praktyce polega ona na wykorzystaniu wiedzy zdobywanej w wyniku prowadzonych badań, do ciągłego usprawniania realizowanych polityk publicznych. Ponad dziesięcioletnie już doświadczenia z ewaluacji programów finansowanych ze środków europejskich pokazują, że zadanie to jest niezwykle trudne w realizacji. Ewaluacja często postrzegana jest raczej jako formalny obowiązek, towarzyszący wydatkowaniu środków unijnych. Jej „życie” często kończy się wraz z oddaniem raportu końcowego z ewaluacji i rozliczeniem badania. Całościowe wyjaśnienie przyczyn tego stanu rzeczy nie jest przedmiotem niniejszego artykułu. Powstały na ten temat osobne opracowania poświęcone m.in. wykorzystaniu wyników ewaluacji w kontekście projektowania polityk publicznych<sup>2</sup>. Zagadnienie to porusza również w bardziej pogłębiony sposób inny artykuł przedstawiony w ramach niniejszej publikacji (por. Piotr Stronkowski). Warto wreszcie dodać, że jest to jeden z głównych problemów podnoszonych w ramach prowadzonej obecnie debaty na poziomie europejskim<sup>3</sup>. Jedną z kluczowych kwestii wskazywanych w dyskusji na temat użyteczności ewaluacji jest postulat związany z koniecznością podniesienia jakości realizowanych badań, co sprzyjać powinno ich większemu wykorzystaniu niż miało to miejsce dotychczas. Szczególny nacisk kładziony jest zwłaszcza na uczynienie z ewaluacji źródła twardych „dowodów”, przemawiających za lub przeciw skuteczności danej interwencji. Sprecyzowanie powyższego oczekiwania względem ewaluacji doprowadziło w ciągu ostatnich lat do popularyzacji rygorystycznych metod szacowania efektów przyczynowych w ramach ewaluacji wpływu (*impact evaluation*), bazujących na tzw. podejściu kontrfaktycznym. Rangi całemu procesowi nadały wprowadzone w ostatnich latach odgórnie wytyczne Komisji Europejskiej, związane z ewaluacją programów operacyjnych dla perspektywy 2014-2020<sup>4</sup>. Komisja oczekuje (a w zasadzie wymaga) od każdego z krajów człon-

- 1 Górnjak J., *Ewaluacja w cyklu polityk publicznych*, [w:] S. Mazur (red.), *Ewaluacja funduszy strukturalnych – perspektywa regionalna*, Uniwersytet Ekonomiczny w Krakowie, Małopolska Szkoła Administracji Publicznej, Kraków 2007.
- 2 Olejniczak K., *Mechanizmy wykorzystania ewaluacji. Studium ewaluacji śródkresowych INTERREG III*, Centrum Europejskich Studiów Regionalnych i Lokalnych UW, Wydawnictwo Naukowe SCHOLAR, Warszawa 2008.
- 3 Por. wystąpienie prof. Elliota Sterna *The Role of Evaluation for the Future of EU Cohesion Policy* na konferencji: „*The result orientation: Cohesion Policy at work*”, Sofia 2016.
- 4 *The Programming Period 2014-2020. Guidance Document on Evaluation Plans*, Komisja Europejska, luty 2015 - [http://ec.europa.eu/regional\\_policy/sources/docoffic/2014/working/evaluation\\_plan\\_guidance\\_en.pdf](http://ec.europa.eu/regional_policy/sources/docoffic/2014/working/evaluation_plan_guidance_en.pdf).

kowskich, by w ramach prowadzonych ewaluacji wykonana została ewaluacja wpływu na poziomie każdej osi priorytetowej danego programu. Dodatkowo Komisja wskazała, że podejście kontryfakcyjne jest zalecanym do tego narzędziem. Tworzone obowiązkowo plany ewaluacyjne dla programów operacyjnych z perspektywy 2014-2020 wskazują, że zalecenia Komisji – przynajmniej na gruncie polskim – zostaną w dużym stopniu zaimplementowane.

Niezależnie od oczekiwań formułowanych na poziomie Komisji Europejskiej, metody kontryfakcyjne przebiły się w ostatnich latach do świadomości zarówno osób zlecających jak i wykonujących ewaluację. Stało się to m.in. za sprawą atrakcyjności tego podejścia, wyrażającego się w „obietnicy” udzielenia odpowiedzi na pytanie o to jakie są faktyczne skutki realizowanych działań. Dotychczasowe doświadczenia w stosowaniu badań ewaluacyjnych z wykorzystaniem podejścia kontryfakcyjnego, nie rozwiązały jednak generalnego problemu – tj. niskiego wykorzystania wyników ewaluacji. Niniejszy artykuł podejmuje próbę wyjaśnienia tego stanu rzeczy oraz wskazuje możliwe sposoby doskonalenia badań ewaluacyjnych w Polsce, tak by podnosić ich faktyczną użyteczność dla projektowanych polityk publicznych. W szczególności zaprezentowany został sposób, w jaki badania kontryfakcyjne powinny być uzupełniane przez podejście oparte na teorii (*theory-based evaluation*). Rozważane są zarówno teoretyczne jak i praktyczne problemy w implementacji obu podejść równocześnie.

### **Pułapka czarnego pudełka**

Przeprowadzona poprawnie ewaluacja wpływu, wykorzystująca metody kontryfakcyjne, pozwala odpowiedzieć na pytanie „czy dana interwencja działa?”. Innymi słowy pozwala zweryfikować, czy pomysł na rozwiązanie danego problemu społeczno-gospodarczego przyniósł oczekiwane skutki, czy też nie. W pewnych sytuacjach taka wiedza będzie wystarczająca do podjęcia wiążących decyzji dotyczących kontynuowania realizacji danego przedsięwzięcia w przyszłości, jednak w większości przypadków będzie to stanowczo za mało. Proces decyzyjny, czy też planistyczny w obszarze projektowanych polityk publicznych, zazwyczaj nie prowadzi do całkowitego odrzucenia danego rozwiązania (programu, działania), lecz jego modyfikacji. Wynika to z faktu, iż problemy jakim stawiają czoła interwencje publiczne (bezrobocie, ubóstwo, wykluczenie społeczne, niska konkurencyjność przedsiębiorstw itp.) mają zazwyczaj charakter trwały. Kolejne programy są zazwyczaj pewną zmodyfikowaną wersją tego co zostało zaimplementowane wcześniej. Jest to szczególnie widoczne w przypadku programów operacyjnych finansowanych ze środków europejskich, w których z łatwością można wskazać działania, których założenia i cele są bardzo zbliżone, niezależnie od perspektywy finansowej, w której były lub są wdrażane. W procesie przechodzenia z jednej do drugiej interwencji pojawia się naturalne miejsce dla zastosowania ewaluacji, z której ustalenia w zakresie skuteczności zaimplementowanych rozwiązań powinny być punktem wyjścia do sformułowania założeń dotyczących kształtu kolejnych programów. Od razu trzeba zauważyć, że sama wiedza dotycząca tego czy

dana interwencja okazała się skuteczna, czy też nie, będzie w większości przypadków niewystarczająca do tego by zaproponować jakiegokolwiek sensowne propozycje dla kolejnych działań. Ewaluacje wpływu już w latach osiemdziesiątych określane były mianem tzw. czarnych pudełek, dzięki którym co prawda wiemy co jest „na wejściu” oraz co jest „na wyjściu” analizowanej interwencji<sup>5</sup>, ale zupełnie nie wiemy co dzieje się w środku. Oszacowany efekt przyczynowy mówi nam jedynie, czy wystąpiła relacja przyczynowa między realizacją danego programu a zaobserwowaną zmianą, jaki jest jej kierunek oraz wielkość. Bazując na tych informacjach nie jest możliwe udzielenie odpowiedzi na pytania „dlaczego interwencja okazała się skuteczna w ustalonym stopniu?” i wreszcie „jak zmienić dany program, by był on w przyszłości bardziej skuteczny lub efektywny?”. Aby móc to zrobić, ewaluacja poza ustaleniem efektu interwencji musi skupić się na mechanizmach przyczynowych, które sprawiają, że oczekiwane skutki materializują się w wyniku podjętych działań.

Powyższych obserwacji bynajmniej nie należy traktować jako krytyki podejścia kontrfaktycznego. Jego zastosowanie jest konieczne, jeśli chcemy zweryfikować występowanie relacji przyczynowej między daną interwencją i obserwowanymi (lub nie) zmianami w rzeczywistości społeczno-ekonomicznej. Jednak, żeby w pełni wykorzystać wyniki badań kontrfaktycznych oraz zrozumieć przyczyny sukcesów lub porażek działań publicznych, konieczna jest równoczesna analiza procesów i mechanizmów, stojących za wywołaniem określonych zmian. Orędownikiem takiego podejścia jest np. Howard White<sup>6</sup>, wskazujący na potrzebę realizacji tzw. ewaluacji wpływu opartych na teorii (*theory based impact evaluation*). Działania badawcze niezbędne do przeprowadzenia takiej analizy znacząco wykraczają poza warsztat metod kontrfaktycznych. W praktyce wymagają również znacznie szerszego zasobu danych włączając w to dane monitoringowe oraz przekrojowe dane o charakterze jakościowym. Co istotne podejście takie sugerowane jest również przez Komisję Europejską w odniesieniu do ewaluacji programów operacyjnych realizowanych w perspektywie 2014-2020.

Dalsza część niniejszego artykułu opisuje podejście oparte na teorii, wskazując na jego główne cechy oraz przedstawia praktyczne rekomendacje do jego implementacji w praktyce badawczej.

### **Ewaluacja oparta na teorii**

Projektowane, a następnie wdrażane interwencje publiczne są za każdym razem swego rodzaju „teorią” na temat tego, w jaki sposób należy rozwiązać określony problem o charakterze społeczno-gospodarczym<sup>7</sup>. Typową sytuacją w procesie tworzenia programów publicznych

5 Tzn. znamy przyczynę i skutek.

6 White H., *Theory-Based Impact Evaluation: Principles And Practice*, Working Paper n. 3, International Initiative for Impact Evaluation, New Delhi 2009.

7 Górniak J., *Ewaluacja w cyklu...*, op. cit.

jest przyjmowanie założenia, że określone działanie publiczne (zdarzenie A) jest warunkiem koniecznym (a często również wystarczającym) do wywołania pewnych, pożądanych zmian na poziomie mikro (zdarzenie B). Zmiany te najczęściej identyfikowane są bezpośrednio na poziomie podmiotów objętych daną interwencją. Zmiany wywołane na poziomie mikro, po osiągnięciu odpowiedniej „masy krytycznej”, powinny przynieść wymierne efekty widoczne również na poziomie makro (zdarzenie C), które zazwyczaj utożsamiane są z rozwiązaniem problemów, będących przyczyną uruchomienia danej interwencji.

Przykładem takiej logiki jest chociażby wsparcie publiczne udzielane przedsiębiorstwom w ramach pierwszej osi priorytetowej Regionalnego Programu Operacyjnego Województwa Kujawsko-Pomorskiego (RPO WK-P) – *Wzmocnienie innowacyjności i konkurencyjności gospodarki regionu*. Działanie 1.6 – *Wspieranie tworzenia i rozszerzania zaawansowanych zdolności w zakresie rozwoju produktów i usług*, zgodnie z zapisami Programu, jest odpowiedzią na niską innowacyjność, a co za tym idzie konkurencyjność przedsiębiorstw sektora MŚP województwa kujawsko-pomorskiego. Chcąc zmienić tę niekorzystną sytuację ustalono, że konieczne są dodatkowe, zewnętrzne bodźce (w postaci wsparcia publicznego), które doprowadzą do uruchomienia w przedsiębiorstwach odpowiednich procesów inwestycyjnych zorientowanych na wdrażanie innowacji. W związku z ograniczonym kapitałem własnym przedsiębiorców, mogą oni liczyć na wsparcie finansowe w ramach Programu udzielane w formie dotacji oraz rozmaitych instrumentów finansowych (pożyczki, poręczenia). Działanie 1.6 stanowi kompletną interwencję (zdarzenie A), która ma zachęcać podmioty gospodarcze do inwestycji w środki trwałe (maszyny i sprzęt produkcyjny), dzięki którym mają zostać opracowane i wprowadzone na rynek nowe lub ulepszone produkty lub usługi. W skutek zrealizowanych inwestycji wzrosnąć ma konkurencyjność przedsiębiorstw, wyrażana przez podniesienie wydajności pracy i efektywności produkcji firm (zdarzenie B). Jednocześnie oczekuje się, że pomoc publiczna udzielona przedsiębiorstwom przełoży się nie tylko na funkcjonowanie samych beneficjentów, ale przyniesie również realne korzyści dla konkurencyjności regionu (zdarzenie C). Przy czym zdarzenie C należy rozpatrywać w kontekście realizacji całej interwencji skierowanej do przedsiębiorstw w ramach pierwszej osi priorytetowej RPO WK-P.

Oczywiście przedstawiona powyżej logika stanowi istotne uproszczenie zasad, na jakich działa wskazana interwencja. W rzeczywistości, relacje między zdarzeniami A, B i C nie są jednoznacznie określone (tj. w sposób przyczynowo-skutkowy), bowiem współzależą od wielu, często powiązanych ze sobą, czynników. Interwencje publiczne są złożonymi działaniami, realizowanymi w skomplikowanym otoczeniu społeczno-gospodarczym, prawnym i instytucjonalnym. Aby schemat „A powoduje B, zaś B powoduje C” faktycznie wystąpił muszą być spełnione dodatkowe założenia i warunki. W rzeczywistości realizacja każdej interwencji wiąże się z przyjęciem założeń, że jej wprowadzenie zakończy się sukcesem (tj. wywoła faktycznie pożądany skutek) w przyjętym systemie realizacji i kształcie interwencji, przy przewidzianych zasobach

i skali, w określonym czasie oraz w danym otoczeniu gospodarczym, społecznym, instytucjonalnym oraz porządku prawnym. Co istotne, większość tego typu założeń nie jest wyrażona wprost na poziomie samej interwencji. Często pozostają one jedynie w głowach decydentów i osób tworzących dany program. Zazwyczaj w oficjalnym (publicznym) obiegu obowiązuje bardzo uproszczony schemat logiczny realizowanych działań, wyrażający się w przywołanej konstrukcji: „A powoduje B, zaś B powoduje C” (czy w skrócie: „jeśli zrealizujemy dane działanie, to problem zostanie w całości lub przynajmniej w części rozwiązany”)<sup>8</sup>. W tym miejscu warto przytoczyć szczegółowe zapisy RPO WK-P obrazujące tę sytuację:

*„Wzmocnienie konkurencyjności przedsiębiorstw z sektora MŚP wymaga wsparcia procesów inwestycyjnych. W szczególności istotne będzie w tym zakresie wsparcie MŚP w zakresie wdrażania innowacji produktowych, procesowych, organizacyjnych oraz marketingowych, tak by uzyskane wsparcie pozwoliło na osiągnięcie trwałej przewagi konkurencyjnej [...]. Realizacja PI 3c przyczyni się do wprowadzenia na rynek nowych i ulepszonych produktów/usług. Podniesiona zostanie również konkurencyjność MŚP poprzez wzrost wydajności pracy i efektywności produkcji/świadczonych usług”<sup>9</sup>.*

Założenia, po spełnieniu których uda się uzyskać założone zmiany, nie zostały nigdzie wskazane – a więc a priori przyjmuje się, że będą prawdziwe. Sytuacja ta dotyczy zarówno efektów oczekiwanych, jak również potencjalnych skutków niezamierzonych. Jeśli jednak przyjąć, że osiągnięcie celów danej interwencji faktycznie zależy od innych czynników, jak również warunkowana jest samą konstrukcją danego instrumentu, w tym sposobem jego wdrożenia, to wiedza na temat tego, jakie założenia muszą zostać spełnione, by cel podjętego działania został osiągnięty, jest tak naprawdę kluczowa patrząc z punktu widzenia oceny danego instrumentu. Ponadto wiedza ta jest niezbędna do wyjaśnienia przyczyn ewentualnego niepowodzenia danej interwencji oraz warunkuje możliwość postawienia trafnych rekomendacji, które umożliwią wprowadzenie niezbędnych korekt w programach projektowanych w przyszłości.

Z tego też powodu, jednym z podstawowych zadań stojących przed ewaluacją jest konieczność ustalenia logiki analizowanej interwencji, wskazanie krytycznych założeń, które warunkują jej poprawne działanie oraz ocena na ile te założenia mogą być (w przypadku oceny o charakterze ex ante) lub zostały (w przypadku oceny o charakterze ex post) spełnione. Analiza założeń powinna dotyczyć po pierwsze występowania zależności pomiędzy elementami zaplanowanej interwencji a oczekiwanymi efektami. Po drugie kluczowe jest zweryfikowanie założeń związanych z czynnikami zewnętrznymi, mającymi potencjalny wpływ na funkcjonowanie danej interwencji. Wreszcie należy rozważyć możliwość wystąpienia skutków niezamierzonych, wywołanych przez dane działanie.

8 Bienias S., Herman-Pawłowska K., (red.) *Ocena wpływu regulacji. Poradnik dla administracji publicznej*, Warszawa 2015.

9 Regionalny Program Operacyjny Województwa Kujawsko-Pomorskiego na lata 2014-2020.

## Rekonstrukcja teorii programu

Proces odtwarzania logiki interwencji ma w dużej mierze charakter koncepcyjny. Składa się na niego przede wszystkim ustalenie kompletnego mechanizmu lub mechanizmów danego programu/działania ze wskazaniem wszystkich założeń krytycznych dla jego funkcjonowania. Warto mieć na względzie, że proces ten ma zazwyczaj charakter iteracyjny – ostateczny model logiczny wypracowywany jest w trakcie trwania całego badania. W tym czasie może być on wielokrotnie aktualizowany, m.in. w wyniku uzgodnień z kluczowymi aktorami zaangażowanymi w funkcjonowanie danego instrumentu. Docelowo model logiczny (teoria interwencji) powinien mieć aprobatę wszystkich stron, których dane rozwiązanie dotyczy.

Koncepcje realizacji badań z wykorzystaniem podejścia opartego na teorii rozwijane były przez ponad 40 ostatnich lat. W tym czasie doszło do wyodrębnienia co najmniej kilku specyficznych podejść metodycznych. Wśród nich wyróżnić można: ewaluację kierowaną teorią – *theory-driven evaluation*<sup>10</sup>; ewaluację opartą na teorii – *theory-based evaluation*<sup>11</sup>; oraz ewaluację realistyczną – *realistic evaluation*<sup>12</sup>. Pomimo pewnych różnic w sposobie akcentowania wybranych elementów podejścia, koncepcje te mają wspólny mianownik w postaci traktowania każdej interwencji publicznej jako teorii<sup>13</sup>. Podejścia te zbliżone są również pod względem metodyki odkrywania, czy też odtwarzania teorii programu. Można wyróżnić co najmniej trzy kluczowe elementy tego procesu:

1. Tworzenie łańcucha przyczynowego interwencji (teorii interwencji/programu) oraz identyfikacja założeń, które warunkują jej skuteczność.
  2. Identyfikacja kluczowych założeń interwencji.
  3. Analiza kontekstu realizacji interwencji.
  4. Identyfikacja obszarów/ wymiarów, w których efekty działań publicznych mogą się różnicować.
- Poniżej kroki te zostały szczegółowo opisane, z wykorzystaniem przywołanego już przykładu pochodzącego z Regionalnego Programu Operacyjnego Województwa Kujawsko-Pomorskiego.

*Tworzenie łańcucha przyczynowego interwencji (teorii interwencji) oraz identyfikacja założeń, które warunkują jej skuteczność.*

Każda interwencja składa się z określonej sekwencji działań przypisanych do różnych podmiotów zaangażowanych w jej realizację. Punktem wyjścia jest najczęściej pewien zdiagnozowany problem, który spotyka się z reakcją (interwencją) ze strony państwa. Z tego powodu, co do zasady, to właśnie działania publiczne nadają bieg zdarzeniom, które mają zaowocować

10 Chen, H. T., Rossi, P. H., *The multi-goal, theory-driven approach to evaluation: A model linking basic and applied social science*, Social Forces, 59/1980, s. 106-122.

11 Weiss, C. H., *Nothing as practical as good theory: Exploring theory-based evaluation for comprehensive community initiatives for children and families*, [w:] Connell J., Kubisch A., Schorr L. B., & Weiss C. H. (red.), *New approaches to evaluating community initiatives: Volume 1, concepts, methods, and contexts*, NY: Aspen Institute, New York 1995.

12 Pawson, R., Tilley, N., *Realistic Evaluation*, Sage Publications, London 1997.

13 Bienias S., Herman-Pawłowska K., (red.) *Ocena wpływu...*, op. cit.

wystąpieniem określonych zmian w rzeczywistości społeczno-gospodarczej. Sekwencja tych zdarzeń tworzy tzw. łańcuch przyczynowy, który łączy zaplanowane działania z oczekiwanymi efektami interwencji<sup>14</sup>. Na etapie odtwarzania logiki interwencji konieczne jest wskazanie, jaki problem ma zostać rozwiązany, jak krok po kroku zaplanowano realizację interwencji, w tym jakie mają być jej kroki pośrednie oraz ich efekty. Te ostatnie można podzielić na efekty pośrednie oraz końcowe, będące odzwierciedleniem pożądanej zmiany, wywołanej przez dany instrument. Istotne jest, aby wskazane zostały wszystkie kluczowe czynniki, czy też przyczyny oczekiwanych zmian, na każdym etapie rozpatrywanego łańcucha przyczynowego. Zmiany mogą dotyczyć w szczególności sposobu działania, czy też funkcjonowania jednostek/podmiotów, które dana interwencja obejmuje. Caroll Weiss, wskazuje, że kluczowe jest wyróżnienie wszystkich kluczowych mini-kroków, które tworzą charakterystyczny łańcuch przyczynowy, w którym jedne zdarzenia, są konsekwencją innych zdarzeń, wynikających z wprowadzenia danego rozwiązania<sup>15</sup>.

W praktyce badawczej podstawowym źródłem danych do ustalenia takiego łańcucha powinien być w pierwszym kroku sam program – dokument opisujący cele i założenia interwencji. Najprawdopodobniej mechanizm przyczynowy opisany na jego podstawie będzie niekompletny. Doświadczenia pokazują, że w przypadku programów operacyjnych finansowanych ze środków unijnych zapisy programowe są dosyć ogólne i znajdują doprecyzowanie dopiero w innych dokumentach, takich jak np. kryteria wyboru projektów<sup>16</sup>. W przypadku programów, logika interwencji najczęściej ukryta jest pod powierzchnią sformułowanych celów i przyjmowanych sposobów działania<sup>17</sup>. Dlatego też, w celu jej odkrycia konieczna jest dodatkowo analiza wszelkiego rodzaju dokumentów o randze strategicznej – wskazujących na zakładane kierunki rozwoju w danym obszarze problemowym. Oprócz tego pomocne będą wcześniejsze ekspertyzy, badania i opracowania poświęcone koncepcjom wprowadzenia określonych instrumentów. Na podstawie tego typu materiałów, możliwe będzie poznanie szczegółowych intencji stojących za podjęciem decyzji o konieczności wprowadzenia danego programu publicznego. Wreszcie, po dokonaniu analizy dostępnych danych zastanych konieczne jest skontaktowanie się z autorami danej interwencji oraz osobami, które miały istotny wpływ na jej kształt. Celem badania powinna być próba ustalenia wszystkich „niewyartykułowanych” wprost w programie zamierzeń i powiązań przyczynowo-skutkowych pomiędzy jej poszczególnymi elementami. Na tym etapie wskazane jest ustalenie, które z działań analizowanej logiki mają charakter kluczowy (są konieczne do wystąpienia zakładanej zmiany), a które tylko pomocnicze. Ustalenia tego etapu analizy prezentuje się najczęściej w formie grafu lub schematu, na którym zaznaczone są wszystkie bloki ujęte w ramach określonej sekwencji zdarzeń.

14 White H., *Theory-Based Impact...*, op. cit.

15 Weiss, C. H., *Nothing as practical...*, op. cit.

16 IBS, *Ewaluacja zmian w systemie wyboru projektów w ramach wybranych działań PO IG 2007-2013*, raport na zlecenie Ministerstwa Infrastruktury i Rozwoju, Warszawa 2013.

17 Górniak J., *Ewaluacja w cyklu...*, op. cit.



Poniżej przedstawiono ogólny graf<sup>18</sup> dla przywołanego wcześniej przykładu Działania 1.6 RPO WK-P. Biorąc pod uwagę zapisy Programu, Szczegółowego Opisu Osi Priorytetowych<sup>19</sup> oraz kryteriów wyboru projektów, można wskazać, że wsparcie udzielane przedsiębiorstwom w postaci dotacji i instrumentów finansowych ma doprowadzić do inwestycji rozwojowych w firmach, realizowanych poprzez zakup środków trwałych, w tym maszyn i sprzętu służącego do produkcji. Oprócz tego oczekuje się, że część firm podejmie się wdrażania wyników zrealizowanych wcześniej prac B+R. Działania te będą skutkować opracowaniem nowych innowacyjnych produktów lub usług lub zmodyfikowany zostanie istotnie proces produkcyjny w firmie. Jednym z oczekiwanych efektów jest również wzrost zatrudnienia we wspieranych przedsiębiorstwach<sup>20</sup>. Wdrożone i prowadzone na rynek innowacje oraz udoskonalone procesy powinny przełożyć się wprost na konkurencyjność przedsiębiorstw, wyrażaną przez wzrost wydajności pracy i efektywności produkcji lub świadczonych usług. Poprawienie konkurencyjności przedsiębiorstw powinno przekładać się z kolei na poprawę wyników finansowych firmy, co pozwoli na jej trwały rozwój i potencjalnie dalszy, trwały wzrost zatrudnienia. Realizacja projektów w ramach działania powinna wpisywać się w całość interwencji realizowanej w ramach pierwszej osi priorytetowej RPO WK-P, polegającej na podnoszeniu konkurencyjności gospodarki województwa kujawsko-pomorskiego. Innymi słowy realizacja modelu logicznego interwencji w praktyce, powinna prowadzić do rozwiązania problemów, które były przyczyną włączenia danego działania do RPO WK-P.

**Schemat 1. Model przyczynowy wsparcia w ramach Działania 1.6 RPO WK-P**



Źródło: Opracowanie własne.

18 Alternatywnym rozwiązaniem jest przedstawienie logiki interwencji wykorzystując narzędzie matrycy logicznej (Kierzkowski, 2002).

19 Szczegółowy Opis Osi Priorytetowych Regionalnego Programu Operacyjnego Województwa Kujawsko-Pomorskiego na lata 2014-2020, wersja z dnia 27 stycznia 2016 r.

20 Wynika to z przyjętych kryteriów wyboru projektów dla Poddziałania 1.6.2: [http://www.mojregion.eu/files/dokumenty%20rpo/konkursy\\_nabory/1.6.2\\_Dotacje\\_dla\\_innowacyjnych\\_MSP\\_nr\\_28\\_11\\_07\\_2016/KM\\_2016.06.02\\_nr\\_28.2016\\_1.6.2.pdf](http://www.mojregion.eu/files/dokumenty%20rpo/konkursy_nabory/1.6.2_Dotacje_dla_innowacyjnych_MSP_nr_28_11_07_2016/KM_2016.06.02_nr_28.2016_1.6.2.pdf).

Opisany powyżej model opisuje idealną sytuację, w której zakłada się, że poszczególne działania i oczekiwane rezultaty następują po sobie, pozostając względem siebie w bezpośrednich, przyczynowych relacjach. W rzeczywistości powyższa teoria będzie mogła się zmaterializować jedynie w określonych warunkach i przy spełnieniu dodatkowych założeń. Jak wspomniano założenia te nie zostały przedstawione wprost w RPO WK-P, ani w innych dokumentach bezpośrednio z nim powiązanych. Dlatego też elementem procesu odtwarzania teorii interwencji musi być identyfikacja i przedstawienie założeń, których spełnienie jest konieczne z punktu widzenia materializacji przyjętej teorii zmiany. W przedstawionym przykładzie owe założenia zostały wskazane na grafie w postaci bloków z kolejnymi numerami. Łącznie wyróżniono 6 założeń, przy czym należy zaznaczyć, że na potrzeby niniejszego artykułu zostały one przedstawione na dosyć ogólnym poziomie. W przypadku faktycznego badania, poziom szczegółowości każdego z założeń powinien być zwiększony. Poniżej opisano wszystkie zaznaczone na schemacie założenia, wskazując na kluczowe kwestie wymagające rozważenia w toku właściwej ewaluacji.

#### *Identyfikacja kluczowych założeń interwencji*

##### *Założenie 1. Kryteria selekcji umożliwiają wybór projektów o założonych cechach oraz podmiotów mających faktyczny potencjał do ich realizacji*

Pierwsze z wyróżnionych założeń (na grafie oznaczone cyfrą 1) dotyczy „serca” każdej interwencji, jakim jest system selekcji i kryteria wyboru projektów. W praktyce kryteria wyboru projektów są ważnym elementem każdej logiki interwencji, decydują bowiem kto i na co tak naprawdę otrzyma wsparcie. Na etapie ich opracowywania dokonuje się faktyczna operacjonalizacja oczekiwań względem projektów, które zostaną dofinansowane w ramach danego instrumentu. Ich sporządzenie rozpoczyna również swoistą grę między instytucjami udzielającymi wsparcia, a podmiotami, które o to wsparcie się ubiegają. Najczęściej jest to gra o pieniądze, a jej regułami są właśnie zasady wyboru projektów. Widoczne jest to zwłaszcza w przypadku programów wsparcia przedsiębiorstw. Firmy, co do zasady kierują się w swym działaniu maksymalizacją zysku, w związku z czym ich interesy mogą (choć nie muszą) różnić się od interesów podmiotów publicznych. Skrajnym przykładem takiej sytuacji bywają programy wsparcia działalności badawczo-rozwojowej (B+R). Obecnie w większości programów operacyjnych finansowany jest zakup sprzętu, aparatury i urządzeń na potrzeby prowadzenia prac B+R przez firmy (Priorytet inwestycyjny 1b: Promowanie inwestycji przedsiębiorstw w B+I...). Z kolei przedsiębiorcy często widzą w tego typu instrumentach szansę na odnowienie własnego parku maszynowego, który służyć ma do celów produkcyjnych (niekoniecznie związanych z działalnością B+R). Aby otrzymać dofinansowanie przedsiębiorcy wskazują oczywiście, że zakupiony sprzęt zostanie wykorzystany zgodnie z intencjami grantodawcy, jednak w praktyce planuje się jego faktyczne wykorzystanie głównie na potrzeby prowadzenia standardowej działalności fir-

my. W rezultacie, co prawda podniesiony może zostać potencjał firm do tworzenia innowacji, ale równocześnie może być on w praktyce niewykorzystany. W przypadku rozważanego instrumentu w ramach Działania 1.6 RPO WK-P tak skrajny przypadek nie wystąpi, bowiem w istocie przewiduje się, że dzięki udzielanemu wsparciu firmy będą inwestować m.in. w maszyny i sprzęt służący zwiększeniu wydajności produkcji (Priorytet inwestycyjny 3c). Niemniej opisany powyżej problem również może być obecny, mając przy tym nieco bardziej subtelny charakter. Firmy mogą bowiem w ramach zaplanowanych inwestycji dokonywać inwestycji, które docelowo nie są w pełni powiązane z przedstawianą we wniosku innowacją. W praktyce wdrożenie zgłaszanej w ramach wniosku innowacji produktowej lub usługowej, może być jedynie pretekstem do odnowienia parku maszynowego, który w przyszłości planuje się wykorzystać na inne cele niż to co zostało wskazane we wniosku o dofinansowanie. Kwestia ta oczywiście powinna podlegać weryfikacji na etapie kontroli projektu, jednak jest to zagadnienie dosyć złożone (co zrobić np. w przypadku gdy zakupiony sprzęt w 30% służy do produkcji zgłoszonej innowacji, a w pozostałych 70% do produkcji, która funkcjonowała również przed otrzymaniem wsparcia). Sednem problemu jest więc zapewnienie takiego systemu selekcji projektów, który będzie w stanie wybrać projekty, które służą w całości wdrożeniu i produkcji innowacyjnych rozwiązań w firmie oraz eliminują po stronie przedsiębiorstw zachowania oportunistyczne, ukierunkowane wyłącznie na pozyskanie wsparcia.

Z drugiej strony nawet jeśli firmy będą miały szczerze intencje wdrożenia rzeczywistych innowacji do własnej praktyki biznesowej, może się okazać, że ich kompetencje i potencjał do osiągnięcia tego celu są niewystarczające. Oprócz tego proponowana przez nie innowacja może nie mieć szans na faktyczne wprowadzenie na rynek i finalnie znalezienie odpowiedniego popytu. Oczywiście system selekcji musi maksymalizować szansę sukcesu poszczególnych pomysłów przedsiębiorstw, wybierając z jednej strony firmy z faktycznym potencjałem do wdrożenia proponowanych rozwiązań, a z drugiej wybierając takie projekty, które mają rzeczywiste szanse osiągnąć sukces rynkowy – jest to jednak zadanie szalenie trudne. Dotychczasowe doświadczenia z realizacji programów strukturalnych pokazują, że projekty często pozostają idealne jedynie na papierze, na etapie aplikowania o wsparcie. Systemy selekcji projektów bywają skuteczne jeśli chodzi o wybór dobrze opracowanych wniosków – wpisujących się w wymagania formalne danego konkursu – jednak często zawodzą jeśli chodzi o wybór dobrych projektów. Firmy w toku ubiegania się o wsparcie w kolejnych programach operacyjnych, nauczyły się skutecznie aplikować o wsparcie (tj. pisać poprawnie projekty). Trzeba również pamiętać, o powstaniu całej sieci wyspecjalizowanych firm doradczych pomagających przedsiębiorcom ubiegać się o dofinansowanie. W wielu przypadkach projekty pisane są w taki sposób by maksymalizować szansę otrzymania wsparcia, a nie szansę sukcesu rynkowego danego pomysłu przedsiębiorcy. W konsekwencji, do komisji oceny projektów trafiają projekty, którym trudno jest cokolwiek zarzucić, jednocześnie ich realizacja i faktyczne efekty po przyznaniu wsparcia

pozostawiają wiele do życzenia. Sprawę dodatkowo komplikuje fakt przejrzystości oceny, która – jak pokazują badania – często jest przedkładana przez instytucje realizujące konkursy ponad innowacyjność projektów<sup>21</sup>. W rezultacie systemy selekcji projektów są zachowawcze i nie premiuje dodatkowo faktycznych innowacji.

Dobry system wyboru projektów jest pierwszym krytycznym warunkiem stojącym na drodze do osiągnięcia celów każdej interwencji. Można to założenie podsumować, iż system selekcji projektów musi prowadzić do wyboru takich przedsięwzięć, które wpisują się w całościową logikę programu biorąc pod uwagę specyfikę wspieranych grup – w tym ich potencjał – oraz charakter realizowanych projektów. Dlatego też, ewaluacja odwołująca się do logiki interwencji, powinna w pierwszym kroku objąć analizą system wyboru projektów i zestaw wykorzystywanych w ocenie kryteriów. Przedmiotem analizy powinny być takie kwestie jak sposób organizacji systemu wyboru, zasady pracy ekspertów oceniających wnioski (ich kompetencje, motywacja i warunki pracy – w tym czas na dokonanie oceny poszczególnych wniosków). Należy również zweryfikować same kryteria wyboru – czy uwzględniono w nich wszystkie kluczowe z punktu materializacji logiki interwencji elementy oraz czy umożliwiały jakościową ocenę projektów i wnioskodawców.

*Założenie 2. W interwencji zastosowano elastyczne, adekwatne dla projektów innowacyjnych (tj. obciążonych wyższym ryzykiem niepowodzenia) warunki wdrażania*

Wdrażanie innowacji w firmie zawsze wiąże się z określonym ryzykiem. Nikt nie ma większej świadomości tego faktu niż firmy, które ponoszą ewentualne koszty nieudanych inwestycji. Wdrażanie innowacji wymaga również dużej elastyczności – stałego obserwowania otoczenia i bieżącego reagowania na zmieniającą się sytuację np. podmiotów stanowiących konkurencję. To co było innowacją wczoraj, jutro może być standardowym produktem oferowanym przez większość podmiotów działających na danym rynku. System wdrażania programów operacyjnych musi uwzględniać ten fakt oraz zapewniać odpowiedni poziom elastyczności w zakresie realizacji projektów, zorientowanych na wprowadzenie na rynek innowacyjnych produktów lub usług. W przeciwnym razie, w przypadku inwestycji realizowanych zazwyczaj na przestrzeni kilku lat, może się okazać że na etapie starania się o wsparcie projekty miały w istocie innowacyjny charakter lecz na etapie jego rozliczania, jest to już kwestia dyskusyjna.

Co więcej „sztywność” procedur w programach europejskich, może w praktyce sprzyjać zachowawczym działaniom przedsiębiorców. Ich wnioski o dofinansowanie i proponowane pomysły mogą mieć tylko pozornie innowacyjny charakter, tak by z jednej strony spełnić wymagania formalne, a z drugiej minimalizować ryzyko porażki projektu i ewentualnej konieczności

<sup>21</sup> WYG PSDB Sp. z o.o. Ocena wpływu Programu Operacyjnego Innowacyjna Gospodarka na zwiększenie innowacyjności przedsiębiorstw, raport na zlecenie Ministerstwa Infrastruktury i Rozwoju, 2014.

zwrotu środków. Osoby projektujące programy finansowane ze środków europejskich zakładają, że bezwrotne lub udzielane na korzystnych warunkach wsparcie finansowe, jest wystarczającym bodźcem do tego by firmy realizowały inwestycje o określonych parametrach (w przypadku Działania 1.6 RPO WK-P, prowadzących do wypracowania i wprowadzenia na rynek nowych lub istotnie ulepszonych produktów i usług). Tymczasem przedsiębiorstwa kierują się w pierwszej kolejności rachunkiem zysków i strat, w którym pierwszy czynnik jest maksymalizowany. Niska elastyczność wdrażania projektów i ograniczone możliwości wprowadzania zmian do projektu na etapie jego realizacji sprawiają, że z punktu widzenia przedsiębiorstw planujących wdrożyć przełomowe innowacje, otoczenie realizacji programów finansowanych ze środków europejskich nie jest bezpieczne, a przez to mało atrakcyjne. Firmy będą więc tak planować inwestycje by minimalizować ryzyko porażki i konieczność zwrotu środków. Należy jednak mieć świadomość, że odbywać będzie się to kosztem poziomu innowacyjności zaplanowanych wdrożeń.

W ewaluacji kwestia ta powinna być zweryfikowana – w jakim stopniu system realizacji Programu, procedury itp. dawały elastyczność firmom przy realizacji inwestycji, a na ile mogły być czynnikiem mającym wpływ na przyjmowanie przez firmy strategii zachowawczych. Zweryfikowany powinien być wreszcie poziom innowacji wprowadzonych na rynek, tak żeby możliwe było określenie finalnych konsekwencji przyjętego systemu realizacji programu.

### *Założenie 3. Firmy podejmują trafne decyzje inwestycyjne, dzięki którym utworzone produkty/ usługi są ekonomicznie opłacalne*

Efektom wprowadzonych na rynek nowych produktów i usług powinno być podniesienie konkurencyjności przedsiębiorstw, wyrażane przez wzrost wydajności i efektywności firm oraz towarzysząca tym procesom poprawa kondycji finansowej. Oczekiwane zmiany interwencji zakładają więc, że poczynione w przedsiębiorstwie inwestycje doprowadzą do utworzenia innowacji, które znajdą na rynku faktycznych nabywców oraz będą generowały dodatni wynik finansowy. Co do zasady przewidywanie popytu na dany produkt czy usługę nie jest zadaniem łatwym – zwłaszcza w przypadku rozwiązań faktycznie innowacyjnych, dla których trudno szukać odpowiedników na rynku. Zakładany w Działaniu 1.6 RPO WK-P wzrost konkurencyjności przedsiębiorstw wymaga jednak, by poczynione inwestycje skończyły się faktycznym sukcesem rynkowym.

Częściowo problem ten weryfikowany jest na etapie analizy założenia pierwszego i drugiego. Jednak nawet jeśli uda się stwierdzić, że: a) system selekcji działań poprawnie (tj. doprowadził do wyboru firm z odpowiednim potencjałem i projektami wpisującymi się w założenia Programu); zaś b) system realizacji był zaprojektowany adekwatnie do danego typu projektów (tzn. był wystarczająco elastyczny), to nie oznacza to jeszcze, że zrealizowane projekty zakończą się sukcesem rynkowym. Nawet najlepszy system wyboru projektów nie pozwoli w całości postawić się ekspertom oceniającym wnioski w roli przedsiębiorcy, by móc zweryfikować, czy dana inwestycja jest optymalna z punktu widzenia rozwoju konkretnego podmiotu gospodarczego.

To przedsiębiorcy znają najlepiej własne firmy, jej potrzeby, jak również branżę, w której działają, w tym swoich potencjalnych konkurentów. W praktyce programy będą więc w takim stopniu skuteczne, w jakim to przedsiębiorcy będą w stanie podejmować świadome oraz trafne decyzje inwestycyjne. Jeśli firmy popełnią na tym etapie błędy (źle ocenią ryzyko rynkowe lub nie przewidzą z odpowiednim wyprzedzeniem zmieniających się trendów lub sytuacji gospodarczej), może się okazać, że wdrożone innowacje nie przyniosą zakładanych rezultatów, w tym poprawy ich wyników finansowych. Firmy decydując się na określone inwestycje, mają swoją własną „teorię inwestycji”, na końcu której znajduje się zazwyczaj lepszy wynik finansowy. Osiągnięcie tego celu, zależy jednak od spełniania wielu założeń, dotyczących m.in. wielkości popytu na ich produkty i usługi, jak również warunków prowadzenia biznesu (otoczenia regulacyjnego i wprowadzanych w nim modyfikacji). Należy pamiętać, że projekty finansowane ze środków europejskich nie są zazwyczaj jedynymi inwestycjami realizowanymi w firmach. Równolegle mogą im towarzyszyć inne działania, które mają również wpływ na kondycję finansową przedsiębiorstw. W skrajnym przypadku może się okazać, że złe wyniki wynikają z błędnej całościowej strategii inwestycyjnej firmy (prowadzącej np. do przeinwestowania, utraty płynności itp.).

W toku ewaluacji konieczne jest więc z jednej strony analiza procesów, stojących za decyzjami firm o realizacji inwestycji – czy bazowały one na wcześniej opracowanych strategiach rozwoju, uwzględniały trendy wynikające z analizy rynku, potencjalnego popytu. Warto również przeanalizować, czy a jeśli tak to jakie inne inwestycje im towarzyszyły. Oprócz tego należy zweryfikować jakie są materialne skutki projektów już po ich wdrożeniu – czy wprowadzone na rynek produkty i usługi znajdują klientów, czy odpowiadają one założeniom deklarowanym na poziomie złożonego wniosku o dofinansowanie. Weryfikacja logiki interwencji powinna finalnie pozwolić odpowiedzieć na pytanie, czy firmy aplikując o wsparcie, a następnie realizując projekty podejmowały właściwe decyzje inwestycyjne.

*Założenie 4. Wielkość inwestycji jest wystarczająca duża, biorąc pod uwagę skalę działalności firmy*

Jako uzupełniające do założenia 3 należy wskazać założenie dotyczące skali zrealizowanych inwestycji oraz wielkości sprzedaży wypracowanych w ramach działania innowacji. Znacząca poprawa konkurencyjności firm, wymaga najczęściej znaczących rozmiarów inwestycji, dzięki którym udaje się uzyskać faktyczny sukces produktu lub usługi na rynku. Program zakłada nie wprost, że udział przychodów ze sprzedaży wprowadzonych innowacji, będzie stanowił istotny udział w dotychczasowej działalności przedsiębiorstwa – w przeciwnym razie trudno mówić o znaczącym podnoszeniu konkurencyjności firmy. Może się jednak okazać, że przyznane dofinansowanie i zrealizowane dzięki nim inwestycje, stanowią zaledwie ułamek prowadzonej przez firmy działalności. W takiej sytuacji brak widocznych efektów wsparcia nie powinien dziwić.

Z tego powodu ewaluacja powinna sprawdzić, jaka jest skala zrealizowanych dzięki wsparciu inwestycji – w porównaniu do inwestycji finansowanych z innych źródeł oraz całościowej działalności przedsiębiorstwa. Weryfikacji powinien podlegać również udział uzyskiwanych przychodów ze sprzedaży nowych produktów i usług w przychodach ogółem wspartych podmiotów.

*Założenie 5. Wdrożenie nowych rozwiązań (innowacji) nie wyklucza wzrostu zatrudnienia*

Element logiki interwencji wskazuje, że na etapie selekcji premiiowane będą projekty prowadzące do wzrostu zatrudnienia w firmach<sup>22</sup>. Warto zauważyć, że oczekiwane efekty wsparcia mogą wydawać się z tego powodu wewnętrznie sprzeczne. Z jednej bowiem strony zrealizowane inwestycje rozwojowe oraz wdrożone innowacje mają prowadzić do podniesienia konkurencyjności przedsiębiorstw (wyrażanych za pomocą wzrostu wydajności pracy oraz efektywności), z drugiej strony oczekuje się wzrostu zatrudnienia, który co do zasady ma negatywny wpływ na wydajność i efektywność pracy. Oba procesy – patrząc z ekonomicznego punktu widzenia – mogą więc wzajemnie się wykluczać. Oczywiście możliwe jest osiągnięcie takiego stanu, w którym rośnie zarówno wydajność firmy jak i poziom zatrudnienia, jest to jednak zadanie znacznie trudniejsze. Wiedza zdobyta w dotychczasowych badaniach pokazuje<sup>23</sup>, że programy ukierunkowane na osiągnięcie obu celów równocześnie, prowadzą w istocie do zwiększenia zatrudnienia (jest to bowiem również najczęściej element kontroli projektów w okresie trwałości) przy jednoczesnym braku rzeczywistych zmian jeśli chodzi o wymiar konkurencyjności firm. Niemniej takie właśnie założenie zostało przyjęte w rozważanej interwencji (Działanie 1.6) – tj. wdrożenie innowacji w firmie nie wyklucza wzrostu zatrudnienia, a jednocześnie sprzyja podnoszeniu wydajności pracy. W praktyce ewaluacja powinna dokonać weryfikacji relacji między tymi dwoma elementami logiki interwencji odwołując się do doświadczeń poszczególnych firm.

*Założenie 6. Otoczenie gospodarcze – sytuacja ekonomiczna, regulacje, polityka podatkowa itp. – wspierają rozwój (a przynajmniej go nie wykluczają).*

Wreszcie materializacja efektów w analizowanej interwencji wymaga spełnienia najbardziej ogólnego założenia, związanego z otoczeniem gospodarczym w jakim działają firmy. Nawet najlepsze inwestycje nie przyniosą spodziewanych efektów, jeśli ich realizacja przypadnie na moment załamania koniunktury gospodarczej. Podobnie zmiany wprowadzane w regulacjach odnoszące się do warunków prowadzenia działalności gospodarczej w sferach takich jak polityka podatkowa, czy prawo pracy, mają bezpośrednie przełożenie na to, czy zakładane

22 Wynika to z przyjętych kryteriów wyboru projektów w ramach Poddziałania 1.6.2 RPO WK-P.

23 GUS (2015). Ustalenie wartości wybranych wskaźników ekonomicznych dla beneficjentów Regionalnych Programów Operacyjnych, Programu Operacyjnego Innowacyjna Gospodarka oraz dla dobranych grup kontrolnych.



efekty zrealizowanych projektów będą miały szansę się w pełni ujawnić. Programujący interwencję zakładają co do zasady neutralny wpływ tego rodzaju czynników zewnętrznych na efekty udzielanego wsparcia. W rzeczywistości będą one jednak wzajemnie powiązane i ewaluacja analizująca efekty wsparcia musi zweryfikować charakter występujących tutaj zależności.

Podsumowując materializacja założonej logiki interwencji wymaga spełnienia dodatkowych założeń. Część z nich odnosi się do systemu realizacji interwencji, inne związane są z zachowaniem grup docelowych wsparcia, część zaś ma charakter niezależny względem systemu wdrażania i wspartych beneficjentów. Należy w tym miejscu zaznaczyć, że powyżej przedstawiona teoria zmiany dla Działania 1.6 RPO WK-P oraz identyfikacja kluczowych dla niej założeń, bazowała przede wszystkim na przeprowadzonej analizie danych zastanych, w której wykorzystano wyłącznie publicznie dostępne dokumenty i opracowania (Regionalny Program Operacyjny, Szczegółowy Opis Osi Priorytetowych, kryteria wyboru projektów, wybrane wyniki badań). Kompletna analiza dokumentów zastanych powinna brać pod uwagę również inne źródła danych – m.in. dokumenty strategiczne i diagnostyczne, inne badania tematyczne, literaturę fachową. Dodatkowo, po fazie badań „gabinetowych”, powinien zostać zrealizowany cykl jakościowych badań terenowych (warsztatów, wywiadów pogłębionych), w której odtworzona teoria skonfrontowana powinna być z głównymi interesariuszami dla rozważanej interwencji (przedstawiciele administracji i przedsiębiorstw).

#### *Analiza kontekstu realizacji interwencji*

Zrozumienie kontekstu realizacji interwencji ma kluczowe znaczenie z punktu widzenia zrozumienia potencjalnego wpływu danego działania publicznego. Jest również ważne z punktu widzenia metodyki planowania ewaluacji oddziaływania programu. Przez kontekst należy rozumieć społeczne, polityczne i ekonomiczne warunki, w jakich dana interwencja jest wdrażana. W szczególności kluczowe jest, aby zidentyfikować te elementy kontekstu, które mogą mieć potencjalny wpływ na skuteczność danego działania<sup>24</sup>. Kwestia ta częściowo powiązana jest z opisanym powyżej założeniem dotyczącym otoczenia gospodarczego. Jednocześnie istotnie poza nie wykracza, biorąc pod uwagę fakt, że interwencja może funkcjonować w różny sposób w różnych kontekstach. Dla przykładu wsparcie działalności B+R będzie najprawdopodobniej przynosiło inne efekty na obszarach wysoko rozwiniętych, posiadających wyspecjalizowaną kadrę oraz sprzyjające warunki infrastrukturalne, a zupełnie inaczej w przypadku obszarów zapóźnionych, położonych w niekorzystnym otoczeniu itd. Podobna sytuacja może wystąpić w przypadku tej samej interwencji realizowanej w tym samym miejscu, ale w innym czasie. Zrozumienie przyczyn tego zjawiska wymaga odpowiedniego zrozumienia kontekstu oraz jego znaczenia z punktu widzenia zaplanowanych efektów interwencji.

24 Bienias S., Herman-Pawłowska K., (red.) *Ocena wpływu regulacji. Poradnik dla administracji publicznej*, Warszawa 2015.



*Identyfikacja obszarów/wymiarów, w których efekty działań publicznych mogą się różnicować*

Zrozumienie kontekstu stanowi punkt wyjścia do identyfikacji zróżnicowania wpływu interwencji. Kwestia ta ma kapitalne znaczenie z punktu widzenia odpowiedniego adresowania wsparcia. Prawdopodobne są bowiem sytuacje, w których dana interwencja nie oddziałuje w sposób jednorodny na wszystkie wsparte podmioty. W skrajnych przypadkach może się okazać, że dla części jednostek, przyjęta logika interwencji rzeczywiście działa poprawnie i przynosi zakładane skutki. Dla innych może się okazać, że przyjęte założenia zawodzą. Podjęcie próby identyfikacji grup podmiotów, wśród których oddziaływanie będzie się różnicować, powinno być zatem kolejnym krokiem w ocenie wpływu wykorzystującej podejście oparte na teorii.

Oprócz powyższego, analiza zróżnicowania oddziaływania interwencji powinna brać pod uwagę aspekt terytorialny oraz czasowy wdrażanych działań. Nie bez znaczenia pozostaje bowiem moment w jakim mierzone są efekty wsparcia. Może się okazać, że początkowo źle funkcjonująca interwencja, zaczyna przynosić oczekiwane efekty dopiero po jakimś czasie od wprowadzenia (i na odwrót – interwencja może z czasem działać coraz gorzej z uwagi na wychwytywanie przez jej odbiorców istniejących luk lub niedoskonałości). Przykładem tego zjawiska jest tzw. efekt *lock-in* zaobserwowany w badaniu wsparcia udzielanego w ramach EFS w perspektywie 2007-2013<sup>25</sup>. Badanie wykazało, że wsparcie osób bezrobotnych przynosi na początku wręcz negatywne skutki, które z czasem są redukowane. Błędne wybranie momentu pomiaru może przynieść zupełnie nietrafne wnioski, stąd na etapie analizy logiki interwencji należy ustalić właściwy moment dla szacowania efektów udzielonego wsparcia.

Wreszcie identyfikacja potencjalnego zróżnicowania efektów interwencji, będzie szczególnie istotna z punktu widzenia prowadzonych analiz kosztów i korzyści. Będzie ona również warunkowała proces generalizacji wyników przeprowadzonej oceny wpływu na całą populację.

**Weryfikacja logiki interwencji**

Po opisanej powyżej fazie koncepcyjnej badania, następuje faza typowo empiryczna, polegająca na testowaniu odtworzonej w pierwszym kroku logiki interwencji. Jak zauważa Prof. Górniak, zadaniem ewaluacji jest sprawdzenie, *czy w świetle uzyskanych efektów teoria programu daje się utrzymać, czy też należy ją odrzucić*<sup>26</sup>. W tym celu powinna być podjęta próba oszacowania rzeczywistych, tj. rozumianych w kategoriach przyczynowo-skutkowych, efektów danej interwencji. Służyć do tego celu mogą analizy zrealizowane z wykorzystaniem podejścia kontrfaktycznego. Należy jednak zauważyć, że metody kontrfaktyczne nie są jedynymi, które powinny być wykorzystane na tym etapie badania. W istocie kompletna weryfikacja logiki in-

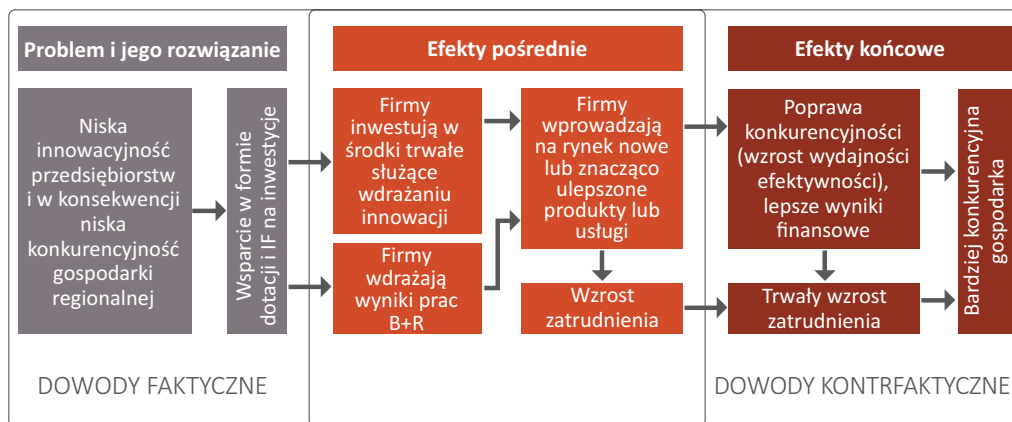
25 PAG Uniconsult, *Wpływ polityki spójności na poziom i jakość zatrudnienia w Polsce*. Raport na zlecenie Ministerstwa Infrastruktury i Rozwoju, Warszawa 2010.

26 Górniak J., *Ewaluacja w cyklu...*, op. cit.

terwencji wymaga zasilenia danymi, czy też dowodami dwojakiego rodzaju. Pierwszą grupę dowodów można określić mianem faktycznych<sup>27</sup>. Zdają one relację z tego „jak jest?” lub „jak było?” w czasie realizacji interwencji, jeśli chodzi o takie kwestie jak kształt i funkcjonowanie systemu wyboru projektów, faktyczny poziom innowacyjności rozwiązań wdrażanych w firmach, czy zakres zrealizowanych inwestycji itp. Zadaniem dowodów faktycznych jest w dużej mierze dostarczenie danych, które pozwolą zweryfikować, czy przyjęte założenia interwencji, zostały faktycznie spełnione. Po części dane te powinny pochodzić z prowadzonego monitoringu i innych analiz lub badań prowadzonych w toku realizacji interwencji. Konieczne będzie jednak przeprowadzenie osobnych badań o charakterze jakościowym (wywiadów z firmami, warsztatów, wykonania studiów przypadku itd.), które pozwolą naocznie poznać funkcjonowanie poszczególnych mechanizmów w ramach odtworzonej logiki interwencji. Co do zasady zalecane jest wykorzystanie na tym etapie badania metod mieszanych (*mix-mode*), odpowiednio dopasowanych do analizowanych zagadnień.

Z kolei dowody kontrfaktyczne zawsze odnoszą się do kwestii tego, czy dana zmiana we wspartej grupie podmiotów (wzrost zatrudnienia, przychodów itd.) wystąpiłby bez udzielenia wsparcia. Możliwe są tu do zastosowania rozmaite metody statystyczne, wykorzystujące do szacowania wpływu interwencji grupy kontrolne i odpowiednie ich dopasowanie. Jak przedstawiono na kolejnym schemacie oba typy dowodów przenikają się wzajemnie w samej logice interwencji i mają względem siebie komplementarny charakter. Konsekwencją braku któregoś z nich będzie brak możliwości weryfikacji całościowej logiki interwencji i będzie negatywnie wpływał na ostateczną użyteczność badania.

**Schemat 2. Typy dowodów w procesie weryfikacji teorii interwencji**



Źródło: Opracowanie własne.

<sup>27</sup> White H., *Theory-Based Impact...*, op. cit.

## Podsumowanie

Wymagania Komisji Europejskiej sformułowane m.in. w wytycznych dotyczących planów ewaluacji na perspektywę 2014-2020 doprowadzą niewątpliwie do większego wzrostu liczby badań, w których zbierane będą twarde dowody związane z szacowaniem wpływu realizowanych polityk publicznych. Proces ten generalnie należy ocenić pozytywnie – wpisuje się on bowiem w koncepcję prowadzenia polityk publicznych opartych na dowodach. O ile zaplanowane ewaluacje zostaną przeprowadzone faktycznie w sposób rygorystyczny, na koniec okresu wdrażania perspektywy 2014-2020 powinniśmy dysponować bogatym źródłem informacji na temat skuteczności podjętych działań projektowych.

Należy jednak pamiętać, że możliwość wykorzystania wyników badań kontrfaktycznych w procesie projektowanych polityk publicznych jest ograniczona. Mają one bowiem stosunkowo wąskie zakresowo zadanie, polegające na odpowiedzi na pytanie, „czy dana polityka działa?”. Jednocześnie nie są one w stanie wskazać ewentualnych przyczyn lub porażek interwencji publicznych. W tym celu konieczne jest zastosowanie odrębnego instrumentarium, które dostępne jest w ramach podejścia opartego na teorii. Badania je wykorzystujące skupiają się na odkrywaniu mechanizmów, które czynią projektowane działania skutecznymi. Zastosowanie obu podejść równocześnie – kontrfaktycznego i opartego na teorii – pozwala na dokonanie całościowej weryfikacji teorii wdrażanych programów. Jednocześnie daje szansę na podniesienie jakości i w konsekwencji użyteczności realizowanych badań ewaluacyjnych.

## Bibliografia

- Bienias S., Herman-Pawłowska K.,** (red.) *Ocena wpływu regulacji. Poradnik dla administracji publicznej*, Warszawa 2015
- Chen, H. T., Rossi, P. H.,** *The multi-goal, theory-driven approach to evaluation: A model linking basic and applied social science*, *Social Forces*, 59/1980, s. 106-122
- Górnica J.,** *Ewaluacja w cyklu polityk publicznych*, [w:] S. Mazur (red.), *Ewaluacja funduszy strukturalnych – perspektywa regionalna*, Uniwersytet Ekonomiczny w Krakowie, Małopolska Szkoła Administracji Publicznej, Kraków 2007
- Główny Urząd Statystyczny, *Ustalenie wartości wybranych wskaźników ekonomicznych dla beneficjentów Regionalnych Programów Operacyjnych, Programu Operacyjnego Innowacyjna Gospodarka oraz dla dobranych grup kontrolnych*, Warszawa 2015
- IBS, *Ewaluacja zmian w systemie wyboru projektów w ramach wybranych działań PO IG 2007-2013*, raport na zlecenie Ministerstwa Infrastruktury i Rozwoju, Warszawa 2013

- Kierzkowski T.**, *Ocena (ewaluacja) programów i projektów o charakterze społeczno-gospodarczym w kontekście przystąpienia Polski do Unii Europejskiej*, Warszawa 2002
- Komisja Europejska, *The Programming Period 2014-2020. Guidance Document on Evaluation Plans*, Brussels 2015
- Leeuw F.**, *Reconstructing Program Theories: Methods Available and Problems to be Solved*, American Journal of Evaluation, n. 24(1)/2003, s. 5-20
- Leeuw F.**, *Theory-based evaluation*, EVALSED, 2012
- Leeuw F.**, *Linking theory-based evaluation and contribution analysis: Three problems and a few solutions*, Evaluation, Vol. 18 no. 3/2012, s. 348-363
- Martini A.**, *Wprowadzenie do podstawowych pytań stawianych w ewaluacji oddziaływania interwencji: Czy interwencja coś zmieniła? A jeśli tak, to dla kogo?*, [w:] Haber A., Trzciński R. (red.), *Ocena wpływu i prognozowanie efektów w badaniach ewaluacyjnych*, Polska Agencja Rozwoju Przedsiębiorczości, Warszawa 2011
- Olejniczak K.**, *Mechanizmy wykorzystania ewaluacji. Studium ewaluacji śródkresowych INTERREG III*, Centrum Europejskich Studiów Regionalnych i Lokalnych UW, Wydawnictwo Naukowe SCHOLAR, Warszawa 2008
- PAG Uniconsult, *Wpływ polityki spójności na poziom i jakość zatrudnienia w Polsce*. Raport na zlecenie Ministerstwa Infrastruktury i Rozwoju, Warszawa 2010
- Pawson, R., Tilley, N.**, *Realistic Evaluation*, Sage Publications, London 1997
- Regionalny Program Operacyjny Województwa Kujawsko-Pomorskiego na lata 2014-2020, wersja z dnia 16 grudnia 2014 r.
- Szczegółowy Opis Osi Priorytetowych Regionalnego Programu Operacyjnego Województwa Kujawsko-Pomorskiego na lata 2014-2020, wersja z dnia 27 stycznia 2016 r.
- Weiss, C. H.**, *Nothing as practical as good theory: Exploring theory-based evaluation for comprehensive community initiatives for children and families*, [w:] Connell J., Kubisch A., Schorr L. B., & Weiss C. H. (red.), *New approaches to evaluating community initiatives: Volume 1, concepts, methods, and contexts*, NY: Aspen Institute, New York 1995
- White H.**, *Theory-Based Impact Evaluation: Principles And Practice*, Working Paper n. 3, International Initiative for Impact Evaluation, New Delhi 2009
- WYG PSDB Sp. z o.o., *Ocena wpływu Programu Operacyjnego Innowacyjna Gospodarka na zwiększenie innowacyjności przedsiębiorstw*, raport na zlecenie Ministerstwa Infrastruktury i Rozwoju, Warszawa 2014



### **RAFAŁ TRZCIŃSKI**

Ewaluator, analityk – absolwent Instytutu Socjologii Uniwersytetu Warszawskiego, o przekrojowym doświadczeniu w zakresie realizacji badań społecznych. Ewaluacją zajmuje się od ponad dziesięciu lat, w tym czasie realizował badania m.in. w Polskiej Agencji Rozwoju Przedsiębiorczości (2005-2014), Instytucie Badań Edukacyjnych (2014-2015), Fundacji IDEA Rozwoju (od 2004 r.) oraz Fundacji ISEE (od 2014 r.). Specjalizuje się w metodologii badań ilościowych, w tym o charakterze quasi-eksperymentalnym. Autor oraz współredaktor publikacji i artykułów w tym zakresie. Posiada wieloletnie doświadczenie trenerskie, obejmujące zagadnienia z obszaru metodyki badań ewaluacyjnych, w szczególności związane z zagadnieniami analiz statystycznych i wykorzystania metod i technik kontrfaktycznych. Członek Polskiego Towarzystwa Ewaluacyjnego.

## MONIKA KWIECIŃSKA-ZDRENKA

### Abstrakt

Ze względu na coraz większą złożoność i dynamikę rozwojową środowiska/rzeczywistości społecznej zmienia się też natura rozwoju i jego planowanie. Mechanicystyczne planowanie rozwoju jest coraz częściej postrzegane jako problematyczne. W niepewnym/nieprzewidywalnym świecie (w kategoriach gospodarczych, społecznych i politycznych) przestają sprawdzać się tradycyjne modele oceny rozwoju (czy rozwój przebiega w kierunku predefiniowanego celu).

Powoduje to, że potrzebna jest nowa generacja ewaluacji: takiej, która lepiej odzwierciedla nieprzewidywalność i złożoność interakcji w systemach; i takiej, która na nowo kładzie nacisk na innowacje, z prototypami i pilotażami, które mogą być ulepszone. Oznacza to ewaluację bardziej czułą, mogącą przechwytywać sygnały zmiany i dostosować się do szybko i stale zmieniających się okoliczności i kulturowej dynamiki. Taką ewaluację określa się jako ewaluację rozwojową (developmental evaluation, znaną też jako real-time evaluation, emergent evaluation, action/adaptive evaluation).

Ma to także konsekwencje dla podejścia do danych, które w tym ujęciu nie są tylko gromadzone, ale także wyszukiwane lub przechowywane, gromadzone w czasie rzeczywistym (w trakcie realizacji innowacji) i z udziałem samych badanych. Po drugie, oznacza to rozszerzenie metod analitycznych, np. o takie, które skupiają się na tworzeniu dalszych wytycznych rozwoju.

Celem artykułu jest przedstawienie specyfiki ewaluacji rozwojowej w porównaniu z ewaluacją tradycyjną ze wskazaniem, jakie możliwości uzupełnienia badań ewaluacyjnych ona tworzy.

**Słowa kluczowe:** ewaluacja rozwojowa, innowacyjne vs. tradycyjne metody badań

### Wprowadzenie

W wytycznych Ministra Infrastruktury i Rozwoju z września 2015 roku dotyczących ewaluacji polityki spójności na lata 2014-2020 znajdujemy zalecenie podwyższenia jakości badań ewaluacyjnych przez zwiększenie wymagań metodologicznych. Celem tego zwiększonego rygoru badawczego jest chęć uzyskania „*bardziej wiarygodnych i użytecznych informacji o (rzeczywistych) efektach polityk spójności (...) oraz odpowiedź na pytania, dlaczego dana zmiana nastąpiła, czyli wychwycenie mechanizmów oddziaływania polityki spójności*”<sup>1</sup>.

1 Wytyczne Ministra Infrastruktury i Rozwoju w zakresie ewaluacji polityki spójności na lata 2014-2020 z 22.09.2015 r. dostępne jako: [https://www.funduszeuropejskie.gov.pl/media/9318/wytyczne\\_090915\\_final.pdf](https://www.funduszeuropejskie.gov.pl/media/9318/wytyczne_090915_final.pdf). s. 8.

Należy się zgodzić, że niezbędne jest dążenie do tego, aby badania ewaluacyjne były coraz wyższej jakości, ale towarzyszyć temu powinno pytanie o cel, któremu ma służyć zwiększanie wymagań wobec badań i jak przekłada się to na poprawę wdrażania polityki spójności. W przywołanych dokumentach ministerialnych zwraca się uwagę na to, że celem ma być „zwiększenie roli i użyteczności ewaluacji w procesie decyzyjnym oraz zwiększenie transparentności polityki spójności poprzez dostarczanie interesariuszom oraz obywatelom wiarygodnych informacji dotyczących rzeczywistych efektów podejmowanych działań”<sup>2</sup>. To podejście podobne do tego, jakie prezentuje m.in. Michael Patton, który zwraca uwagę, że „w ewaluacji skoncentrowanej na wykorzystaniu nacisk kładzie się na zamierzone wykorzystanie badań przez docelowych użytkowników. (...) Ewaluacja skoncentrowana na wykorzystaniu wymaga przejścia od elementów ogólnikowych i abstrakcyjnych, tj. możliwych odbiorców i potencjalnych sposobów wykorzystania, do elementów rzeczywistych i konkretnych: faktycznych głównych użytkowników docelowych i ich wyraźnie sprecyzowanego zobowiązania do przyjęcia konkretnych, ściśle określonych sposobów wykorzystania. Ewaluacja skoncentrowana na wykorzystaniu odpowiada na pytanie o to, na czyich wartościach ewaluacja ma się opierać, poprzez współpracę z jasno określonymi głównymi użytkownikami docelowymi, odpowiadającymi za wykorzystywanie wyników ewaluacji i wdrażanie rekomendacji”<sup>3</sup>.

To, co odróżnia założenia ewaluacji skoncentrowanej na wykorzystaniu od założeń ministerialnych to, że nie faworyzuje ona „żadnego konkretnego zagadnienia, modelu, metody, teorii czy nawet wykorzystania. Stanowi ona raczej proces, którego celem jest pomoc głównym użytkownikom docelowym w wyborze zagadnień, modelu, metod, teorii i sposobów wykorzystania najbardziej odpowiednich w ich konkretnej sytuacji”<sup>4</sup>. Tymczasem dokument ministerialny wyznacza rodzaj „złotego standardu” wskazując, że zwiększanie jakości badań powinno następować przez „rozwój oraz stosowanie zaawansowanych metod badawczych, w tym w szczególności metod kontrfaktycznych oraz ewaluacji opartej na teorii, [co] stanowić będzie jeden z priorytetów ewaluacji w okresie 2014-2020”<sup>5</sup>. Próby wyznaczenia takiego „złotego standardu” są zresztą podejmowane nie tylko przez urzędników, ale także (lub głównie?) przez organizacje zrzeszające ewaluatorów. Istnieją jednak dla tego postulowanego wzoru badań alternatywy. Jedną z nich może być stosowanie podejścia ewaluacji rozwojowej, chociaż trafniejsze byłoby ujmowanie jej nie jako alternatywy dla głównego nurtu badań ewaluacyjnych w obecnej perspektywie finansowej, ale raczej jako swoistego pomostu, szpachli do załatania dziury, jaka może pojawiać się, gdy testujemy dedukcyjnie sformułowane hipotezy w procedurze eksperymentalnej.

2 Wytyczne Ministra Infrastruktury i Rozwoju w zakresie ewaluacji ..., op. cit., s. 7.

3 Patton, M. Q., *Ewaluacja skoncentrowana na wykorzystaniu*, [w:] *(R) ewaluacja. Poszukiwanie nowych metod oceny efektów*, (red.) Haber A., Popis Z., PARP, Warszawa 2013, dostępne jako: <https://www.parp.gov.pl/files/74/81/626/18724.pdf>. s. 17-18.

4 Patton, M. Q., *Ewaluacja skoncentrowana* ..., op. cit., s. 18.

5 Wytyczne Ministra Infrastruktury i Rozwoju w zakresie ewaluacji ..., op. cit., s. 8.



### **Ewaluacja rozwojowa – potrzeba, wymóg czasu czy zbyteczny wysiłek badawczy?**

Pomysł ewaluacji rozwojowej to efekt wyrażenia potrzeb dwójakiego rodzaju: potrzeby myślenia systemowego oraz przekonania o konieczności uwzględniania w badaniach perspektywy złożoności rzeczywistości.

O potrzebie myślenia systemowego. Myślenie systemowe stanowić może ramę wyjaśniającą zmianę, jaka zachodzi pod wpływem interwencji. Oparte jest ono na kilku założeniach – system złożony jest z części (podsystemów), powiązanych ze sobą wzajemnie w taki sposób, że zmiany jednych elementów wpływają na wszystkie pozostałe i na relacje między nimi, a całość jest czymś większym niż suma tych poszczególnych części. Przyjęcie tego założenia (w badaniach socjologicznych obecnego od blisko stu lat) ma konsekwencje praktyczne: zrozumienie całości procesu interwencji nie jest możliwe bez analizy kontekstu węższego i szerszego oraz analizy wzajemnych powiązań między poszczególnymi elementami (interesariuszami, ich wyobrażeniami o celach interwencji, podejmowanymi działaniami, itp.), określenia granic systemu (gdzie nie należy spodziewać się skutków działań) oraz analizy splątania różnych systemów, to jest sposobów przenikania, nachodzenia i wzajemnego oddziaływania na siebie różnych systemów<sup>6</sup>.

O potrzebie dostrzegania złożoności. Dominującą cechą rzeczywistości, w której funkcjonujemy stała się nieprzewidywalność procesów społecznych i gospodarczych i wiążąca się z tym niepewność (w reakcji na nią kilka lat temu pojawiła się między innymi fala społecznego sprzeciwu wobec dominacji gospodarki neoliberalnej nad polityką i państwem w Europie Zachodniej, USA, Afryce Północnej), która wiąże się z takimi coraz silniej uświadamianymi cechami jak nielinearność i współzmiennność (nie jest oczywiste, jakie – nawet odległe i pozornie bez znaczenia dla nas – zdarzenia mogą wywołać zmianę i w jakim kierunku się ona potoczy – jak trzepot skrzydeł motyla, który powoduje burzę piaskową w odległym miejscu i z opóźnieniem w czasie), dynamika i wyłanianie się – w wyniku złożoności rzeczywistości, kierunek zmiany nie może być ustalony z góry, co powoduje, że naszymi zachowaniami kieruje potrzeba adaptacyjności (próba dopasowania się do okoliczności, a nie podporządkowanie biegu wydarzeń naszym potrzebom czy woli). Niepewność złożonej rzeczywistości oznacza, że wszelkie plany mogą okazać się nietrafne (spóźnione, zdeaktualizowane) lub że trzymanie się planu może powodować, że przegapiane są istotne fakty i zdarzenia.

Takie wyobrażenie o rzeczywistości prowokuje charakterystyczny zestaw czynności badawczych: konieczność rozpoznawania wpływu czynników historycznych, ekonomicznych, politycznych, społeczno-kulturowych, administracyjnych, ekologicznych, klimatycznych, prawnych na kierunek interwencji i cechy systemu, w którym wprowadzana jest zmiana (wzajemne we-

6 Ramalingam, B., Jones, H., Reba, T. and Young, J., *Exploring the science of complexity: Ideas and implications for development and humanitarian work*. Working Paper 285, Overseas Development Institute, London 2008, dostępne jako <http://bit.ly/10ZeLRn>.



wewnętrzne powiązania składowych systemu, jego granice, dynamika); uchwycenie przyczynowości pośród dynamicznych, wzajemnie powiązanych, nieliniarnie przebiegających procesów zmiany; rozpoznania różnic w sposobie rozumienia celu i założeń zmiany z perspektywy różnych aktorów zmiany (mogą mieć różne interesy lub teorie wyjaśniające zmianę) ulokowanych w zróżnicowanych strukturalnie organizacjach (różniących się zasięgiem, typem działań, lokalizacją, źródłami finansowania działalności i innymi cechami, w dodatku wzajemnie na siebie oddziałującymi).

Te dwie cechy interwencji (ich ulokowanie w systemie o wysokim poziomie złożoności i sama złożoność realizowanych projektów zmiany) zmuszają do pytania o to, w jaki sposób możliwa jest ewaluacja tych procesów. Według coraz liczniejszych autorów (patrz Bamberger i in., Patton) niezbędne staje się poszukiwanie form ewaluacji, które pozwolą na uniknięcie problemów takich jak niesatysfakcjonujące wyjaśnienia źródeł zmiany lub niezdolność do ujawnienia efektów niezamierzonych podjętych działań. W odpowiedzi na ten apel najczęściej pojawiają się postulaty, aby prowadzić badania w oparciu o zasady odwołujące się do teorii, ale z wykorzystaniem mieszanych metod, a więc oferujące pewien pomysł na wyjaśnienie procesu zmiany, jednak nie zamykający się na różne paradygmaty, modele i metody wyjaśnień oraz oferujący wsparcie (zrozumienie) procesu zmiany już w trakcie jego zachodzenia. Takie walory wydaje się mieć ewaluacja rozwojowa.

### **Ewaluacja rozwojowa – czym jest i czym nie jest**

Ewaluacja rozwojowa to badanie, którego celem jest wspieranie procesu rozwoju innowacji, interwencji lub programu. Służy ono tworzeniu informacji zwrotnej w czasie rzeczywistym, to znaczy w czasie, w którym badany proces ma miejsce. Michael Patton wskazuje na pięć głównych celów i sposobów korzystania z wyników ewaluacji rozwojowej:

1. dostosowywanie projektu, programu, strategii, polityki lub innej inicjatywy rozwojowej (innowacji) do nowych warunków w złożonych, dynamicznych systemach,
2. dostosowywanie ogólnych zasad do nowego kontekstu – ewaluacja rozwojowa lokuje się pomiędzy oddolnymi i odgórnymi siłami zmiany,
3. sformułowanie natychmiastowej odpowiedzi na istotne zmiany lub kryzys (katastrofa naturalna, załamanie rynku finansowego), badanie podejmowanych rozwiązań, innowacyjne i pomocne interwencje,
4. (performatywne) rozwijanie innowacji (wyłaniających się idei i wizjonerskich nadziei) w potencjalne modele, które są lepiej skonceptualizowane i których działanie jest potencjalnie lepiej mierzalne,
5. tworzenie informacji zwrotnej na temat tego, jak przebiega zmiana, jakich zmian będzie wymagał ten proces, aby mógł zwiększyć oddziaływanie i aby mógł być powtarzalny w szerszej skali?

7 Patton, M. Q., *Developmental Evaluation: Applying Complexity Concepts to Enhance Innovation and Use*, The Guilford Publications, New York 2010, s. 21-22).

Różnice między ewaluacją tradycyjną, a rozwojową dotyczą kilku sfer:

1. celu i kontekstu, w którym warto po nią sięgać – to ewaluacja przydatna, gdy w złożonych, dynamicznym środowisku wprowadzana jest innowacja i badanie służyć ma nie testowaniu czy i jak program działa (w tym jaki jest jego wpływ na rzeczywistość), ale temu, żeby badać możliwości oddziaływania,
2. przedmiotu ewaluacji – ewaluacja rozwojowa ma dostarczać informację zwrotną, zachęcać do uczenia się i wspierać działania w ramach procesu rozwoju, który początkowo jest małej skali, a z czasem coraz większej, którego główną siłą napędową jest zmiana systemowa (a nie system postrzegany jako kontekst zmiany), który jest efektem ścierania się pomysłów oddolnych i odgórnych,
3. modelowania i metod badawczych – zastąpienie linearnych modeli przyczynowo-skutkowych mapowaniem złożonej dynamiki systemowej; zrezygnowania z badań kontryfaktycznych (w złożonej rzeczywistości jest zbyt wiele zmiennych i wyłaniających się możliwości wpływów), nastawienie na wykrywanie mechanizmów, gdy się wyłaniają (a nie efekty działania), nastawienie na to, co niespodziewane i nieprzewidywalne; nastawienie na użyteczność i pragmatyzm zarówno jeśli chodzi o podejście metodologiczne, jak i interpretację i wnioskowanie,
4. ról i relacji między badanymi i badaczem – ewaluator jest częścią zespołu realizującego interwencję, jest coachem skłaniającym do uczenia się i ułatwiającym proces, który nie rezygnuje z wartościowania, zaufanie nie jest budowane na niezależności badacza, ale na wzajemnym poszanowaniu i zrozumieniu różnic oraz przyjmowanych perspektyw i wartości,
5. wynikami i wpływem na badane procesy – wskazanie zasad kierujących praktykami, które będzie można dostosować do kontekstu lokalnego, inny sposób raportowania wyników (nastygmiasowy, zaangażowany, aktywny, aktywny głos w czasie rzeczywistym interwencji),
6. podejściem do złożoności – akceptacja niepewności i nieprzewidywalności i uczenie się reagowania w sytuacjach niekontrolowanych,
7. podejściem ewaluatorów – elastyczność, eklektyzm i adaptacyjność metodologii, zdolność do takiego budowania refleksji opartej na dowodach, która będzie użyteczna dla kształtowania działań, ale przy identycznych standardach zawodowych i etycznych, jak w podejściu tradycyjnym<sup>8</sup>.

<sup>8</sup> Patton, M. Q., *Developmental Evaluation*, op. cit., s. 23-26.

**Tabela 1. Ewaluacja tradycyjna i rozwojowa – porównanie**

Ewaluacja tradycyjna	Ewaluacja rozwojowa
Oceny sukcesów i porażek	Informacja zwrotna, nauka, wspieranie zmian
Założone cele a sukces	Nowe miary, mechanizmy monitorowania towarzyszące powstawaniu i zmianie celów
Niezależność i obiektywność ewaluatora	Ewaluacja częścią działania
Linearny model logiczny	Uchwycenie dynamiki systemu, wzajemne zależności, modele i wyłaniające się powiązania
Wyniki – możliwe do uogólnienia	Sensy uwzględniające kontekst – informacja zwrotna dla trwających innowacji
Zewnętrzne autorytety, interesariusze, płatnicy	Wartości i zaangażowanie innowatorów
Kontrola i przypisywanie odpowiedzialności	Uczenie się reagowanie na brak kontroli, reakcje strategiczne
Ewaluator kontroluje zmianę	Ewaluator współpracuje z wprowadzającymi zmianę
Jej wynikiem jest ustalenie sukcesów i porażek interwencji	Wspiera proces uczenia się

Źródło: Gamble J. A. A., 2008, s. 62.

Warto podkreślić jeszcze kilka cech ewaluacji rozwojowej, wokół których narastają rozmaite mity. Ewaluacja rozwojowa nie zastępuje innych rodzajów ewaluacji – stanowi raczej ich dopełnienie. Ewaluacja rozwojowa nie jest ewaluacją procesu, ale skupia się ostatecznie na wyniku wprowadzanej zmiany. Może sięgać po każdą odpowiednią metodę badawczą – nie opiera się wyłącznie na miękkich (jakościowych) metodach pomiarowych, subiektywnych opowieściach, nie musi (ale może) mieć charakter uczestniczący. Ewaluacja rozwojowa nie bagatelizuje problemu odpowiedzialności – ta związana jest z tym, na ile wyniki badania mogą wspierać rozwój<sup>9</sup>.

Jest to też typ ewaluacji, której nie należy stosować w wybranych przypadkach (np. gdy interwencja nie przebiega w złożonych, dynamicznych systemach lub nie dotyczy procesu rozwojowego, w którym niemożliwy jest do określenia z góry cel ani środki niezbędne do jego osiągnięcia). Sprawdzi się więc np. dla ewaluacji programów edukacyjnych (jak w warszawskim programie Wars i Sawa kierowanym do uczniów zdolnych; lub w kanadyjskim projekcie Jump – kierowanym do uczniów z dysfunkcjami), programów rozwoju lokalnego, zwłaszcza tych, które rozwiązują złożone problemy społeczne, zwiększają jakość życia, dobrostan, rozwiązują pro-

9 Gamble J. A. A., *A Developmental Evaluation Primer*, The J.W. McConnell Family Foundation, Montreal 2008, dostępne jako: <http://www.mcconnellfoundation.ca/assets/Media%20Library/Publications/A%20Developmental%20Evaluation%20Primer%20-%20EN.pdf>. s. 19-24.

blem nierówności (np. program aktywizacji młodzieży ze środowisk zaniedbanych, aktywizacja zawodowa, izolacja społeczna np. ze względu na chorobę psychiczną, biedę, przemoc rodzinną) lub problemy środowiska.

#### *Przebieg badania i jego efekty. Metody badań*

Badanie realizowane jako ewaluacja rozwojowa ma nieco inny przebieg od tego, który charakteryzuje tradycyjne badania ewaluacyjne. Inna jest rola badacza, który ma być z jednej strony zanurzony i zaangażowany w cały proces interwencji, ale jednocześnie musi zachować pewną odrębność po to, żeby jego głos miał znaczenie i żeby mógł przez ocenę podejmowanych działań wpływać na przebieg interwencji. Ewaluacja rozwojowa zaczyna się już w momencie planowania działań, towarzyszy ich realizacji i trwa po ich zakończeniu po to, żeby ocenić wpływ podjętych działań.

Jednym z pytań, które wymusza dyskusja o możliwościach stosowania ewaluacji rozwojowej jest to, czy można posługiwać się w tych badaniach tradycyjnym instrumentarium badawczym. Metodolodzy opisujący ten typ ewaluacji, pozostawiają badaczom pełną dowolność metodologiczną – można stosować każdą metodę, która da użyteczne informacje, daje wiarygodne dane. Jednocześnie jednak ze względu na to, że dane muszą być pozyskiwane na bieżąco i dotyczą zmieniającego się i złożonego podmiotu zachęca się, aby adaptować tradycyjne metody do tych wymogów. W efekcie w ewaluacji rozwojowej chętniej bazować będziemy z jednej strony na metodach empirycznych, na danych wytwarzanych przez badacza (w wersji tradycyjnej lub z użyciem nowych technologii – jak wywiady grupowe on-line), przez uczestników samego procesu zmiany (jak monitoring partycypacyjny - crowdsourcing, czyli ochotnicze wprowadzanie danych do systemu informatycznego, autonarracje, foto-blogi, foto-relacje, „keypad technologies”) oraz pozyskiwane bez świadomości ich wytwarzania przez uczestników rzeczywistości społecznej (tzw. big data, w tym także logowania do sieci, statystyki użytkowania stron internetowych).

### **Możliwość realizacji ewaluacji rozwojowej w polskiej rzeczywistości instytucjonalnej**

#### *Rozwiązania organizacyjne*

Mimo funkcjonowania w rzeczywistości złożonej i nieprzewidywalnej możliwe jest planowanie. Nie wiąże się ono jednak z wyznaczeniem sztywnych ścieżek prowadzących do celu, bo ten nie może być precyzyjnie określony, ale dostosowywanie planu do tego, jak projekt się rozwija, jakie pojawiają się szanse lub wyzwania w jego realizacji. R. Hummelbrunner i H. Jones proponują, aby przy takim planowaniu kierować się trzema wytycznymi: 1. zastąpienie planowania statycznego – planowaniem dynamicznym (plan traktowany jest jako hipotetyczny i tworzony w oparciu o teorię zmiany, nie tyle dostarcza odpowiedzi, co raczej stawia kluczowe pytania i wyraża założenia, które mają być rozstrzygane w trakcie realizacji projektu, ale muszą zostać

stworzone reguły i procedury, umożliwiające właściwe dostosowywanie planu do postępów działań); 2. zastąpienie planowania preskryptywnego – planowaniem elastycznym (raczej planowanie wielu i zróżnicowanych dróg dojścia do celu lub określenie zasad działania pozwalających ukierunkowywać samoistnie dalsze działania lub podejście ewolucyjne polegające na wdrażaniu równoległe różnych planów, które mają potencjał rozwijania się); 3. zastąpienie planowania całościowego – planowaniem zróżnicowanym, wielopoziomowym, zhierarchizowanym<sup>10</sup>.

Planowanie w praktyce będzie wiązało się raczej z wyznaczeniem kamieni milowych: momentów wyznaczonych dynamiką procesu zmiany lub z góry określonymi w scenariuszu zmiany punktami, w których dokonana będzie weryfikacja planu i kierunku jego realizacji. Będzie to przekładało się na sposób projektowania budżetu na badanie. Przypomnijmy, ewaluacja towarzyszy projektowi już od fazy jego opracowania koncepcyjnego, ale nie są w pełni przewidywalne rodzaje niezbędnych pomiarów, ponieważ nie ma z góry określonego programu działań, celu, do którego one służą. Powoduje to, że możliwe są trzy rozwiązania budżetowania tego rodzaju ewaluacji:

1. kontrakt czasowy z ewaluatorem (zakres działań pozostaje otwarty, a kontrakt dotyczy liczby godzin przepracowanych w tygodniu w okresie przeznaczonym na realizację projektu, np. 20 godzin tygodniowo w miesiącu przez 6 miesięcy trwania projektu),
2. finansowanie schodkowe (stopniowe) – długość kontraktu ewaluatora związana z kolejnymi fazami projektu i tylko na czas trwania danej fazy (np. odstępy kwartalne lub półroczne zawierania umów z ewaluatorem),
3. spekulacja i przyzwoleństwo na nieokreśloność – określone z góry 25-50% kosztów, a reszta obliczana progresywnie<sup>11</sup>.

### *Gotowość do zmiany*

Poza przyjęciem konkretnych rozwiązań organizacyjnych, ważniejsze jest rozstrzygnięcie, na ile polskie organizacje są gotowe na ewaluację rozwojową. Co miałyby być wyznacznikiem takiej gotowości? Byłyby to takie cechy, jak:

1. akceptacja dla takiej ewaluacji,
2. chęć uczenia się, radzenie sobie z porażkami, gotowość do podejmowania ryzyka,
3. gotowość do inwestowania w innowację,
4. mandat do podejmowania decyzji,

10 Hummelbrunner R., Jones H., *A guide for planning and strategy development in the face of complexity*. Overseas Development Institute, London 2013, dostępne jako: <https://www.odi.org/sites/odi.org.uk/files/odi-assets/publications-opinion-files/8287.pdf>.

11 Budgeting for Developmental Evaluation. An Interview with Michael Quinn Patton by Heather Britt. BetterEvaluation, April 2012, dostępne jako: <http://betterevaluation.org/sites/default/files/Budgeting%20for%20Developmental%20Evaluation.pdf>.

5. pewność finansowania (nieuzależniony od osiągnięcia celu),
6. doświadczenie w pracy ze złożonymi problemami,
7. chęć adaptacji i zastosowania nowego sposobu pracy,
8. niski poziom konfliktów wewnętrznych.

Czy funkcjonowanie w rzeczywistości odbiurokratyzowanej, w której podejmowane są samodzielne, adaptowane do potrzeb i przebiegu zmiany decyzje, w atmosferze otwartości i innowacji jest możliwe w polskich instytucjach? Na ile polskie instytucje są gotowe do podejmowania działań innowacyjnych i innowacyjnej ewaluacji? W dotychczasowych analizach<sup>12</sup> dominują raczej pesymistyczne wnioski, ale znajdujemy też oznaki zmian (np. zwiększanie się refleksyjności odgórnego i oddolnego w instytucjach). Te pytania zostawiamy tu jednak nierozstrzygnięte, jako wymagające osobnego studium lub przedmiot kolejnej konferencji ewaluacyjnej.

---

12 Olejniczak K. (red.), *Organizacje uczące się. Model dla administracji publicznej*. Warszawa: Wydawnictwo Naukowe Scholar, Warszawa 2012, dostępne jako: [http://www.euroreg.uw.edu.pl/dane/web\\_euroreg\\_publications\\_files/2329/organizacje\\_uczce\\_si\\_2012.pdf](http://www.euroreg.uw.edu.pl/dane/web_euroreg_publications_files/2329/organizacje_uczce_si_2012.pdf).

## Bibliografia

**Bamberger, M., J. Vaessen, E. Raimondo** (red.), *Dealing with Complexity in Development Evaluation Approach*, Sage Publications 2016

Budgeting for Developmental Evaluation. An Interview with Michael Quinn Patton by Heather Britt. BetterEvaluation, April 2012. Dostępne jako: <http://betterevaluation.org/sites/default/files/Budgeting%20for%20Developmental%20Evaluation.pdf>

**Dozois E., Langlois M., Blanchet-Cohen N.**, *DE 201: A Practitioner's Guide to Developmental Evaluation*, The J.W. McConnell Family Foundation Montreal 2010, dostępne jako: <https://morgridge.wisc.edu/documents/DE%20201%20EN.pdf>

**Gamble J. A. A.**, *A Developmental Evaluation Primer*, The J.W. McConnell Family Foundation, Montreal 2008, dostępne jako: <http://www.mcconnellfoundation.ca/assets/Media%20Library/Publications/A%20Developmental%20Evaluation%20Primer%20-%20EN.pdf>

**Hummelbrunner R., Jones H.**, *A guide for planning and strategy development in the face of complexity*. Overseas Development Institute, London 2013, dostępne jako: <https://www.odi.org/sites/odi.org.uk/files/odi-assets/publications-opinion-files/8287.pdf>

**Olejniczak K.** (red.), *Organizacje uczące się. Model dla administracji publicznej*. Warszawa: Wydawnictwo Naukowe Scholar, Warszawa 2012, dostępne jako: [http://www.euroreg.uw.edu.pl/dane/web\\_euroreg\\_publications\\_files/2329/organizacje\\_uczce\\_si\\_2012.pdf](http://www.euroreg.uw.edu.pl/dane/web_euroreg_publications_files/2329/organizacje_uczce_si_2012.pdf)

**Patton, M. Q.**, *Developmental Evaluation: Applying Complexity Concepts to Enhance Innovation and Use*, The Guilford Publications, New York 2010

**Patton, M. Q.**, *Ewaluacja skoncentrowana na wykorzystaniu*, [w:] (R) ewaluacja. Poszukiwanie nowych metod oceny efektów, (red.) Haber A., Popis Z., PARP, Warszawa 2013, dostępne jako: <https://www.parp.gov.pl/files/74/81/626/18724.pdf>

**Patton, M. Q.**, *Przyszłe trendy w ewaluacji*, [w:] (R) ewaluacja. Poszukiwanie nowych metod oceny efektów, (red.) Haber A., Popis Z., PARP, Warszawa 2013, dostępne jako: <https://www.parp.gov.pl/files/74/81/626/18724.pdf>

**Ramalingam, B., Jones, H., Reba, T. and Young, J.**, *Exploring the science of complexity: Ideas and implications for development and humanitarian work*. Working Paper 285, Overseas Development Institute, London 2008, dostępne jako <http://bit.ly/10ZeLRn>

Wytoczne Ministra Infrastruktury i Rozwoju w zakresie ewaluacji polityki spójności na lata 2014-2020 z 22.09.2015 r. dostępne jako:

[https://www.funduszeuropejskie.gov.pl/media/9318/wytoczne\\_090915\\_final.pdf](https://www.funduszeuropejskie.gov.pl/media/9318/wytoczne_090915_final.pdf)



### **MONIKA KWIECIŃSKA-ZDRENKA**

Doktor socjologii, związana naukowo i badawczo z Instytutem Socjologii Uniwersytetu Mikołaja Kopernika w Toruniu, w którym od kilku lat uczy podstaw badań ewaluacyjnych. Ma bogate doświadczenie badawcze, jest autorką lub współautorką kilkudziesięciu badań ewaluacyjnych realizowanych przez Pracownię Badań „Soma” z Torunia. Była ewaluatorem zewnętrznym m.in. programów regionalnych. Poza licznymi publikacjami, także z zakresu metodologii badań, jest też autorką i realizatorem szkoleń ewaluacyjnych.





**DARIUSZ MAJEREK**

## **Wstęp**

Mnogość programów społecznych realizowanych w ramach wsparcia płynącego z Unii Europejskiej lub środków wewnętrznych państwa, powoduje pośrednio wzrost zainteresowania technikami estymacji efektu oddziaływania interwencji na zmienną wynikową. Raporty podsumowujące projekty czy programy społeczne niestety zawierają czasem zasadnicze błędy metodologiczne w sposobie szacowania wspomnianego efektu. Główne wady wiążą się z nieeksperymentalnym charakterem danych, które są materiałem do estymacji efektów przy czynowo-skutkowych. Analizy ex-post prowadzone w ramach badań ewaluacyjnych, dotyczą często programów o charakterze wolontarystycznym, co skutkuje brakiem informacji na temat mechanizmu przyporządkowania bodźca do poszczególnych jednostek. W takim przypadku estymacja efektu interwencji oparta o różnicę średnich zmiennej wynikowej w grupach: eksperymentalnej (zastosowano bodziec) i kontrolnej (nie stosowano bodźca), jest niewłaściwa. W ten sposób wyznaczony estymator efektu jest obciążony i nie jest zgodny, co oznacza, że jego wartość może się znacząco różnić od prawdziwej wartości efektu interwencji.

Poprawnie wykonane oszacowanie wpływu bodźca na wielkość zmiennej wynikowej powinno być oparte o techniki kontrfaktyczne<sup>1</sup>.

Jedną z takich metod - powszechnie stosowaną - jest *Propensity Score Matching*. Niniejszy artykuł ukazuje metodologiczne podstawy stosowania PSM na bazie dostępnej literatury przedmiotowej oraz osobistego doświadczenia autora. Zostaną w nim przedstawione różne algorytmy dopasowania prób w ramach PSM, zalety i wady tej techniki, jak również aplikacyjny charakter w zakresie badań ewaluacyjnych. Na podstawie przykładowych danych pochodzących ze zbioru przygotowanego przez Lalonde'a<sup>2</sup> zostanie przeprowadzona ewaluacja wpływu programu społecznego przeprowadzonego w Stanach Zjednoczonych na poziom zarobków z wykorzystaniem technik PSM. Posłuży ona, jako ilustracja konieczności wykorzystywania technik kontrfaktycznych, a przy okazji będzie okazją do porównania wyników pochodzących z różnych algorytmów dopasowania.

1 Nazwa pochodzi od stanu kontrfaktycznego, który chcemy oszacować.

2 Lalonde R., *Evaluating the econometric evaluations of training programs with experimental data*, "The American Economic Review", 76(4) 1986, s. 604-620.

### Metoda eksperymentalna

Założmy, że w pewnej populacji przeprowadzamy ewaluację efektu oddziaływania pewnej interwencji. Zbiór jednostek objętych interwencją określamy grupą eksperymentalną, a pozostałe elementy grupą kontrolną. Niech  $Y_i$  będzie zmienną wynikową<sup>3</sup>  $i$ -tej jednostki,  $T_i$  zmienną informującą o ewentualnym stosowaniu interwencji<sup>4</sup> wobec  $i$ -tej jednostki (0 – brak interwencji, 1 – wystąpiła interwencja). Dodatkowo założmy, że zostały zmierzone również wartości cech towarzyszących<sup>5</sup> dla  $i$ -tej jednostki i oznaczamy je przez  $X_i$ . Należy przy tym zaznaczyć, że wartości zmiennych wchodzących w skład wektora  $X$  nie mogą zależeć od przynależności do interwencji. Wówczas możemy określić model przyczynowy Neymana-Rubina<sup>6</sup> następującym równaniem

$$Y_i = T_i Y_i(1) + (1 - T_i) Y_i(0), \quad (1)$$

gdzie  $Y_i(1), Y_i(0)$  oznaczają wartości zmiennej wynikowej  $i$ -tej z uwzględnieniem czy zastosowano wobec niej bodziec.

W ocenie wielkości efektu interwencji zależy nam na oszacowaniu różnicy  $Y_i(1) - Y_i(0)$  dla każdej mierzonej jednostki. Każdy badany element znajduje się wyłącznie w jednym z dwóch stanów (interwencja lub brak interwencji), zatem w zagadnieniu oceny wielkości efektu interwencji mamy do czynienia z sytuacją braku danych. Wyjściem z tej sytuacji jest szacowanie przeciętnego poziomu efektu oddziaływania bodźca<sup>7</sup>

$$\tau_{ATE} = E[Y(1)] - E[Y(0)], \quad (2)$$

gdzie  $E[Y(1)]$  jest wartością oczekiwaną zmiennej wynikowej, gdy wszystkie jednostki są wystawione na działanie bodźca, natomiast  $E[Y(0)]$  jest wartością oczekiwaną zmiennej wynikowej, podczas gdy wszystkie jednostki nie brały udziału w interwencji. ATE jest wielkością przeciętną efektu oddziaływania bodźca dla losowej jednostki w próbie. Z punktu widzenia ewaluacji programów bardziej nas interesuje ile wynosi przeciętny efekt oddziaływania bodźca, ale na jednostkach, do których bodziec został zastosowany<sup>8</sup>

$$\tau_{ATT} = E[Y(1)|T = 1] - E[Y(0)|T = 1], \quad (3)$$

gdzie  $E[Y(1)|T = 1]$  jest oczekiwanym przeciętnym wynikiem po zastosowaniu interwencji w grupie eksperymentalnej, a  $E[Y(0)|T = 1]$  jest oczekiwanym przeciętnym wynikiem nie stosując bodźca w grupie kontrolnej. Druga z wymienionych wielkości nieobserwowalna w rzeczywistości nazywa się stanem kontrfaktycznym.

3 Ang. *outcome*.

4 Ang. *treatment*.

5 Ang. *covariates*.

6 Znany też, jako model Royo-Rubina.

7 ATE – ang. Average Treatment Effect.

8 ATT – Average Treatment Effect on Treated.

W przypadku, gdy w przeprowadzonym badaniu dane będą zbierane z zachowaniem pierwszej i drugiej zasady randomizacji<sup>9</sup>, wówczas nieobciążonym estymatorem ATE i ATT jest różnica średnich grupy eksperymentalnej i kontrolnej. Wynika to z faktu, iż rozkłady wektora zmiennych towarzyszących w obu grupach (eksperymentalnej i kontrolnej) są takie same, przy posiadaniu dostatecznie dużej próby. Co więcej, w tym przypadku również zmienne, których badacz nie obserwuje, mają te same rozkłady. Jedyna istotna różnica pomiędzy grupami to fakt zastosowania lub braku bodźca.

Wielkości ATE i ATT można również oszacować, na podstawie modelu liniowego

$$\hat{Y}_i = \beta_0 + \beta_\tau T_i + \varepsilon_i, \quad (4)$$

gdzie  $\beta_\tau$  jest wielkością efektu, a  $\varepsilon_i$  jest zakłóceniem losowym. Nieco bogatszy model liniowy może posłużyć również do oszacowania wielkości efektu, również w przypadku, gdy badanie nie zostało przeprowadzone metodą eksperymentalną<sup>10</sup>

$$\hat{Y}_i = \beta_0 + \beta_\tau T_i + \beta_1 X_i + \varepsilon_i. \quad (5)$$

Statystyczna kontrola kowariantów ma za zadanie wydzielić wpływ interwencji na zmienną wynikową. Podejście regresyjne ma także pewne wady:

- Trudno jednoznacznie określić, czy postać modelu została poprawnie określona.
- W przeciwieństwie do technik opartych na dopasowaniu jednostek kontrolnych do eksperymentalnych, brakuje tu wyraźnego rozdziału między przygotowaniem doświadczenia a analizą wyników. W modelach regresyjnych obie te cechy są realizowane jednocześnie. Wgląd w poziom dopasowania modelu regresji, może powodować u badacza dążenie do osiągnięcia pewnego, z góry założonego, poziomu zależności zmiennej wynikowej od interwencji.
- W przypadku, gdy pewne wartości zmiennej wynikowej (binarnej) rzadko występują, przy jednoczesnej powszechności interwencji, model regresji może być niewłaściwym narzędziem oceny efektu. Ponadto sugeruje się<sup>11</sup> dobór 10 przypadków na każdą zaobserwowaną zmienną zakłócającą  $X$ .

Jak pokazał Rosenbaum<sup>12</sup> metody oparte na dopasowaniu grup kontrolnych charakteryzują się nieco mniejszym błędem szacunku.

9 I zasada – dobór elementów do próby jest losowy, II zasada – dobór jednostek, wobec których będzie stosowany bodziec jest losowy.

10 Cox D.R., *A note on weighted randomization*, "The Annals of Mathematical Statistics", 27(4) 1956, s. 1144–1151. Robinson J., *The Analysis of Covariance Under a Randomization Model*, "Journal of the Royal Statistical Society. Series B (Methodological)", 35(2) 1973, s. 368–376.

11 Peduzzi P., Concato J., Feinstein A.R., Holford T.R., *Importance of events per independent variable in proportional hazards regression analysis: II. Accuracy and precision of regression estimates*. "Journal of Clinical Epidemiology", 48, 1995, s. 1503–1510.

12 Rosenbaum P., *Covariance adjustment in randomized experiments and observational studies*, "Statistical Science", 17(3) 2002, s. 286–304.

## Metody quasi-eksperymentalne

Przeprowadzenie badań metodą eksperymentalną nie zawsze jest możliwe. Czasami względy etyczne, logistyczne, finansowe lub czasowe nie pozwalają na wykonanie eksperymentu. W badaniach ewaluacyjnych mechanizm przydzielania jednostek do interwencji najczęściej nie jest znany. Ponadto można przypuszczać, że nie jest on losowy, ponieważ charakter uczestnictwa w ewaluowanych programach jest często woluntarystyczny. Brak równości rozkładu cech towarzyszących (takich jak płeć, wiek, miejsce zamieszkania, motywacja, stan cywilny, podwyższony próg ryzyka), a które mogą mieć wpływ na zmienną wynikową oraz na przystąpienie do programu, skutkuje tzw. obciążeniem selekcyjnym

$$E[Y(0)|T = 1] - E[Y(0)|T = 0]. \quad (6)$$

W przypadku metody eksperymentalnej powyższa wielkość jest równa zero, natomiast w badaniach na podstawie danych obserwowanych może się ona istotnie różnić od zera. Gdy wielkość określona równością (6) jest równa zero, to stan kontrfaktyczny możemy oszacować wartością obserwowaną  $E[Y(0)|T = 0]$ , a w konsekwencji estymować ATT przy pomocy modelu liniowego lub różnicy średnich.

Metody quasi-eksperymentalne dążą do wyeliminowania wspomnianego obciążenia, poprzez zbalansowanie dystrybucji cech towarzyszących w obu grupach. Wszelkiego rodzaju techniki *matchingu* dążą do wyeliminowania obciążenia selekcyjnego przez dobór grupy kontrolnej w taki sposób, aby spełnione były założenia warunkowej niezależności<sup>13</sup>

$$Y(0), Y(1) \perp T | X \quad (7)$$

$$0 < P(T = 1|X) < 1. \quad (8)$$

Istotą tych założeń jest zapewnienie, że rozkłady zmiennej wynikowej zarówno w grupie eksperymentalnej, jak i kontrolnej są niezależne od bodźca, jeśli znane są wartości kowariantów oraz niemożliwe jest, aby na podstawie  $X$  jednoznacznie określić, do której grupy należy obiekt.

Pierwszą i zarazem jedną z najmniej praktycznych metod *matchingu* jest metoda łączenia ze względu na cechy. Oznacza to, że dla każdego elementu grupy eksperymentalnej poszukuje się elementu puli kontrolnej<sup>15</sup>, którego cechy towarzyszące są identyczne. Jej mało użyteczny charakter objawia się tym, iż w przypadku dużej ilości kowariantów znalezienie jednostki o takich samych charakterystykach jest bardzo trudne.

13 Ang. *Conditional Independence Assumption* lub *Strongly Ignorable Treatment Assignment*.

14 Rosenbaum P., Rubin D., The central role of the propensity score in observational studies for causal effects, "Biometrika", 70(1)1983, s. 41–55.

15 Zbiór wszystkich elementów próby, które nie zostały poddane interwencji.

Dlatego Rosenbaum i Rubin zaproponowali w swojej pracy<sup>16</sup> nieco inne podejście. Wprowadzili oni pojęcie funkcji balansującej<sup>17</sup>, która z definicji spełnia następujący warunek

$$X \perp T | b(X). \quad (9)$$

Powyższe założenie oznacza, że dystrybucja  $X$ -ów zarówno w grupie eksperymentalnej, jak i w kontrolnej jest taka sama. Funkcją *propensity score* będziemy nazywali funkcję następującej postaci

$$e(x) = P(T = 1 | X = x). \quad (10)$$

Wyraża ona prawdopodobieństwo zakwalifikowania obiektu do grupy eksperymentalnej, jeśli znamy wartości zmiennych towarzyszących. Rosenbaum i Rubin udowodnili, że *propensity score*<sup>18</sup> jest funkcją balansującą.<sup>19</sup> Pokazali też, że jeśli założenie o warunkowej niezależności jest spełnione względem  $X$ , to również względem  $e(X)$ .

Dyskusja na temat, które zmienne towarzyszące powinny być włączane do modelu szacującego wartości funkcji *PS* trwa do dziś. Możemy rozpatrywać cztery zbiory potencjalnych zmiennych:

1. Wszystkie zmienne towarzyszące;
2. Zmienne towarzyszące skorelowane z  $T$  ale nie mające wpływu na  $Y$
3. Zmienne towarzyszące skorelowane z  $Y$  ale nie mające wpływu na  $T$
4. Zmienne towarzyszące skorelowane zarówno z  $T$ , jak i  $Y$

Ostatnie badania z wykorzystaniem techniki Monte Carlo wskazują<sup>20</sup>, że zmienne wymienione w punktach 2 i 4, dają w rezultacie estymator efektu o większej precyzji. Ponieważ w praktyce trudno ocenić, które zmienne będą miały wpływ na zmienną wynikową lub czy będą skorelowane z interwencją, to stosuje się podejście uwzględniające zmienne mogące mieć potencjalny związek ze zmienną wynikową lub interwencją. Oczywiście poszukiwanie zmiennych, których warto użyć w modelu, powinno być poprzedzone dokładną analizą *desk research*. Ze względu na założenie (8) nie powinno się włączać do modelu zmiennych, które mogą jednoznacznie wskazać, do której grupy ma należeć obiekt. Przykładowo, jeśli badamy wpływ programu mającego na celu aktywizację zawodową kobiet, to nie możemy do modelu włączyć zmiennej *płeć*. Należy dodatkowo nadmienić, że w doborze zmiennych do modelu przychodzą z pomocą techniki statystyczne, takie jak: *Hit or Miss*, testowanie istotności statystycznej modeli brzegowych oraz sprawdziany krzyżowe, ale nie są one przedmiotem niniejszego opracowania.

16 Rosenbaum P., Rubin D., *The central...*, op. cit., s. 42.

17 Ang. *balancing score*.

18 W skrócie *PS*.

19 Rosenbaum P., Rubin D., *The central...*, op. cit., s. 44.

20 Austin P., Grootendorst P., Anderson G., *A comparison of the ability of different propensity score models to balance measured variables between treated and untreated subjects: A Monte Carlo study*, "Statistics in medicine", 26, 2007, s. 734-753.

## Propensity Score Matching

Ogólny schemat techniki *Propensity Score Matching* można przedstawić w następujących krokach:

1. Wyznaczenie modelu PSM;
2. Wybranie algorytmu łączenie elementów;
3. Sprawdzenie założeń;
4. Ocena jakości łączenia;
5. Estymacja efektu interwencji.

Najczęstszą praktyką w wyznaczaniu funkcji *PS* jest użycie regresji logistycznej lub probitowej  $T$  względem  $X$ . W przypadku, gdy zmienna  $T$  może przyjąć więcej niż dwa stany, modelowanie funkcji  $e(X)$  odbywa się za pomocą wielomianowej regresji logistycznej lub probitowej.

W ramach techniki *Propensity Score Matching* wyróżniamy kilka metod łączenia grupy kontrolnej z grupą eksperymentalną. Najbardziej znaną i najpowszechniej stosowaną jest metoda „najbliższego sąsiada”<sup>21</sup>. Polega ona na dopasowaniu do jednostki z grupy eksperymentalnej, takiej jednostki z puli kontrolnej, aby ich wartości *PS* były jak najbliższe. Zapewnia to, że rozkład cech towarzyszących uwzględnionych w modelu *PS* będzie zbliżony w obu grupach. Oczywiście możliwe jest dla dwóch konkretnych jednostek, aby mimo równości *PS* cechy im towarzyszące były inne, ale w ujęciu globalnym, jeśli funkcja  $e(X)$  była poprawnie określona, to rozkłady zmiennych towarzyszących powinny być zbliżone.

Łączenie elementów grupy kontrolnej metodą NN można przeprowadzić na kilka sposobów:

- z powtórzeniami - każdy element z puli kontrolnej może być wykorzystany kilka razy, jako element w grupie kontrolnej;
- bez powtórzeń - każdy element z puli kontrolnej może być tylko raz użyty w grupie kontrolnej;
- 1: 1 - jednemu elementowi grupy eksperymentalnej jest przyporządkowany jeden element puli kontrolnej najbliższy w sensie odległości *PS*;
- 1:  $n$  - jednemu elementowi grupy eksperymentalnej jest przyporządkowane  $n$  elementów puli kontrolnej najbliższych w sensie odległości *PS*.

Wpływ na to, która technika zostanie zastosowana w badaniu może mieć kilka elementów. Jeśli próba, którą dysponuje badacz jest niewielka, a dodatkowo cech towarzyszących jest stosunkowo wiele, to wybór łączenia metodą z powtórzeniami 1: 1 wydaje się być optymalny. W sytuacji, gdy badacz ma do dyspozycji pulę kontrolną znacząco przekraczającą liczebność grupy eksperymentalnej, można wówczas stosować technikę bez zwracania 1:  $n$ . Wybór algorytmu łączenia ma zasadniczy wpływ na obciążenie estymatora efektu oraz jego precyzję. Rozwiązanie ze zwracaniem 1: 1 charakteryzuje się mniejszym obciążeniem selekcyjnym ale jednocześnie mniejszą precyzją. Jest to spowodowane faktem, iż łatwiej jest znaleźć pojedyncze

21 NN – ang. Nearest Neighbor.

elementy (czasami występujące wielokrotnie) puli kontrolnej, które są „najbliższymi sąsiadami” w sensie odległości  $PS$ . Wówczas, jakość dopasowania grupy kontrolnej wyrażona obciążeniem selekcyjnym jest lepsza. Spada natomiast precyzja oszacowania efektu, ponieważ jest ona oparta o mniejszą liczbę jednostek grupy kontrolnej. Z analogicznych powodów rozwiązanie bez zwracania 1:  $n$  jest obciążone większym obciążeniem selekcyjnym, ale za to ma większą precyzję. Warto w tym przypadku nadmienić, że kolejność ustawienia elementów w grupie eksperymentalnej może mieć istotne znaczenie, dlatego zaleca się dokonać losowego przedstawienia elementów przed dopasowaniem. Niektórzy polecają też ustawić elementy grupy eksperymentalnej od najmniejszej do największej wartości  $PS$  (lub odwrotnie)<sup>22</sup>.

Zbalansowanie cech towarzyszących w obu grupach jest pochodną bliskości wartości  $PS$  łączonych jednostek. Może się jednak zdarzyć, że dla pewnych elementów objętych interwencją nie da się znaleźć dostatecznie bliskiej, w sensie odległości  $PS$  jednostki puli kontrolnej. Wówczas można zastosować technikę PSM z limitem<sup>23</sup>. W tej metodzie w grupie kontrolnej znajdują się tylko te elementy, których odległość łączenia nie przekracza pewnej, z góry ustalonej, wartości. Jeśli dla jakiegoś elementu nie znajdzie się najbliższego sąsiada o odległości nieprzekraczającej limitu, to taką jednostkę usuwa się z grupy eksperymentalnej. Niedogodnością tej metody jest konieczność arbitralnego ustalenia poziomu limitu. Najczęściej ustala się go na 0.25 odchylenia standardowego wielkości  $\text{logit } e(X)$ .

Pewną modyfikacją metody z limitem jest technika zaproponowana przez Dehejia i Wahba<sup>24</sup> zwaną metodą z promieniem<sup>25</sup>. O ile w metodzie z limitem liczba najbliższych sąsiadów branych pod uwagę była ustalona, to w metodzie z promieniem do grupy kontrolnej bierzemy wszystkie te jednostki, które spełniają warunek

$$|e_1(x_i) - e_0(x_j)| < r, \quad (11)$$

gdzie  $e_1(x_i), e_0(x_j)$  są odpowiednio wartościami  $PS$  dla grupy eksperymentalnej i kontrolnej.

Kolejną odmianą PSM jest metoda warstwowa<sup>26</sup>, w której nie następuje formalne łączenie elementów, a redukcja obciążenia selekcyjnego jest realizowana poprzez warstwowanie wartości  $PS$ . Rosenbaum i Rubin wykazali, że podział kwantylowy na kilka warstw w 90% redukuje obciążenie selekcyjne.<sup>27</sup> Oszacowanie wielkości efektu odbywa się wówczas przez uśrednienie wyników oceny efektu w poszczególnych warstwach.

22 Dehejia R., Wahba S., *Propensity Score Matching Methods for Nonexperimental Causal Studies*, "Review of Economics and Statistics", vol. 84, 2002, s. 151-161.

23 Ang. *Propensity Score Matching with Caliper*.

24 Dehejia R., Wahba S., *Propensity Score...*, op. cit., s. 9.

25 Ang. *Propensity Score Matching with Radius*.

26 Ang. *Stratification on the Propensity Score*.

27 Rosenbaum P., Rubin D., *Reducing bias in observational studies using subclassification on the propensity score*. „Journal of the American Statistical Association”, 79, 1984, s. 516–524.



Imbens<sup>28</sup> wskazuje, że jeśli się weźmie warstwy o jednakowej liczebności, to do oszacowania efektu ATE użyjemy średniej różnic średnich w poszczególnych warstwach. Natomiast do oszacowania ATT zaleca się użycia średniej ważonej różnic średnich w poszczególnych warstwach, z wagami proporcjonalnymi do liczby elementów wystawionych na działanie bodźca w danej warstwie.

Jeszcze inną metodą wykorzystującą wszystkie elementy puli kontrolnej w szacowaniu efektu interwencji jest ważona metoda PSM<sup>29</sup>. W tym przypadku każdy element próby jest użyty w ocenie efektu interwencji z pewną wagą. Wagi te są wyznaczone, jako odwrotność prawdopodobieństwa przynależności do grupy eksperymentalnej

$$\omega_i = \frac{T_i}{e_i} + \frac{1 - T_i}{1 - e_i}, \quad (12)$$

gdzie jest wartością  $PS$  dla -tej jednostki. Uwzględnienie wag w próbie badawczej spowoduje, że jednostki puli kontrolnej, które są zbliżone w sensie  $PS$  do jednostek objętych interwencją będą miały większe wagi, natomiast te dalsze – mniejsze. W ocenie efektu ATT używa się nieco innego ważenia

$$\omega_i = T_i + \frac{(1 - T_i)e_i}{1 - e_i} \quad (13)$$

Ostatnią metodą z rodziny technik PSM jest model regresji z uwzględnieniem wartości  $PS$ <sup>30</sup>. Model regresji przedstawia się wówczas następująco

$$\tilde{Y}_i = \beta_0 + \beta_\tau T_i + \beta_1 e(X_i) + \varepsilon_i. \quad (14)$$

Dopuszczalne jest też stosowanie nieliniowych form modelu regresji względem  $PS$  oraz dołączanie kowariantów.

Wszystkie wspomniane techniki z rodziny *Propensity Score Matching* wykazują pewne różnice w zbalansowaniu grup, a co za tym idzie, również w wielkości efektu interwencji. Ostatnia z wymienionych metod różni się znacząco od pozostałych pod względem realizacji oceny efektu interwencji. Metody najbliższego sąsiada, warstwowania ze względu na  $PS$ , czy ważona metoda PSM, mają wyraźnie określony podział na przygotowanie materiału porównawczego, w którym możemy przeprowadzić ocenę dopasowania grup oraz etap estymacji efektu. W metodzie regresyjnej z uwzględnieniem  $PS$  takiego podziału nie ma<sup>31</sup>.

Warunkiem stosowalności techniki PSM jest spełnienie założenia wspólnej dziedziny dla obu grup. Polega ono na tym, że aby dokonać poprawnego łączenia elementów zbioru war-

28 Imbens G., *Nonparametric estimation of average treatment effects under exogeneity: A review*. "Review of Economics and Statistics", 2004, s. 5-29.

29 Ang. *Inverse Probability of Treatment Weighting*.

30 Ang. *Covariate Adjustment with Propensity Score*.

31 Rubin D., *Using propensity scores to help design observational studies: Application to the tobacco litigation*, "Health Services and Outcomes Research Methodology", 2(3), 2001, s.169–188.

tości *PS* dla obu grup powinny mieć część wspólną. Inaczej dla pewnych jednostek objętych interwencją nie znajdziemy elementów puli kontrolnej o zbliżonej wartości *PS*. Te jednostki nie zostałyby poprawnie zbalansowane, a w konsekwencji wzrost obciążenia selekcyjnego pogorszyłby ocenę efektu.

Kolejnym ważnym elementem diagnozy, czy model PSM został poprawnie zbudowany a dopasowanie jest dostateczne, jest porównanie standaryzowanych różnic<sup>32</sup> przed i połączeniu. Różnica standaryzowana wyraża się następującym wzorem

$$SB = \frac{\bar{X}(1) - \bar{X}(0)}{\sqrt{0.5(S_{X(1)}^2 + S_{X(0)}^2)}} 100\%, \quad (15)$$

gdzie  $\bar{X}(1)$  oznacza średnią wartość kowariantu w grupie bodźcowanej,  $\bar{X}(0)$  oznacza średnią wartość kowariantu w grupie poza interwencją, a  $S_{X(1)}^2, S_{X(0)}^2$  oznaczają wariancje kowariantów odpowiednio w grupie eksperymentalnej i kontrolnej. Powyższe wartości w próbie przed i po przeprowadzeniu łączenia powinny się znacznie różnić. Czasami do badania istotności tej różnicy używa się testu. Wówczas przed łączeniem test wskazuje na istotność, natomiast po łączeniu różnice średnich wszystkich kowariantów nie powinny wykazywać istotności.

### Przykład

Dane użyte w tym przykładzie częściowo pochodzą z badania przeprowadzonego w USA w latach 1975-1980 i dotyczą programu National Support Work Demonstration, którego zadaniem było sprawdzenie, jaki jest poziom zarobków, po odbyciu stażu w zakładzie pracy chronionej. Badaniem były objęte osoby z problemami: byli przestępcy, byli narkomani, kobiety, które dłuższy czas przebywały na zasiłku socjalnym oraz osoby, które porzuciły naukę w wieku 17-20 lat. Oszacowanie efektu interwencji, w postaci stażu, zostało przeprowadzone metodą eksperymentalną, a wynik przeciętnego poziomu płac po zakończeniu programu wśród badanych wyniósł 1794\$. W roku 1999 Dehejia i Wahba<sup>33</sup> przeprowadzili badanie efektu interwencji korzystając z tej samej grupy eksperymentalnej, natomiast pula kontrolna pochodziła z dwóch spisów powszechnych *Panel Study of Income Dynamics* i *Current Population Survey*. Naturalnie rozkład cech towarzyszących w obu grupach był inny, z racji nieeksperymentalnego doboru grupy kontrolnej. Celem takiego podejścia była próba weryfikacji poziomu dokładności oszacowania efektu interwencji przy zastosowaniu PSM.

Grupa eksperymentalna składała z 185 mężczyzn objętych interwencją, a pula kontrolna zawierała dane ze spisu powszechnego PSID na temat 2490 mężczyzn w wieku poniżej 55 lat,

<sup>32</sup> Ang. *standardized bias*.

<sup>33</sup> Dehejia R., Wahba S., *Causal effects in nonexperimental studies: Reevaluating the evaluation of training programs*. "Journal of the American Statistical Association", 94(448), 1999, s. 1053–1062.

którzy nie byli jeszcze na emeryturze. Dla obu grup obserwowane były zmienne: wiek (AGE), lata edukacji (EDUC), czy mężczyzna przerwał edukację przed 12 rokiem życia (NODEGREE), czy jest czarnoskóry (BLACK), czy jest Latynosem (HISP), czy jest żonaty (MARR), czy był bezrobotny w roku 1974 (U74), czy był bezrobotny w roku 1975 (U75), poziom zarobków w 1974 (RE74), poziom zarobków w 1975 (RE75), poziom zarobków w 1978 (RE78). Ostatnia zmienna była traktowana, jako wynikowa.

Rozpatrywano następujące modele:

- Regresja wieloraka

$$\widehat{RE78}_i = \beta_0 + \beta_\tau T_i + \beta_1 X_i, \quad (16)$$

gdzie  $X = AGE, AGE^2, EDUC, NODEGREE, BLACK, HISP, RE74, RE75$ .

- Propensity score

$$\text{logit } e(X_i) = \alpha_0 + \alpha_1 X_i, \quad (17)$$

gdzie  $X = AGE, AGE^2, EDUC, EDUC^2, NODEGREE, BLACK, HISP, MARR, RE74, RE75, RE74^2, RE75^2, U74, U75, U74 \cdot HISP$

Oszacowanie efektu interwencji przeprowadzono w oparciu o kilka technik:

- Różnica średnich bez dopasowania;
- Regresję wieloraką (16);
- Łączenie względem cech;
- PSM metodą najbliższego sąsiada (1:1 z powtórzeniami);
- PSM metodą z promieniem ( $r = 0.001, 0.0001, 0.00001$ );
- PSM z warstwowaniem (6 warstw);
- Regresja wieloraka z *PS*;

Tabela 1 przedstawia poziomy zbalansowania grup stosując wymienione wyżej techniki.

Na jej podstawie możemy jednoznacznie stwierdzić, że najlepsze zbalansowanie kowariantów otrzymujemy stosując łączenie wg cech. W przypadku tego badania liczba zmiennych towarzyszących nie była zbyt wielka, dlatego można było sobie pozwolić na tą technikę. Pomimo, iż zbalansowanie w metodach PSM jest nieco gorsze, to szacunki efektów interwencji są bardziej zbliżone do efektu wyznaczonego metodą eksperymentalną. Ze względu na przejrzystość tabeli nie zostały w niej umieszczone wyniki średnie dla zmiennych towarzyszących w metodzie warstwowej, choć należy dodać, że wykazywały podobny poziom zbalansowania jak inne algorytmy PSM.

W tabeli 2 zaprezentowane są wyniki oszacowania efektu interwencji wraz z błędem standardowym estymacji, który dla łączy 1:1 został wyliczony metodą bootstrap.

**Tabela 1. Średnie zmiennych towarzyszących w podziale na techniki łączenia**

Metoda	Liczba obserwacji	Średnia PS	AGE	EDUC	BLACK	HISP	NODE-GREE	MARR	RE74	RE75	U74	U75
Eksperyment	260	0,37	25,82	10,35	0,84	0,06	0,71	0,19	2095	1532	0,29	0,4
Bez dopasowania	2490	0,02	34,85	12,12	0,25	0,03	0,31	0,87	19429	19063	0,1	0,09
łączenie wg cech	185	-	26,82	10,45	0,84	0,06	0,66	0,19	2213	2282	0,7	0,6
PSM NN	185	0,7	24,81	10,72	0,78	0,09	0,53	0,14	2206	1801	0,54	0,69
PSM r=0.001	2021	0,7	24,98	10,74	0,79	0,09	0,53	0,14	2398	1882	0,53	0,69
PSM r=0.0001	337	0,7	24,92	10,73	0,78	0,09	0,53	0,14	2228	1763	0,54	0,7
PSM r=0.00001	193	0,7	24,85	10,72	0,78	0,09	0,53	0,13	2216	1819	0,54	0,69

Źródło: Opracowanie własne na podstawie: Dehejia R., Wahba S., *Propensity Score Matching Methods for Nonexperimental Causal Studies*, "Review of Economics and Statistics", vol. 84, 2002, s. 28.

**Tabela 2. Szacunki efektu interwencji**

Metoda	ATT	SE
Eksperyment	1794	633
Bez dopasowania	-15205	657
Regresja wieloraka	218	768
łączenie wg cech	2037	1731
PSM NN	1890	1202
PSM r=0.001	1824	1187
PSM r=0.0001	1973	1191
PSM r=0.00001	1893	1198
PSM warstwowa	1452	1141
Regresja z PS	1149	2000

Źródło: Opracowanie własne na podstawie: Dehejia R., Wahba S., *Propensity Score Matching Methods for Nonexperimental Causal Studies*, "Review of Economics and Statistics", vol. 84, 2002, s. 28.

Na podstawie powyższych wyników możemy stwierdzić, że najlepsze oszacowania wielkości efektu zastosowania bodźca, otrzymujemy stosując techniki PSM. Oceny uzyskane techniką warstwową i regresją z włączeniem wartości PS charakteryzują się większym błędem.

## Zakończenie

Z pewnością szeroko rozumiane techniki *Propensity Score Matching* są pomocne w szacowaniu wielkości efektów przyczynowo-skutkowych. Ich stosowanie w kontekście badań ewaluacyjnych jest jak najbardziej uprawnione, a w niektórych przypadkach konieczne. Rozwój tej techniki, wyrażony również rosnącą liczbą publikacji z tego zakresu, jest w ostatnich latach szczególnie zintensyfikowany. Szeroki horyzont zastosowań i aplikacyjny charakter tej techniki statystycznej zostaje często doceniany, co owocuje kolejnymi opracowaniami z tego zakresu, podejmując zarówno aspekty teoretyczne, jak i praktyczne.

Przedstawione w pracy różne techniki *matchingu* wykazują znaczące różnice w podejściu do usuwania obciążenia selekcyjnego. Nie ma wśród nich obiektywnie najlepszej metody, a ich użycie jest często podyktowane charakterem danych, którymi dysponuje badacz. Wiele opracowań wykorzystujących technikę PSM wskazuje na ich komplementarny charakter i poleca stosować je w ten sposób.

Technika PSM nie jest pozbawiona wad, które coraz częściej są sygnalizowane w rozlicznych pracach z tego tematu<sup>34</sup>. Są w nich zawarte przykłady danych, w których po użyciu PSM balans pomiędzy grupami się pogarsza. Krytycy zarzucają także, że te metody balansują dystrybucję wyłącznie zmiennych obserwowanych, a zmienne ukryte mogą być dalej niezbalansowane.

Alternatywne metody *matchingu*, ale oparte o inną miarę odległości to:

- Mahalanobis Distance Matching;
- Coarsened Exact Matching.

34 King G., Nielsen R., Coberley C., Pope J., Wells A., *Comparative effectiveness of matching methods for causal inference*, 2011; King G., Nielsen R., *Why propensity scores should not be used for matching*. Working Paper, 2016.

**Bibliografia**

- Austin, P., Grootendorst, P., & Anderson, G.,** *A comparison of the ability of different propensity score models to balance measured variables between treated and untreated subjects: A Monte Carlo study*, *Statistics in medicine*, 26, 2007, s. 734-753
- Cox, D.,** *A note on weighted randomization*, *The Annals of Mathematical Statistics*, 27(4), 1956, s. 1144-1151
- Dehejia, R., & Wahba, S.,** *Causal effects in nonexperimental studies: Reevaluating the evaluation of training programs*, *Journal of the American Statistical Association*, 94(448), 1999, s. 1053–1062
- Dehejia, R., & Wahba, S.,** *Propensity score matching methods for non-experimental causal studies*, *Review of Economics and Statistics*, 84, 2002, s. 151-161
- Imbens, G.,** *Nonparametric estimation of average treatment effects under exogeneity: A review*, *Review of Economics and Statistics*, 86(1)/2004, s. 4-29
- King, G., Nielsen, R.,** *Why propensity score should not be used for matching*. Working Paper, 2016
- King, G., Nielsen, R., Coberley, C., Pope, J., & Wells, A.,** *Comparative effectiveness of matching methods for causal inference*, 2011
- Lalonde, R.,** *Evaluating the econometric evaluations of training*, *The American Economic Review*, Vol. 76, No. 4/1986, s. 604-620
- Peduzzi, P., Concato, J., Feinstein, A., & Holford, T.,** *Importance of events per independent variable in proportional hazard regression analysis: II. Accuracy and precision of regression estimates*, *Journal of Clinical Epidemiology*, 48(12)/1995, s. 1503-1510
- Robinson, J.,** *The analysis of covariance under a randomization model*, *Journal of the Royal Statistical Society. Series B (Methodological)*, 35(2)/1973, s. 368-376
- Rosenbaum, P.,** *Covariance adjustment in randomized experiments and observational studies*, *Statistical Science*, 17(3)/2002, s. 286-304
- Rosenbaum, P., & Rubin, D.,** *The central role of the propensity score in observational studies for causal effects*, *Biometrika*, 70(1)/1983, s. 41-55
- Rosenbaum P., Rubin D.,** *Reducing bias in observational studies using subclassification on the propensity score*. "Journal of the American Statistical Association", 79, 1984, s. 516–524
- Rubin, D.,** *Using propensity scores to help design observational studies: Application to the tobacco litigation*, *Health Services and Outcomes Research Methodology*, 2(3)/2001, s. 169-188



### **DARIUSZ MAJEREK**

Od 2000 r. pracownik Katedry Matematyki Stosowanej Wydziału Podstaw Techniki Politechniki Lubelskiej, od 2008 r. doktor nauk matematycznych. Praca doktorska pt. „Warunkowe prawa wielkich liczb” dotyczyła uogólnienia praw wielkich liczb rachunku prawdopodobieństwa na ich odpowiedniki warunkowe. Obecna działalność naukowa skupia się wokół statystycznej analizy danych zarówno od strony teoretycznej, jak i praktycznej. Zajmuje się modelowaniem zjawisk masowych za pomocą narzędzi statystycznych w różnych dziedzinach nauki, przede wszystkim ekonomii, ale także inżynierii środowiska, medycynie oraz naukach społecznych. Obecnie prowadzi międzywydziałowe seminarium statystyczne na Politechnice Lubelskiej skupiające naukowców zainteresowanych wykorzystaniem technik statystycznych w różnych dyscyplinach naukowych.

## BOGDAN JANKOWSKI I MAGDALENA BŁASZCZYK

Artykuł wskazuje nowe możliwości dokonywania oceny programów i projektów za pomocą technik socjologii wizualnej i analizy dużych zbiorów danych tzw. big data. Oba narzędzia ewaluacyjne, chociaż operujące w odmiennych obszarach analizy, umożliwiają pozyskanie alternatywnych lub uzupełniających informacji i zdobycie nowej wiedzy. Poszukiwanie nowych metod i technik ewaluacji wydaje się jednym z istotniejszych wyzwań dla ewaluacji programów operacyjnych w perspektywie 2014–2020 i może przyczynić się do rozwoju całej dyscypliny.

### Nowe metody i techniki badawcze a praktyka ewaluacyjna

Konieczność poszukiwania nowych metod i technik badawczych w procesie oceny realizowanych projektów czy programów wydaje się oczywista i nie wymaga szczególnego uzasadnienia. Jednak zarówno instytucje zamawiające ewaluacje oraz firmy badawcze zaangażowane w system ewaluacji w Polsce, nie wykazują w tym zakresie szczególnej aktywności. Każdy kto uważnie obserwuje krótką historię tworzenia i rozwoju systemu instytucjonalnego ewaluacji w Polsce (związanego głównie z Funduszem Spójności) może zauważyć pewien „socjologiczny konserwyzm” metod i technik badawczych stosowanych w trakcie procesu badawczego. Chociaż ewaluatorzy deklarują korzystanie z osiągnięć nauk społecznych, a w szczególności socjologii<sup>1</sup>, to w zasadzie od 2004 roku, kiedy utworzono Krajową Jednostkę Oceny, katalog stosowanych technik badawczych nie uległ większej zmianie. Metodologiczna codzienność to CATI i CAWI na przemian z PAPI i CAPI z dodatkiem IDI i FGI. Oczywiście jest to duże uproszczenie bo mamy przecież analizy ekonometryczne, logiczne czy semantyczne – szczególnie w trakcie prowadzenia ewaluacji *ex-ante*. Pojawiają się również propozycje nowych podejść, jak np. analizy behawioralne<sup>2</sup> a Michael Quinn Patton energicznie upowszechnia wiedzę dotyczącą korzyści płynących z podejścia rozwojowego<sup>3</sup> jednak polska codzienna praktyka ewaluacyjna nie lubi badawczych eksperymentów. Wynika to z ustalonych już, wydaje się trwale, relacji pomiędzy Zamawiającymi a Wykonawcami na rynku badań ewaluacyjnych. Zamawiający – nie zawsze dobrze przygotowani metodologicznie – wiedzą

- 1 Współczesna socjologia jest ogromnie zróżnicowana, zarówno jeżeli chodzi o preferowane paradygmaty, jak i stosowane techniki badawcze.
- 2 Olejniczak K., Śliwowski P., *Nadchodzi rewolucja? Analizy behawioralne w interwencjach publicznych*. [w:] *(R)ewaluacja 2. Wiedza w działaniu*. (red.) Haber A., Olejniczak K., Polska Agencja Rozwoju Przedsiębiorczości, Warszawa 2014.
- 3 Jankowski B., *Ewaluacja rozwojowa. Wspieranie innowacyjności w sytuacji niepewności*. [w:] *Ewaluacja programów operacyjnych. Aspekty teoretyczne, metodologiczne i praktyczne*. Urząd Marszałkowski Województwa Kujawsko-Pomorskiego w Toruniu, Toruń 2013.



czego mogą się spodziewać po znanym już *menu*. Większość instytucji zamawiających ewaluację, niezależnie od tego czy same przygotowują metodykę badania czy też zdają się w tym względzie na wykonawcę, preferuje rozwiązania proste i sprawdzone. Decydują tu: presja czasu, wypracowane już kryteria oceny i względna przewidywalność rezultatów. Z drugiej strony firmy badawcze konkurujące ze sobą na rynku badań ewaluacyjnych przystosowały się do oczekiwań zamawiających gdyż szybko zorientowały się, że przedstawianie niestandardowych podejść badawczych oraz nowych technik nie jest skuteczne przy zdobywaniu kontraktów. Tak więc ambicje badawcze wielu ewaluatorów i chęci stosowania innowacyjnych podejść ustąpiły pod presją praw rynku.

Może jednak to, co powyżej zostało opisane jako ewaluacyjna rutyna, to poszukiwany przez Michaela Quinna Pattona „złoty standard”<sup>4</sup> i dalsze metodologiczne innowacje nie mają sensu? Ale chyba byłby to wniosek nazbyt optymistyczny. Bardziej realistyczne i użyteczne wydaje się zwrócenie w kierunku nowych metod w naukach społecznych i podjęcie próby ich zastosowania w ewaluacji. Zadaniem niniejszego artykułu jest prezentacja dwóch z tych metod: socjologii wizualnej i narzędzi big data.

## Socjologia wizualna

Popularność technik badania obrazów (w szczególności fotografii) lub posługiwanie się obrazem jako narzędziem analizy systematycznie wzrasta. Wizualność stała się obszarem, w który śmiało wkroczyła socjologia wizualna, stając się nowym nurtem (subdyscypliną?) socjologii, a jej entuzjaści<sup>5</sup> coraz częściej korzystają z ilustracji i zdjęć dokumentujących przedstawiane zagadnienia. Chociaż jej użyteczność jest różnie oceniana<sup>6</sup> wizualna strona socjologii uzyskuje w ostatnich latach duże znaczenie nie tylko poprzez coraz częstsze korzystanie z ilustracji, ale właśnie dzięki rozwijaniu metod i technik badawczych właściwych dla socjologii wizualnej<sup>7</sup>. W przesyconym różnorodnymi obrazami współczesnym świecie przesunięcie w kierunku wizualnych aspektów społecznego świata wydaje się uzasadnione, a szczególne zainteresowanie fotografią – trafne, gdyż „siła obrazów fotograficznych wynika stąd, że sama fotografia jest materialną rzeczywistością, bogatym złożem informacji pozostawionym przez to, co ją utrwaliło, znakomitym sposobem odwracania sy-

4 Patton M. Q., *Przyszłe trendy w ewaluacji*. [w:] (R)ewaluacja. Poszukiwanie nowych metod oceny efektów. (red.) Haber A., Popis Z., Polska Agencja Rozwoju Przedsiębiorczości, Warszawa 2013, s. 39.

5 Sztompka P., *Socjologia. Analiza społeczeństwa*. Wydawnictwo Zak, Kraków 2002;

Bauman Z., *Między chwilą a pięknem. O sztuce w rozpadzonym świecie*. Wydawnictwo Oficyna, Łódź 2010;

Bauman Z., Kubicki R., Zeidler-Janiszewska A., *Życie w kontekstach. Rozmowy o tym, co za nami i o tym, co przed nami*. Wydawnictwo Akademickie i Profesjonalne Sp. z o.o., Warszawa 2009.

6 David Silverman prezentując różne sposoby wizualnego myślenia w socjologii (Silverman, 2008, s. 214-236) przestrzega jednak przed ich skomplikowanymi metodami analizy i ich „przeteoretyzowaniem” (Silverman, 2006, s. 163).

7 Sztompka P., *Socjologia wizualna. Fotografia jako metoda badawcza*. Wydawnictwo Naukowe PWN, Warszawa 2005; Pink S., *Etnografia wizualna. Obrazy, media i przedstawienie w badaniach*, Wydawnictwo Uniwersytetu Jagiellońskiego, Kraków 2009; Banks M., *Materiały wizualne w badaniach jakościowych*. Wydawnictwo Naukowe PWN, Warszawa 2009; Konecki K. T., *A Methodological Outline and Examples from Empirical Work, Revija za sociologiju, Vol. 41 No 2, August 2011*.

tuacji na niekorzyść rzeczywistości – sposobem zmieniania r z e c z y w i s t o ś c i w c i eń”<sup>8</sup>. Postawa uwzględniająca różnorodne sposoby wizualnego myślenia, szczególnie w badaniach społecznych nakierowanych na znaczenia i sensy, wydaje się tym bardziej uzasadniona, że „kulturowe uniwersum symboliczne uobecnia się dzisiaj coraz częściej wyłącznie pod postacią zwizualizowanych znaczeń nieposiadających już swoich słownych odpowiedników ani nie tłumaczonych na słowa, oraz o to, że tradycyjne językowe narracje wypierane są przez narracje obrazowe”<sup>9</sup>. *Visual sociology*, która określana jest jako „młodsza siostra etnografii wizualnej” i nazywana bywa także „socjologią obrazu”<sup>10</sup>, choć wciąż znajduje się na etapie poszukiwań - może zaferować badaczom enklawy techniki, które nie są tylko włączaniem wizualności w istniejącą dotychczas metodologię socjologiczną. Badania prowadzone w ramach oceny realizacji projektów czy programów nie muszą mieć ambicji stawiania wizualności w roli głównej, ale nawet pozostając na drugim planie techniki wizualne spełniają kryterium „wzbogacania naszego poznania poprzez obrazy wizualne”<sup>11</sup>. Z wymienianych różnych strategii fotograficznych możliwych do przyjęcia na polu socjologii wizualnej<sup>12</sup> czy też funkcji, jakie fotografia może pełnić wobec socjologii<sup>13</sup>, w procesie oceny projektów i programów można wykorzystać trzy, które można określić jako: dokumentacja faktów wizualnych, inspiracja badawcza oraz ilustracja.

Istotą **dokumentacji faktów wizualnych**<sup>14</sup> jest inwentaryzacja istotnych, ale podlegających zmianie obiektów materialnych. Zdjęcia wykorzystane są tu dla oddania „charakteru” miejsca<sup>15</sup>, które jest obiektem analizy. Możliwość rejestracji różnorodnych, ulotnych wizualnych wrażeń zmysłowych na materialnej powierzchni fotografii, pozwala utrwalić atmosferę, której słowa nie potrafią

8 Sontag S., *O fotografii*, Wydawnictwo Karakter, Kraków 2009, s. 190.

9 Drozdowski R., *Obraz na obrazy. Strategie społecznego oporu wobec obrazów dominujących*. Wyd. II rozszerzone. Zysk i S-ka Wydawnictwo, Poznań 2009, s. 7.

10 Kowalewicz K., *O czterech socjologiach wizualnych*. [w:] *Przestrzenie fotografii. Antologia tekstów*. (red.) Ferenc Tomasz, Makowski Krzysztof. Wydawnictwo Galeria f5 & Księgarnia Fotograficzna, Łódź 2005; Olechnicki K., *Teoria, praktyka i sztuka eseju fotograficznego*. [w:] *Obrazy w działaniu. Studia z socjologii i antropologii obrazu*. (red.) Olechnicki K., Wydawnictwo Uniwersytetu Mikołaja Kopernika w Toruniu, Toruń 2003, s. 8.

11 Pink S., *Etnografia wizualna. Obrazy, media i przedstawienie w badaniach*, Wydawnictwo Uniwersytetu Jagiellońskiego, Kraków 2009, s. 15.

12 Konecki K. T., *Wizualne wyobrażenia. Główne strategie badawcze w socjologii wizualnej a metodologia teorii ugruntowanej*. [w:] *Przegląd Socjologii Jakościowej, Tom I, Numer 1 – Grudzień 2005*, [www.qualitativesociologyreview.org](http://www.qualitativesociologyreview.org), s. 44.

13 Rose G., *Interpretacja materiałów wizualnych. Krytyczna metodologia badań nad wizualnością*. Wydawnictwo Naukowe PWN SA, Warszawa 2010, s. 289-292; Sztompka P., *Socjologia wizualna. Fotografia ...*, op. cit., s. 70-74 - Krzysztof Konecki wymienia cztery strategie z użyciem fotografii: jako podstawowa i celowa strategia badawcza; jako zastany obiekt o znaczeniu społecznym; jako materiał zastany wywołujący narracje werbalne; jako ilustracja dla wniosków lub uzupełnienie. Piotr Sztompka wskazuje na sześć funkcji fotografii: stymulowanie uwagi i wyobraźni; inspiracja heurystyczna; rejestracja; materiał pomocniczy do prowadzenia wywiadu; ilustracja pojęć, kategorii i prawidłowości; propaganda i perswazja. Ogólny podział zastosowany przez Gillian Rose to: zdjęcia jako wsparcie w naukach społecznych i zdjęcia jako uzupełnienie w naukach społecznych.

14 Już w 1910 roku na V Międzynarodowym Kongresie Fotografii stwierdzono, że pojęcie dokumentu odnosi się tylko do tych fotografii, które mogą zostać „wykorzystane dla wszelkiego rodzaju badań naukowych” Rouillé A., *Fotografia. Między dokumentem a sztuką współczesną*. TAIWPN UNIVERSITAS, Kraków 2007, s. 63.

15 Rose G., *Interpretacja materiałów...*, op.cit., s. 291.

równie dobrze przekazać. Możliwość „zatrzymania czasu” i utrwalenia szczegółów trudno poddających się słownemu opisowi pozwala na traktowanie zdjęć jako dodatkowego lub alternatywnego źródła informacji. Dane te mogą być przechowywane i używane do powtórnej analizy i porównywania z innymi - w innej przestrzeni lub innym czasie. Fotografia może być stosowana jako uzupełnienie tekstu pisanego lub równie efektywnie wykorzystana w formie foto-narracji/foto-eseju<sup>16</sup>.

**Rysunek 1. Trasa W-Z Łódź 2012 r.**



Źródło: Bogdan Jankowski, 2012.

**Rysunek 2. Trasa W-Z Łódź 2013 r.**



Źródło: Bogdan Jankowski, 2013.

<sup>16</sup> Banks M., *Materiały wizualne...*, op. cit., s. 161-164; Harper D., *Co nowego widać?* [w:] *Metody badań jakościowych. Tom 2*, (red.) Denzin Norman K., Lincoln Yvonna S., Wydawnictwo Naukowe PWN, Warszawa 2009, s. 156; Olechnicki K., *Teoria, praktyka i sztuka...*, op. cit., s. 15-37; Pauvels L., *The visual essay: Affinities and divergences between the social scientific and social documentary modes*. *Visual Anthropology*, Volume 6, Issue 2/1993, s. 199-210.

**Rysunek 3. Trasa W-Z Łódź 2015 r.**

Źródło: Bogdan Jankowski, 2015.

Materiały zastane, czyli np. fotografie, mogą też pełnić funkcję **inspiracji badawczej**. Są to bowiem światy nie tylko nam obce, o których chcemy wzbogacić naszą wiedzę, ale też często takie, o których istnieniu nie mieliśmy pojęcia. *„Obdarza nas ona [fotografia] możliwością wszechobecności, bycia w innej przestrzeni i czasie niż te, w których jesteśmy, ale w takich, które były, które naprawdę istniały. Pozwala ona uważać siebie za element pośredniczący, dzięki któremu jesteśmy natychmiast obecni w określonej przeszłości lub gdzie indziej”*<sup>17</sup>. Udostępnione nam zdjęcia otwierają często przed nami obszary zmuszające do przedefiniowania lub tworzenia pojęć oraz stawiania nowych hipotez.

Funkcja **ilustracji** pomaga w prezentacji pozyskanych danych i wzmacnia tekst drukowany. *„Problemem nie jest to, że ludzie pamiętają przez fotografie, lecz że zapamiętują wyłącznie fotografie. Zapamiętanie poprzez fotografie przesłania inne formy rozumienia i pamiętania”*<sup>18</sup>. W tej pracy funkcja ilustracyjna nie jest realizowana poprzez obrazowania ogólnych pojęć za pomocą zdjęć pozyskanych z dostępnych w internecie *image banks*<sup>19</sup>, ale prezentuje ilustracje konkretnych sytuacji, osób lub przedmiotów związanych z prowadzonymi badaniami. Tego rodzaju fotografię Francois Soulage nazywa definicyjną lub informatywną, gdyż „ma na celu oddanie w sposób jednoznaczny sensu odniesienia”<sup>20</sup>.

17 Soulages F., *Estetyka fotografii. Strata i zysk*. TAIWPN UNIVERSITAS, Kraków 2005, s. 16.

18 Sontag S., *Widok cudzego cierpienia*. Wydawnictwo Karakter, Kraków 2010, s. 107.

19 *Image banks* oferują produkcje fotograficzne, tzw. infra-obrazy, wypożyczają i zarządzają zdjęciami dostarczonymi przez fotografów w różnych tematycznych obszarach takich jak: sport, scena, podróże, przemysł, postaci, abstrakcje, spełniając zasadę uniformizacji, kierują się zasadą użyteczności. Jednym z pierwszych takich przedsięwzięć była amerykańska firma Stanley’a Kanney’a (obecnie notowana na nowojorskiej giełdzie) *The Image Bank* założona w 1974 roku (Rouillé, 2005, s. 142).

20 Soulages F., *Estetyka fotografii...*, op. cit., s. 296.

## Big data

Pojęcie Big Data w ostatnich latach stało się niezwykle popularne, a jego rozumienie, jako zbiorów o dużej objętości, różnorodności i szybkim tempie narastania rozpało wyobraźnię i nadzieję wielu badaczy na znaczne ograniczenie problemów związanych z doбором i realizacją próby. Niestety, deficyt narzędzi i metod do przetwarzania, zarządzania i przechowywania tych danych skutkowało niewielkim wzrostem ilości badań realizowanych z wykorzystaniem dużych zbiorów. Ponadto do problemów natury technicznej dołączyły czynniki związane z dostępem do danych, wśród których jednym z głównych, w przypadku badań ewaluacyjnych, okazał się problem na poziomie instytucjonalno-prawnym, rozumiany tu jako brak uregulowań i procedur umożliwiających instytucjom publicznym przekazywanie zanonimizowanych danych. W rzeczywistości ewaluacyjnej jednymi z najbardziej pożądanymi są dane zbierane przez ZUS i Urzędy Skarbowe gdyż pozwalają one na prowadzenie badań z niespotykaną do tej pory dokładnością. Zbiory te są pożądane również z uwagi na swoją wysoką jakość - są uporządkowane, rzetelne i precyzyjne, czyli posiadają wszystkie cechy charakterystyczne dla tradycyjnego podejścia do zbierania danych. Niestety, dostęp do nich uzyskało do tej pory niewielu ale dzięki coraz liczniejszym próbom podejmowanym przez różne zespoły badawcze należy mieć nadzieję, iż instytucje te wypracują odpowiednie mechanizmy organizacyjno-prawne pozwalające na sprawne i bezpieczne przekazywanie informacji zgromadzonych w swoich bazach. Takich uporządkowanych zbiorów w świecie big data nie ma zbyt wiele dlatego stały się one swoistym świętym gralem dla ewaluatorów przesłaniając inne, może już nie tak idealne ale wciąż bardzo użyteczne źródła danych. W większości są one nieuporządkowane i mnóstwo w nich błędów pomiaru ale jest to immanentna cecha big data mająca znacznie mniejsze znaczenie niż sądzono. Big data to w rzeczywistości zmiana sposobu myślenia - przejście od postrzegania badań przez pryzmat doboru próby, do myślenia w kategoriach  $n=całość$ <sup>21</sup>. Eliminacja błędów wynikających z doboru próby pozwala nam na większą akceptację błędów pomiaru. Rewolucja Big Data polega właśnie na zmianie podejścia polegającej na pogodzeniu się z większym nieuporządkowaniem i różnorodnością danych. Oczywiście wymaga to wypracowania nowych metod i narzędzi badawczych, ale nagrodą będzie odkrycie olbrzymiej ilości zależności, które z uwagi na ograniczone zbiory danych umykały do tej pory badaczom.

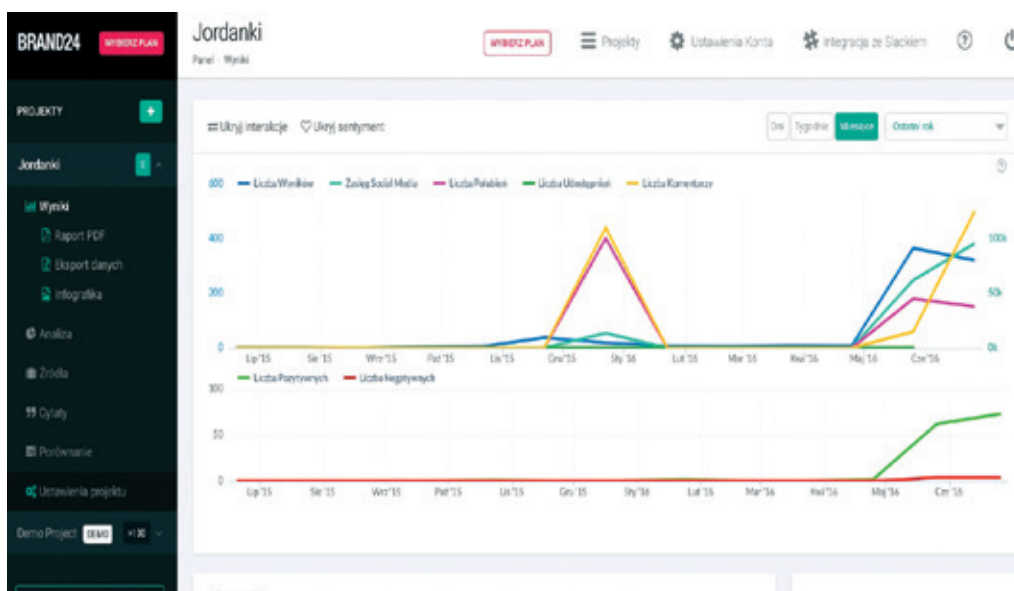
Efektom coraz większej popularności big data jest wzrost liczby oferowanych narzędzi do zarządzania, analizowania i przechowywania dużych zbiorów danych. Niestety większość z nich nie jest zbyt często wykorzystywana w badaniach ewaluacyjnych. Brak popularności wynika w dużej mierze z braku przygotowania do korzystania z rozwiązań wymagających w wielu przypadkach dużej wiedzy z zakresu algorytmiki, baz danych czy znajomości języków programowania. W efekcie narzędzia takie jak Hadoop, MapReduce (przeznaczone do rozproszonego składowania i przetwarzania wielkich zbiorów danych) czy roboty przeszukujące Internet nie pojawiają się na liście narzędzi wy-

21 Cukier K., *The Big Data is better data*. TED. Talks [ 23.09.2014], <https://www.youtube.com/watch?v=8pHzROP1D-w>.

korzystywanych w procesie badawczym. Sytuacje mają szansę zmienić dynamicznie rozwijające się narzędzia analityczne pod warunkiem, że ich obsługa będzie intuicyjna, a wymagania stawiane użytkownikowi nie będą znacząco odbiegać od tych stawianych analitykom czy badaczom społecznym.

Jednym z przykładów tego typu rozwiązań są narzędzia przeznaczone do monitoringu Internetu i mediów społecznościowych takie jak Brand24, Monitori czy SentiOne. Zostały one stworzone z myślą o monitorowaniu marek jednak nic nie stoi na przeszkodzie aby wykorzystać je do oceny zasięgu epidemii grypy, poparcia dla kandydatów na prezydenta czy ewaluacji programów. Narzędzia te poprzez zdefiniowane przez użytkownika słowa kluczowe pozwalają na obserwowanie powiązanych z nimi wypowiedzi publikowanych w Internecie. Mogą to być pojedyncze słowa (np. konferencja, Toruń, Jordanki) lub też bardziej złożone frazy określające grupę produktów bądź zjawisk (np. konferencje ewaluacyjne). Konieczny jest tu dobór odpowiednich słów czy fraz kluczowych, gdyż będą one podstawą uzyskanych wyników. Stworzone w ten sposób projekty można monitorować przy użyciu wbudowanych w programy narzędzi, warunkujących przydatność i popularność danego rozwiązania programistycznego. Program Brand24 oferuje na przykład jedną z bardziej użytecznych funkcji w kontekście badań ewaluacyjnych pozwalającą na dostęp zarówno do bieżących wypowiedzi (monitoring w czasie rzeczywistym) jak i do danych historycznych. Pozwala to na ocenę w czasie interesującego nas projektu, np. budowy Centrum Kulturalno-Kongresowe Jordanki, na długo po jego zakończeniu. Dla badaczy jest to cenne narzędzie pozwalające na niemal nieograniczony dostęp do danych historycznych.

Rysunek 4. Widok wyników monitoringu słowa kluczowego #Jordanki

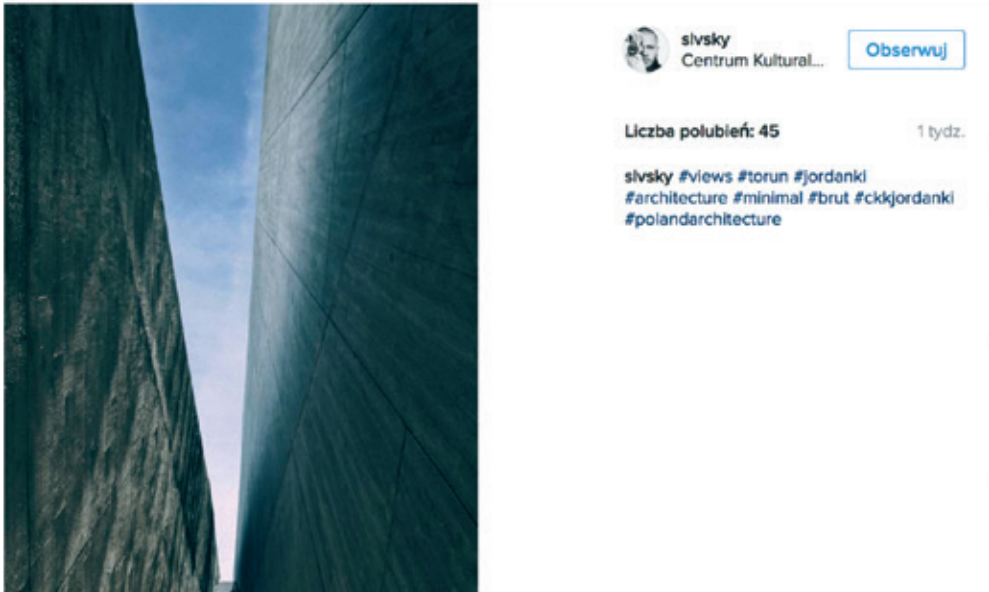


Źródło: Brand24.



Przedstawiony na poprzedniej stronie panel wynikowy obrazuje część charakterystycznych dla programu Brand24 funkcji. W trakcie użytkowania istnieje nie tylko możliwość podsumowania łącznej liczby wypowiedzi pojawiających się w kontekście monitorowanego słowa kluczowego ale również dostępny jest ich rozkład, ze względu na platformę na jakiej zostały opublikowane. Analizie można poddać również zasięg Social Media, liczbę polubień, liczbę udostępnień i komentarzy (na rysunku 4. pierwszy z widocznych na panelu wykresów). Warto w tym miejscu zaznaczyć, iż narzędzie szacuje zasięg na bazie potencjalnej ilości obserwujących, fanów etc. w związku z czym jest to tylko estymacja, a nie rzeczywista wartość wskaźnika. Program pozwala również badać tzw. sentyment czyli poddaje analizie nacechowanie emocjonalne publikowanych wypowiedzi, klasyfikując je jako pozytywne, negatywne bądź neutralne (drugi z widocznych na rysunku 4. wykresów). Pozwala też przyjrzeć się konkretnej wypowiedzi bądź zdjęciu, które z jakichś powodów nas zainteresowało (np. jest odpowiedzialne za duży wzrost zasięgu).

**Rysunek 5. Widok zdjęcia opisanego przez zdefiniowane słowo kluczowe**



Źródło: Brand24.

Program Brand24 dając dostęp do zaawansowanych wskaźników pozwala na lepsze zrozumienie badanego przez nas projektu czy zjawiska. Obok analizy w czasie, interesująca w kontekście badań wydaje się być możliwość prezentacji wyników z użyciem filtrów, dzięki którym można analizować wyniki wg wybranych grup, domen, autorów, wartości wypowiedzi czy ptci. Możliwości jest znacznie więcej, a poprzez eksport (Excel, PDF) uzyskujemy w zasadzie nieograniczone możliwości analiz i łączenia danych. Brand24 nie jest jedynym narzędziem służącym do monitorowania Internetu, choć jednym z bardziej popularnych. Łatwość dokonywania analiz i dostępność oferowanych przez tego typu narzędzia rozwiązań gwarantuje nie tylko wzrost ich popularności ale również rozwój w kierunkach interesujących środowisko badawcze. Pozostaje tylko kwestią czasu kiedy ulepszone narzędzia do automatycznej analizy fotografii zaimplementowane zostaną do programów monitorujących Internet otwierając drogę do eksploracji danych na niespotykaną dotąd skalę.

### **Zakończenie**

Socjologię wizualną i analizy wykorzystujące duże zbiory danych łączy, oprócz niewielkiego stopnia w jakim wykorzystują je na co dzień badacze, chęć dostrzeżenia ukrytych znaczeń i związków. Narzędzia socjologii wizualnej wpływają na lepsze rozumienie charakteru badanych zjawisk, a udostępnione zdjęcia otwierają często przed nami obszary zmuszające do przededefiniowania lub tworzenia pojęć oraz stawiania nowych hipotez. Immanentną cechą analiz big data jest odkrywanie korelacji w niezliczonej ilości danych pochodzących z różnych źródeł takich jak transakcje zakupowe, dane sieci komórkowych, portale społecznościowe, różnego rodzaju czujniki, internet rzeczy, satelity etc. Wyzwanie stojące przed ewaluatorami to zhumanizowanie tych danych i uczynienie ich przydatnymi dla lepszego zrozumienia mechanizmów stojących za wdrażanymi interwencjami publicznymi.



## Bibliografia

- Banks M.**, *Materiały wizualne w badaniach jakościowych*. Wydawnictwo Naukowe PWN, Warszawa 2009
- Bauman Z.**, *Między chwilą a pięknem. O sztuce w rozpadzonym świecie*. Wydawnictwo Officyna, Łódź 2010
- Bauman Z., Kubicki R., Zeidler-Janiszewska A.**, *Życie w kontekstach. Rozmowy o tym, co za nami i o tym, co przed nami*. Wydawnictwo Akademickie i Profesjonalne Sp. z o.o., Warszawa 2009
- Cukier K.**, *The Big Data is better data*. TED. Talks [23.09.2014], <https://www.youtube.com/watch?v=8pHzROP1D-w>
- Drozdowski R.**, *Obraza na obrazy. Strategie społecznego oporu wobec obrazów dominujących*. Wyd. II rozszerzone. Zysk i S-ka Wydawnictwo, Poznań 2009
- Harper D.**, *Co nowego widać?* [w:] *Metody badań jakościowych*. Tom 2, (red.) Denzin Norman K., Lincoln Yvonna S., Wydawnictwo Naukowe PWN, Warszawa 2009
- Jankowski B.**, *Ewaluacja rozwojowa. Wspieranie innowacyjności w sytuacji niepewności*. [w:] *Ewaluacja programów operacyjnych. Aspekty teoretyczne, metodologiczne i praktyczne*. Urząd Marszałkowski Województwa Kujawsko-Pomorskiego w Toruniu, Toruń 2013
- Katal A., Wazid M., Goudar R. H.**, *Big Data: Issues, Challenges, Tools and Good Practices*, 2013 *Sixth International Conference on Contemporary Computing (IC3), IEEE, Noida, s. 404-409*.
- Konecki K. T.**, *A Methodological Outline and Examples from Empirical Work*, *Revija za sociologiju*, Vol. 41 No 2, August 2011
- Konecki K. T.**, *Wizualne wyobrażenia. Główne strategie badawcze w socjologii wizualnej a metodologia teorii ugruntowanej*. [w:] *Przegląd Socjologii Jakościowej*, Tom I, Numer 1 – Grudzień 2005, [www.qualitativesociologyreview.org](http://www.qualitativesociologyreview.org)
- Kowalewicz K.**, *O czterech socjologiach wizualnych*. [w:] *Przestrzenie fotografii. Antologia tekstów*. (red.) Ferenc Tomasz, Makowski Krzysztof. Wydawnictwo Galeria f5 & Księgarnia Fotograficzna, Łódź 2005
- Mayer-Schönberger V., Cukier K.**, *Big Data, Rewolucja która zmieni nasze myślenie, pracę i życie*, MTBiznes, Warszawa 2014
- Olechnicki K.**, *Teoria, praktyka i sztuka eseju fotograficznego*. [w:] *Obrazy w działaniu. Studia z socjologii i antropologii obrazu*. (red.) Olechnicki K., Wydawnictwo Uniwersytetu Mikołaja Kopernika w Toruniu, Toruń 2003
- Olejniczak K., Śliwowski P.**, *Nadchodzi rewolucja? Analizy behawioralne w interwencjach publicznych*. [w:] *(R)ewaluacja 2. Wiedza w działaniu*. (red.) Haber A., Olejniczak K., Polska Agencja Rozwoju Przedsiębiorczości, Warszawa 2014
- Patton M. Q.**, *Przyszłe trendy w ewaluacji*. [w:] *(R)ewaluacja. Poszukiwanie nowych metod oceny efektów*. (red.) Haber A., Popis Z., Polska Agencja Rozwoju Przedsiębiorczości, Warszawa 2013

- Pauvels L.**, *The visual essay: Affinities and divergences between the social scientific and social documentary modes*. Visual Anthropology, Volume 6, Issue 2/1993, s. 199-210
- Pink S.**, *Etnografia wizualna. Obrazy, media i przedstawienie w badaniach*, Wydawnictwo Uniwersytetu Jagiellońskiego, Kraków 2009
- Rose G.**, *Interpretacja materiałów wizualnych. Krytyczna metodologia badań nad wizualnością*. Wydawnictwo Naukowe PWN SA, Warszawa 2010
- Rouillé A.**, *Fotografia. Miedzy dokumentem a sztuką współczesną*. TAIWPN UNIVERSITAS, Kraków 2007
- Silverman D.**, *Doing Qualitative Research. A Practical Handbook*. SAGE Publication, London, Thousands Oaks, New Delhi 2006
- Silverman D.**, *Interpretacja danych jakościowych. Metody analizy rozmowy, tekstu i interakcji*. Wydawnictwo Naukowe PWN, Warszawa 2008
- Sontag S.**, *O fotografii*, Wydawnictwo Karakter, Kraków 2009
- Sontag S.**, *Widok cudzego cierpienia*. Wydawnictwo Karakter, Kraków 2010
- Soulages F.**, *Estetyka fotografii. Strata i zysk*. TAIWPN UNIVERSITAS, Kraków 2005
- Sztompka P.**, *Socjologia. Analiza społeczeństwa*. Wydawnictwo Znak, Kraków 2002
- Sztompka P.**, *Socjologia wizualna. Fotografia jako metoda badawcza*. Wydawnictwo Naukowe PWN, Warszawa 2005



### **BOGDAN JANKOWSKI**

Bogdan Jankowski – ewaluator i socjolog. Od 2014 r. pełni funkcję zastępcy dyrektora Regionalnego Centrum Polityki Społecznej w Łodzi. W latach 2007-2013 w Jednostce Ewaluacyjnej Urzędu Marszałkowskiego Województwa Łódzkiego. Doktor nauk humanistycznych w zakresie socjologii. Członek International Visual Sociology Association.

### **MAGDALENA BŁASZCZYK**

Z wykształcenia socjolog, ewaluator i trener. Do niedawna zatrudniona w Jednostce Ewaluacyjnej Urzędu Marszałkowskiego Województwa Łódzkiego, obecnie związana z Regionalnym Obserwatorium Integracji Społecznej przy Regionalnym Centrum Polityki Społecznej w Łodzi. Brała udział w kilkudziesięciu projektach badawczych oraz szkoleniowych. Zajmuje się badaniami społecznymi, w tym ewaluacyjnymi oraz analizami z zakresu komercjalizacji nowych technologii.

# W POSZUKIWANIU LEPSZYCH REKOMENDACJI... JAK WYKORZYSTAĆ PODEJŚCIE DESIGN THINKING W PRACY EWALUATORA.

BARBARA LESZCZYŃSKA

*Prawdziwa podróż odkrywcza nie polega na szukaniu nowych lądów, lecz na nowym spojrzeniu.*

Marcel Proust *W poszukiwaniu straconego czasu*

## Wstęp

Istotą badań ewaluacyjnych jest dokonanie oceny (efektów, sposobów wdrażania, przyjętych założeń interwencji publicznej, itp.). Zadaniem tej oceny jest zaś rozwój, który operacyjnie można zdefiniować jako wprowadzanie zmian w działaniu, tak by prowadziły do celu stawianego przed interwencją publiczną. Narzędziem do tego służącym są właśnie rekomendacje.

Najogólniej można stwierdzić, że rekomendacje stanowią zalecenia do zmian. Ta definicja jest bardzo szeroka. Stosowana w ewaluacji środków pochodzących z funduszy Unii Europejskiej tabela rekomendacji idzie o krok dalej. Od ewaluatora oczekuje wskazania: sposobu wdrożenia rekomendacji, a więc już **konkretnych rozwiązań**, które należy zastosować; instytucji za to odpowiedzialnej czy nawet terminu realizacji. Daje to wrażenie, że zarówno zamawiający, jak i wykonawcy zakładają, że w wyniku ewaluacji możliwe jest zaproponowanie określonych działań: możliwie szczegółowych, realistycznych i terminowych. Być może z tego powodu, obie strony tak boją się rekomendacji innowacyjnych, nieszablonowych, a przede wszystkim bez 100% gwarancji na powodzenie.

Projektując odpowiednio proces ewaluacji możliwe jest zwiększenie jakości rekomendacji: wyjście poza dotychczasowe doświadczenia (zarówno ewaluatora po stronie zamawiającego i wykonawcy), przeanalizowanie kilku możliwych scenariuszy, wzorowanie się na rozwiązaniach stosowanych w innych branżach i obszarach tematycznych. Jedną z metod, która może w tym pomóc jest Design Thinking. Jest to podejście wykorzystywane głównie zagranicą do tworzenia innowacyjnych rozwiązań tak w biznesie<sup>1</sup>, jak i w polityce publicznej. Zyskuje coraz większą popularność również w Polsce.

1 Jest wykorzystywana m.in. przez Google, Intel, Ideo, SAP itp., w polityce publicznej sięgnęły po nią rządy Wielkiej Brytanii i Danii.

Celem niniejszego artykułu jest pokazanie, w jaki sposób metodyka Design Thinking może wzbogacić warsztat ewaluatora w przygotowywaniu rekomendacji – trafnych, konkretnych, bardziej szczegółowych i wychodzących poza oczywiste wnioski płynące z badania.

### Jak powstają rekomendacje?

By móc wdrożyć rozwiązania Design Thinking w proces tworzenia rekomendacji, warto wpiąć się dotychczas stosowanemu sposobowi powstawania rekomendacji. Jak pokazuje 8-letnie doświadczenie autorki artykułu, ewaluatorzy przyjmują nieme założenie, że rekomendacje wypływają wprost (lub niemal wprost) z wyciągniętych wniosków, wnioski zaś z przeprowadzonych badań i analiz. Ewaluatorzy (tak zlecający i wykonujący badanie) zakładają w ten sposób, że przeprowadzony prawidłowo proces zbierania i analizy danych pozwoli nie tylko ocenić interwencję publiczną, ale również zaproponować adekwatne zmiany w jej wdrażaniu. Wydaje się jednak, że to założenie jest błędne, a przynajmniej nie zawsze prawdziwe. Przeprowadzona przez autorkę artykułu analiza wniosków stawianych w raportach ewaluacyjnych wskazuje, że badania rzeczywiście pozwalają ocenić, czy dane rozwiązanie działa lub nie, choć nie zawsze jednoznacznie pozwalają ustalić przyczynę takiego stanu rzeczy. Znany wszystkim ewaluatorom przykładem może być diagnozowany w olbrzymiej liczbie badań problem nadmiernej biurokracji, będącej znaczącym obciążeniem dla beneficjentów. Z tego wniosku dość łatwo sformułować rekomendację o treści: zmniejszenie obciążeń biurokratycznych dla beneficjentów. Jednakże, jak zapewne sam czytelnik zauważa, treść rekomendacji jest bardzo ogólna, właściwie nie wskazuje żadnych konkretów. Jej zoperacjonalizowanie wciąż wymaga dużych nakładów pracy: poznania tak potrzeb beneficjentów, jak i instytucji zarządzających i wdrażających, zaproponowania nowych rozwiązań, z których zostanie wybrane najlepsze, przetestowania zmian i wprowadzenia korekt. Być może na koniec procesu okaże się, że inne rozwiązanie byłoby bardziej adekwatne.

Badania ewaluacyjne diagnozują również nowe niezaspokojone potrzeby beneficjentów interwencji publicznej (za przykład można wybrać potrzebę dostosowania sposobów rekrutacji do osób niepełnosprawnych). Jednakże dość rzadko badania wprost pokazują, w jaki konkretny sposób dana potrzeba powinna być zaspokojona lub jakie konkretnie rozwiązanie powinno być wdrożone. W tym miejscu warto zacytować Henriego Forda, który podobno powiedział: *Gdybym na początku swojej kariery jako przedsiębiorcy zapytał klientów, czego chcą, wszyscy byliby zgodni: chcemy szybszych koni. Więc ich nie pytałem.* Ewaluacja również nie może poprzestać na pytaniu o to beneficjentów, urzędników, samorządów itp. Oczekiwanie, że respondenci w ankietach, czy nawet w badaniach jakościowych, sami wskażą rozwiązania: nowe, innowacyjne, czy dla nich najlepsze, wydaje się z góry nietrafione. Zazwyczaj odpowiedzi uczestników badań często są tylko wskazówką, jakie są ich potrzeby. Potrzebą klientów Henriego Forda była potrzeba szybszego swobodnego przemieszczania się. W końcu

XIX wieku i początku XX pociągi dawały możliwość szybkiego pokonywania dużych odległości, ale miały swoje ograniczenia, jakim był rozkład jazdy czy docieranie tylko do wybranych miejsc. Konie natomiast dawały swobodę – zawsze były gotowe, by zawieźć jeźdźca do wybranego przez niego celu. Koszt ich utrzymania też był niewielki. Dalsza podróż jednak była długa i nużąca. Produkowane przez Forda auta łączyły potrzeby: swobody, szybkości i przystępnej ceny.

Bliższe przyjrzenie się procesowi powstawania rekomendacji pokazuje również, że jest to moment kończenia badań terenowych i powstawania raportu końcowego. Ewaluatorzy słyszają już tykający zegar odmierzający dni do zakończenia umowy zawartej pomiędzy zamawiającym a wykonawcą. Doświadczenie autorki w realizacji ewaluacji w różnych zespołach, różnych firmach i dla różnych zamawiających pokazuje, że wiąże się to z jednej strony z pośpiechem, by oddać produkty na czas, z drugiej ze zwyczajnym zmęczeniem i chęcią zamknięcia projektu. Z tego powodu czas na operacjonalizację rekomendacji – wypracowania rozwiązań, które będą zarówno skuteczne, jak i możliwe do wdrożenia – jest stosunkowo krótki. Pamiętajmy, że zamawiający oczekuje, że wyniki ewaluacji wskażą konkretnie: co, kto, jak i kiedy powinien zrobić, by ulepszyć prowadzone działania i osiągnąć zamierzone efekty. Zatem podkreślimy raz jeszcze: przygotowanie rekomendacji jest momentem kluczowym dla powodzenia całego procesu ewaluacji.

Czy zatem ewaluacja jest skazana na mało twórcze rekomendacje? Wydaje się, że nie. Wymaga to jednak, po pierwsze, wydzielenia w harmonogramie czasu na pracę nad rekomendacjami. Po drugie, wykorzystanie do tego nowych narzędzi. Po trzecie, wcielenie się przez ewaluatora w rolę konsultanta, a nie tylko badacza czy eksperta<sup>2</sup>. Z tego punktu widzenia zasadne wydaje się sięganie przez ewaluatorów po metody stosowane w konsultingu. Mogą one znacząco wpłynąć na jakość i użyteczność rekomendacji. Jedną z takich technik jest metoda Design Thinking, która jest przedmiotem niniejszego referatu.

### **Czym zatem jest Design Thinking?**

W literaturze przedmiotu i licznych publikacjach internetowych Design Thinking bywa ściśle definiowane jako metoda tworzenia innowacyjnych *produktów i usług* na podstawie zrozumienia potrzeb klienta – jego zwyczajów, odczuć, myśli. Podejście Design Thinking jest stosowane przez największe koncerny – w tym Google, Intel, Ikea, Bayern, czy też przez prekursorów wykorzystania myślenia projektowego w biznesie SAP i IDEO. Jednakże przykład duńskiego MindLab, czyli międzyministerialnego laboratorium innowacji, pokazuje, że jest to podejście,

---

2 W trakcie III Międzyregionalnej Konferencji Ewaluacyjnej w Toruniu padła teza, że ewaluator oprócz bycia badaczem i ekspertem, powinien pełnić funkcję konsultanta, a jego narzędziem w tym obszarze są między innymi rekomendacje.

które może być wykorzystywane tak do projektowania produktów biznesowych, struktur organizacyjnych, jak i usług społecznych. Założenia DT mogą być również z łatwością wprowadzane w proces ewaluacji, co zostało przetestowane przez zespół ewaluacyjny firmy Danae i Fundacji Idea Rozwoju w badaniu dla Narodowego Centrum Badań i Rozwoju<sup>3</sup>. Wyniki tego doświadczenia będą ilustracją dla dalszych rozważań.

By zrozumieć podejście Design Thinking należy poznać 3 kluczowe zasady, na których się opiera. Po pierwsze, projektowanie metodą Design Thinking jest **skoncentrowane na człowieku** (*Human-centered design*), czyli w centrum całego procesu tworzenia rozwiązań jest człowiek, dla którego one powstają. W Design Thinking wykorzystuje się wiele technik, które mają za zadanie zadbać, by lepiej poznać potrzeby, oczekiwania, ale też zwyczaje, emocje i myśli. Jedną z nich jest technika *persony*. Na potrzeby procesu kreatywnego tworzy się charakterystyki osób, dla których się projektuje (cechy społeczno-demograficzne, zwyczaje, codzienne doświadczenia, motywacje, wartości, którymi się kierują). Najlepszymi personami są tzw. *skrajni użytkownicy*, czyli osoby, które najbardziej mają skorzystać na wprowadzeniu nowego rozwiązania. Przykładem skrajnych użytkowników mogą być osoby niepełnosprawne, czy też rodzice z dziećmi lub rodziny wielodzietne. Zakłada się w ten sposób, że jeśli rozwiązania będą dla nich odpowiednie, będą również odpowiednie dla przeciętnych użytkowników. Przykładem produktu, który został stworzony na potrzeby osób niepełnosprawnych, a stał się produktem masowego użytku, jest szczoteczka elektryczna<sup>4</sup>. Skoncentrowanie na człowieku powinno być wartością szczególnie bliską administracji publicznej. Jednakże dość powszechnie występujące dysfunkcje biurokratyczne, np. skupienie się na celach autonomicznych organizacji takich jak chociażby osiągnięcie wskaźników, mogą powodować, że instytucje i pracujący w niej ludzie zapominają, komu ma służyć dane działanie. Wykorzystanie Design Thinking w ewaluacji i programowaniu polityk publicznych pozwala systemowo wrócić do tych wartości, które są podstawą służby publicznej.

Po drugie, proces Design Thinking ma **charakter interdyscyplinarny**, włączani są do niego eksperci reprezentujący różne podejścia: socjologowie, etnografowie, inżynierowie, politycy społeczni, informatycy. Ich współpraca daje szansę zaistnieć nowym jakościom, które mogą mieć innowacyjny charakter. Różnorodność członków zespołu, to różnorodność perspektyw i różnorodność wymyślanych rozwiązań. W ewaluacji IV Priorytetu PO KL poza członkami zespołu badawczego i przedstawicielami zamawiającego do prac włączyliśmy osoby o różnorod-

3 Leszczyńska B. (red) *Badanie ewaluacyjne o charakterze on-going identyfikujące dobre praktyki w realizacji praktycznych elementów kształcenia w projektach dofinansowanych w ramach IV Priorytetu PO KL oraz w innych działaniach w obszarze szkolnictwa wyższego*. Raport Końcowy z badania przeprowadzonego dla Narodowego Centrum Badań i Rozwoju przez Danae sp. z o.o. i Fundacja Idea Rozwoju.

4 Broxodent, pierwsza na świecie szczoteczka elektryczna, została wynaleziona w 1939 roku w Szwajcarii przez dr Philippe-Guy Woog'a. Powstała ona z myślą o osobach z ograniczonymi funkcjami motorycznymi oraz o pacjentach ortodontycznych. Miała na celu umożliwiać i ułatwiać im codzienne mycie zębów.

nych doświadczeniach zawodowych: artystów, studentów czy menadżerów, a więc osoby, od których praca wymaga kreatywności na co dzień. Ich rolą było świeże spojrzenie na problem i generowanie niestandardowych rozwiązań.

Po trzecie, Design Thinking to **eksperymentowanie i testowanie hipotez**. Z jednej strony szukanie nowych rozwiązań, z drugiej szybka ocena ich potencjału do wdrożenia. Metodyka DT uwzględnia ryzyko porażki, dlatego też każdy z wybranych pomysłów poddawany jest szybkiej ocenie, w wyniku której jest ulepszany lub porzucany. Ta zasada DT może wydawać się trudna do wdrożenia w jednostkach administracji publicznej ze względu na ich hierarchiczność czy też brak odpowiednich wewnętrznych procedur, wydaje się jednak, że są to przeszkody do pokonania.

### **Przebieg procesu tworzenia rekomendacji z wykorzystaniem Design Thinking**

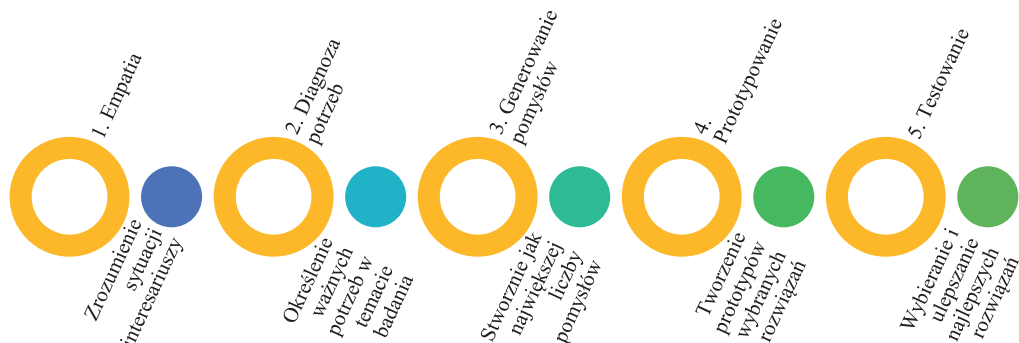
Założenia metodyki Design Thinking doskonale wpisują się w oczekiwania, jakie zamawiający stawiają przed rekomendacjami, które powstają w wyniku przeprowadzonej ewaluacji. Pojawia się pytanie, jak praktycznie zorganizować badanie, by na koniec otrzymać wartościowe i konkretne rekomendacje. Poniżej przedstawiano praktyczne wskazówki i doświadczenia wynikające z badań prowadzonych przez autorkę artykułu.

Przyjmuje się, że metodyka Design Thinking jest 5 etapowym heurystycznym procesem. Rozpoczyna się **Empatią**, czyli głębokim zrozumieniem potrzeb osób, dla których będą projektowane rozwiązania. Dla ewaluacji oznacza to wyjście poza informacje statystyczne i dane deklaratywne, a włączenie do badań etnografii: wywiadów swobodnych<sup>5</sup>, obserwacji zachowań, głębszego poznania środowiska (zarówno społecznego, jak fizycznego). Zebrane dane są porządkowane za pomocą *mapy empatii*. W przeprowadzonych badaniach na potrzeby ewaluacji on-going Priorytetu IV PO KL zrealizowanej dla Narodowego Centrum Badań i Rozwoju wykorzystane zostały do tego głównie studia przypadku prowadzone w szkołach wyższych, jak również wyniki badań ankietowych. W ich ramach wytypowano kategorie osób, dla których były projektowane rozwiązania mające za zadanie ulepszyć kształcenie praktyczne z udziałem pracodawców. Byli to: student kierunku turystyka w średniej wielkości mieście, studentka polonistyki na uniwersytecie, kierownik Akademickiego Biura Kariery, prezes firmy przetwórstwa spożywczego przyjmującej stażystów.

5 Wywiady swobodne rozumiem tutaj jako wywiady z mniej sformalizowanymi pytaniami niż powszechnie stosowane wywiady pogłębione skupione wokół zagadnień scenariusza.



Rysunek 1. Proces Design Thinking



Źródło: Opracowanie własne.

Drugim etapem w procesie Design Thinking jest **definiowanie problemu**. Jest to moment syntezy informacji zebranych na etapie empatii i sformułowanie problemu rzeczywiście istotnego z punktu widzenia grup docelowych, dla których projektowane są rozwiązania. Należy zwrócić uwagę, że problemy będą inaczej zdefiniowane niż te, które są formułowane w badaniach ewaluacyjnych. Powinny odnosić się wprost do potrzeb osób, których dotyczą. W badaniu prowadzonym dla NCBR sformułowano między innymi następujące problemy/pytania: W jaki sposób zachęcić studentów do podjęcia staży i praktyk bez podejmowania kosztownych działań? (problem z punktu widzenia kierownika Akademickiego Biura Kariery) Jak powinno być zorganizowane kształcenie (w tym praktyczne), by zaspokajało potrzeby studentów (było ciekawe i przydatne)? (problem z punktu widzenia studentów) Jak zorganizować staż/praktykę, by uczestniczący w nim studenci byli zaangażowani i chętni do nauki? (problem z punktu widzenia pracodawcy).

Sformułowanie problemu pozwala na wykonanie 3. kroku, jakim jest **generowanie pomysłów**. Wykorzystuje się do tego burzę mózgow (może ona przyjąć różnorodną postać), której celem jest wymyślenie jak największej liczby koncepcji, idei, rozwiązań. W tym miejscu należy podkreślić, że w technice *brainstorm* nie ma miejsca na krytykę czy też promowanie swoich pomysłów. Liczy się praca zespołowa, budowa rozwiązań na bazie wypowiedzi innych osób, kreatywność. Przydatna bywa również nuta szaleństwa i dużo uśmiechu, dzięki którym wychodzi się poza utarte szlaki myślenia. Po etapie generowania pomysłów następuje **faza oceny**, gdzie zespół demokratycznie wybiera rozwiązania do dalszej pracy. W założeniu proces kreatywny nie musi się ograniczać do jednej burzy mózgow, może mieć charakter cykliczny, angażować wiele osób itp. Kształt tego etapu zależy od celów, które przed nim postawimy. W badaniu przeprowadzonym przez Danae i Fundację Idea Rozwoju etap generowania pomysłów połączony został z kolejnymi etapami procesu Design Thinking. Przybrał formę warsztatu, na którym 4 interdyscyplinarne zespoły pracowały nad 4 wcześniej wymienionymi problemami. Każda z grup miała za zadanie w ciągu 25 minut przygotować listę minimum 50 pomysłów odpowiadających na

problem. Z jednej strony krótki czas sprzyja nieszablonowemu myśleniu, z drugiej nie daje możliwości bieżącej analizy powstających pomysłów ograniczając tym samym krytykę.

Następnym etapem pracy z wykorzystaniem metodyki Design Thinking jest **prototypowanie**, czyli tworzenie fizycznych modeli rozwiązań, które zostały wybrane przez zespół. Celem nie jest tworzenie gotowego produktu czy usługi, a raczej wizualizacja, która pomoże szybko zebrać opinie o pomysle. Prototypowanie może przybrać formę rysunku czy scenki opisującej działanie rozwiązania, modelu zrobionego z kartonu, styropianu czy innych fizycznych materiałów. Jest to kolejny element procesu kreatywnego, wiąże się z tzw. *myśleniem rękami*. Dzięki temu pobudzona zostaje prawa półkula mózgu odpowiedzialna za twórcze myślenie, obrazy, rytm, ale również za szersze spojrzenie. Kluczowe w tym etapie jest zebranie **informacji zwrotnej**, dzięki której energia i zasoby nie będą kierowane na rozbudowę nietrafionych pomysłów. Oczywiście nie gwarantuje to sukcesu, ale zwiększa szanse na jego osiągnięcie. W trakcie ewaluacji prowadzonej dla Narodowego Centrum Badań i Rozwoju zespoły przygotowywały prototypy rozwiązań, a feedback był na bieżąco zbierany od pozostałych uczestników. W ten sposób wypracowano rozwiązania, które zostały potem opisane w raporcie końcowym jako propozycje do wykorzystania przez uczelnie aplikujące o uzyskanie dofinansowania ze środków PO WER.

Końcowym etapem procesu Design Thinking jest **testowanie wybranego rozwiązania w środowisku użytkownika**. W klasycznych badaniach ewaluacyjnych wykonawcy stosunkowo rzadko dochodzą do tego momentu, gdyż ich rola dość często kończy się na etapie przygotowania rekomendacji, a więc wskazania konkretnego rozwiązania do wdrożenia. Jednak zbieranie informacji zwrotnej i ulepszanie prototypów jest już swoistą formą testu. Z pewnością zwiększa to wartość rekomendacji, która w trakcie pracy może ewoluować, dzięki uzyskiwanym w trakcie pracy informacjom zwrotnym.

### Jakie korzyści daje wykorzystanie Design Thinking w ewaluacji?

Po przetestowaniu podejścia Design Thinking w badaniu ewaluacyjnym można dojść do wniosku, że jest to metodyka wzbogacająca warsztat ewaluatora. Poprzez samo zastosowanie podejścia w procesie ewaluacji wygospodarowuje czas na pracę nad rekomendacjami. Jak pokazuje doświadczenie autorki, ma to niebagatelne znaczenie dla powodzenia całego procesu ewaluacji. Opracowanie dobrych rekomendacji wymaga czasu, a więc **zaplanowania w harmonogramie**. Włączenie do tego warsztatów Design Thinking w krótkim czasie dostarczających wiele pomysłów i rozwiązań, multiplikuje efekty.

Design Thinking pozwala na stworzenie szerokiej bazy pomysłów, które będą mogły zostać poddane analizie przez zespół badawczy, czy też dostarczą materiału do dyskusji na panelu ekspertów, jak to miało miejsce w prowadzonej przez autorkę ewaluacji. Włączenie DT pozwala spojrzeć na badane zagadnienie oczami osób, które korzystają z szeroko rozumianych usług publicznych. W badaniu ewaluacyjnym prowadzonym przez Danae i Fundację Idea Rozwoju

dla Narodowego Centrum Badań i Rozwoju zarówno ewaluatorzy po stronie wykonawcy, jak i zamawiającego zwrócili uwagę na problemy, z jakimi się stykają odbiorcy poddziałania 4.1.1 PO KL. Przestaliśmy mówić o uczelniach, a zaczęliśmy mówić o opiekunach praktyk, czy kierownikach Akademickich Biur Karier, ale przede wszystkim o studentach, studentkach i ich potrzebach. Pracodawcy stali się opiekunami stażystów i osobami zarządzającymi firmą.

Jednocześnie warsztat przeprowadzony wg metodyki Design Thinking oprócz funkcji kreatywnej, spełnia funkcję pogłębiającą zrozumienie. Dyskusje prowadzone w jego trakcie pozwoliły w badaniu dla NCBR jeszcze głębiej zdiagnozować badany problem i twórczo wykorzystać płynące z tego wnioski. Kolejne iteracje pomysłów pogłębiły analizę potrzeb, rozbudowały i urealniały kolejne pomysły.

Jak pokazało przeprowadzone badanie ewaluacyjne IV Priorytetu PO KL nie wszystkie pomysły i rozwiązania mają charakter innowacyjny czy rewolucyjny. Jednakże każdorazowo odpowiadają na niezaspokojone potrzeby. Jako przykład można przytoczyć wymyślone w trakcie warsztatu Laboratoria Zawodowe, które zyskały duże poparcie wśród uczestników warsztatu DT. To rozwiązanie funkcjonuje w części szkół wyższych na wybranych wydziałach (np. radio Kampus, gdzie studenci prowadzą własną stację radiową ucząc się jednocześnie dziennikarskiego rzemiosła, gazety wydziałowe), jednakże dyskusja nad pomysłem zaowocowała konkretnymi wskazówkami, jak można to doświadczenie wykorzystać na innych kierunkach.

## Zakończenie

Design Thinking jest podejściem, które może być szeroko wykorzystywane w tworzeniu społecznych innowacji czy nowych rozwiązań stosowanych przez administrację publiczną. Jednym z obszarów, w którym z pewnością można je z powodzeniem wykorzystać, jest ewaluacja. Włączenie Design Thinking do procesu oceny interwencji publicznych przede wszystkim pozwala spojrzeć na nią z wielu perspektyw, w szczególności z perspektywy osób, które z niej korzystają. W ten sposób powracamy do paradygmatu skoncentrowania projektowania na człowieku, który powinien być kwintesencją służby publicznej. Samo włączenie warsztatów Design Thinking staje się nie tylko źródłem pomysłów na rekomendacje, ale również polem eksperymentalnym, gdzie można testować odważne i innowacyjne rozwiązania. Tym samym warsztaty w podejściu Design Thinking mogą być – obok stosowanego powszechnie panelu ekspertów – jedną z technik wspomagających proces wnioskowania i przygotowywania rekomendacji.

## Bibliografia

**Bason C.**, *Leading Public Sector Innovation: Co-creating for a Better Society*, The Policy Press, Bristol 2010

**Brown T.**, *Change by design*, Harper Collins Publishers, New York 2009

*Design thinking starter*. Design Thinking Poznań.

*Design Thinking for Educators Toolkit* <https://www.ideo.com/work/toolkit-for-educators> (dostęp w dniu 13.06.2016)

**Leszczyńska B.** (red.), *Badanie ewaluacyjne o charakterze on-going identyfikujące dobre praktyki w realizacji praktycznych elementów kształcenia w projektach dofinansowanych w ramach IV Priorytetu PO KL oraz w innych działaniach w obszarze szkolnictwa wyższego*. Raport Końcowy z badania przeprowadzonego dla Narodowego Centrum Badań i Rozwoju przez Danae sp. z o.o. i Fundacja Idea Rozwoju, Warszawa 2015

**Tonhauser P.**, *Design Thinking Workshop*, 2016

*Co to jest design thinking?* [w:] <http://designthinking.pl/co-to-jest-design-thinking/> (dostęp w dniu 13.06.2016)

*The MindLab methods* [w:] <http://mind-lab.dk/en/> (dostęp w dniu 13.06.2016)



### BARBARA LESZCZYŃSKA

Dyrektor ds. Badań Ewaluacyjnych w firmie DANA E sp. z o.o. Absolwentka Instytutu Socjologii Uniwersytetu Warszawskiego. Członkini Polskiego Towarzystwa Ewaluacyjnego. Moderatorka warsztatów Design Thinking. Jej zainteresowania zawodowe skupiają się wokół metodologii badań społecznych, wykorzystania technik kreatywnych w ewaluacji oraz gender studies. Posiada ponad 8-letnie doświadczenie w projektowaniu i realizacji badań społecznych i ewaluacyjnych. W latach 2007-2014 związana z WYG PSDB sp. z o.o., gdzie pełniła funkcję Dyrektora Działu Badań Społecznych. W 2015 r. pracowała jako Project Manager w grupie IQS, gdzie była odpowiedzialna za rozwój segmentu badań dla administracji publicznej. W wolnej chwili prezeska stowarzyszenia Pozytywna Przestrzeń i działaczka Warszawskiej Kooperatywy Spożywczej. Wierzy, że partycypacja wspiera skuteczne działania i dobre rozwiązania.











## WERONIKA FELCIS

### Wstęp

Standardy ewaluacji w Polsce wspólnymi wysiłkami, koordynowanymi przez Polskie Towarzystwo Ewaluacyjne, wypracowało szerokie grono osób, będących nie tylko jego członkami w 2008 roku w ramach projektu ze środków Unii Europejskiej w ramach Programu Środki Przejściowe 2005 „Wzmocnienie zdolności ewaluacyjnych w sektorze publicznym”. Zaproszono do tego procesu akademików, pracowników administracji zlecających ewaluacje, badaczy z kluczowych wówczas firm ewaluacyjnych, przedstawicieli organizacji pozarządowych i niezależnych ewaluatorów także spoza Towarzystwa. Biorąc pod uwagę wysoką jakość wypracowanego materiału warto dokonać refleksji nad znaczeniem standardów dla kultury ewaluacyjnej w Polsce oraz jak powinien wyglądać system profesjonalizacji zbudowany na ich fundamencie. Niniejszy artykuł zawiera pewną propozycję wypływającą z konkretnego doświadczenia autorki, jest jednak równocześnie zaproszeniem do szerszej dyskusji.

### Znaczenie standardów dla kultury ewaluacyjnej

Maria Bustelo<sup>1</sup> wyróżnia trzy rodzaje norm regulujących badania ewaluacyjne: *zasady*, które wskazują czym powinni się kierować ewaluatorzy, *standardy* wskazujące jakie przymioty powinna mieć dobra ewaluacja i *wytyczne* tłumaczące w jaki sposób osiągnąć ideał. Artykuł ten omawia temat *standardów* oraz w jaki sposób wyprowadzić z nich *wytyczne* możliwe do zastosowania w praktyce zawodowej.

Przez dziesięć miesięcy w 2008 roku kilkudziesięciu przedstawicieli zaangażowanych było w wypracowanie standardów ewaluacji nie tylko na potrzeby oceny wdrażania polityk publicznych finansowanych ze środków unijnych, ale także na potrzeby sektora prywatnego i pozarządowego. Dokument podzielono na 16 punktów procesu badawczego w etapach planowania, realizacji i wykorzystania, a następnie przetłumaczono na język angielski. Polskie standardy ewaluacji są wzorem dla kolejnych towarzystw ewaluacyjnych w regionie Europy Środkowo-Wschodniej. Kiedy dokument był tworzony, środowisku w Polsce brakowało jasnego drogowskazu jak prowadzić proces ewaluacji. Jest on napisany normatywnym językiem, mając na celu edukację i promocję jasnych norm, przedstawiające standardy pracy dla zamawiających i badaczy. Jeszcze w 2011 roku

1 Bustelo M. *The potential role of standards and guidelines in the development of evaluation culture in Spain* [w:] *Evaluation* 12(4):437-453, 2006.

Polskie Towarzystwo Ewaluacyjne podjęło bardzo ciekawą próbę diagnozy rozwoju kultury ewaluacyjnej w Polsce<sup>2</sup>, jednak bez bezpośredniego odniesienia do stosowania *Standardów* w praktyce.

### Zarys sukcesów i porażek w procesie ewaluacyjnym w Polsce

Niniejsza część jest zbiorem refleksji opartych na rozmowach indywidualnych i grupowych, dyskusjach konferencyjnych i kongresowych, a także oddaniem głosu uczestnikom sesji konferencyjnej, której efektem jest niniejsza publikacja<sup>3</sup>. Niniejszy podział odzwierciedla strukturę *Standardów ewaluacji*, które są szeroko omawiającym zagadnienie ewaluacji dokumentem<sup>4</sup>.

1. **Zasadność podejmowania ewaluacji** – z pewnością dla zasadności zlecanych badań wsparciem są plany ewaluacji, które nie tylko wskazują argumenty za potrzebą podjęcia konkretnego badania i ukazują je w kontekście innych ewaluacji. Nie zawsze jednak są zgodne z rzeczywistą potrzebą w danym momencie. Częstym zagrożeniem pozostaje to, że ewaluacja ma badać sztucznie wyprodukowaną rzeczywistość środków unijnych, a nie realne potrzeby (wynikającą z odgórnie zaplanowanych planów finansowych, częściowo pokrywających się z potrzebami lokalnych wspólnot). Często też ewaluacje przeprowadza się w nieodpowiednich momentach cyklu polityki – tzn. takim, w którym nie da się zastosować większości wyników badania.
2. **Precyzyjne określenie celu, przedmiotu i zakresu badania** – mimo iż element ten jest bardzo dokładnie opisywany w dokumentach przetargowych i rozliczany w umowie, jest on często zbyt szeroko zakrojony, a wymagania za wysokie wobec pojedynczego badania (jak np. konieczność odpowiedzi na ponad 30 pytań badawczych, a przy tym przyznawanie punktów dla wykonawcy za dodanie kolejnych pytań w przeciągu zazwyczaj 3-6 miesięcy badania). Jeśli nawet wynika to z oszczędności środków publicznych, to tak szeroko zakrojone zakresy badawcze powodują wręcz ich marnowanie, bo badania takie nie są w stanie dostarczyć tak solidnych dowodów jakich dostarczają długie i skoncentrowane na wybranych celach badania naukowe.
3. **Realistyczne określenie sposobu przeprowadzenia ewaluacji** – przyznać należy, że znacznie sposobu przeprowadzania ewaluacji, a nie tylko cena, odgrywają obecnie znacznie większą rolę. Dostęp do wiedzy dzięki wieloletniej aktywności publikacyjnej kluczowych instytucji publicznych i naukowych jest duży, istnieje kilka sprawdzonych kursów i studiów z zakresu ewaluacji dla pracowników administracji i badaczy na poziomie wprowadzającym, ale wciąż ewaluatorzy nie stosują szerokiej gamy podejść (najczęściej stosowane jest „podejście deskryptywne” – tzw. *opisz to, co widzisz*). Wydaje się, że znacznie częściej należałoby

2 Katalog PTE 2011. Wydaje się, że było wówczas jeszcze za wcześnie na takie podsumowania. Obecnie jednak można by podjąć taką próbę wspólnym wysiłkiem.

3 By móc jednak stawiać twarde wnioski należałoby przeprowadzić pogłębioną metaewaluację rynku w oparciu o badania terenowe z interesariuszami omawianego procesu.

4 Przedmiotem refleksji są jedynie badania oceniające wdrożenie i rezultaty Polityki Spójności w Polsce.

stosować tryb dialogu technicznego lub negocjacji z ogłoszeniem lub uelastyczyć PZP<sup>5</sup> by rozbudzić kreatywność obu stron.

4. **Określenie przejrzystych kryteriów wyboru ewaluatora** – jest to bolączka wielu zamawiających. Mimo stosowania różnych rozwiązań wciąż trudno o sprawiedliwe, transparentne i użyteczne kryteria wyboru ewaluatora. Z pomocą może tu przyjść do pewnego stopnia stosowanie rozwiązania zaproponowanego w innym artykule tego tomu – podziału na role seniora i juniora i jednocześnie restrykcyjne domaganie się uczestnictwa specjalistów z obszarów interwencji w zespole badawczym.
5. **Zapewnienie niezależności ewaluacji** – obok wielu nacisków, z którymi zmagają się ewaluator podstawą konfliktu jest fakt iż płatnikiem za oddane badanie jest ten, kogo ewaluujemy. W tym kontekście bardzo istotne jest prowadzenie badania partycypacyjnego, włączającego zamawiającego w proces badawczy, ale także przydatna byłaby rola zewnętrznego arbitra w sytuacjach konfliktowych<sup>6</sup>.
6. **Zapewnienie zgodności z zasadami etycznymi** – niestety etyka w polskich badaniach ewaluacyjnych to często biała plama (złośliwi sugerują, że czarna dziura). Zapewnienie przestrzegania warunków umowy, przestrzegania anonimowości osób i poufności materiału czy apolityczność badań wydają się oczywistościami, o których nie trzeba pisać. W rozmowach bezpośrednich i nieformalnych pojawia się jednak wiele dylematów, których rozwiązań powinniśmy szukać wspólnie w szerokim środowisku, również poprzez interdyscyplinarne cykle dyskusji, szkoleń i kursów.
7. **Zapewnienie wysokiej jakości ewaluacji** – nie da się zapewnić wysokiej jakości rezultatów badania ewaluacyjnego, gdy biorąc pod uwagę kryteria wyboru ofert konieczne jest rozszerzenie i tak bardzo rozległego zakresu badania o trzecią jego część by wygrać zamówienie. Czas, w którym trzeba odpowiedzieć na tak wiele pytań badawczych często nie daje możliwości odpowiedniego pogłębienia tematu. Dlatego by zapobiegać wszelkim nierealistycznym anomaliiom, ważną jest rolę zewnętrznego eksperta przy tworzeniu dokumentów przetargowych, a w trakcie badania wymaganie raportów cząstkowych by wymagać od ewaluatora wyczerpujących odpowiedzi na każde z pytań.
8. **Zapewnienie trafności i rzetelności badań ewaluacyjnych** – by zapewnić trafne wyniki zacząć należy od stawiania hipotez opartych na badaniu gabinetowym bądź odtwarzaniu teorii programu, a nie tylko od tego, co wpisane zostało w dokumenty przetargowe. Z drugiej strony zlecający powinien posiadać narzędzia monitoringu rzetelności zbieranych podczas badań terenowych danych niewykraczających poza normy etyczne.

5 To ostatnie rozwiązanie to temat skomplikowany, który jednak notorycznie powraca w wielu rozmowach i warto zastanowić się nad organizacją spotkań w gronie zamawiających, badaczy i przedstawicieli Urzędu Zamówień Publicznych.

6 Po ustaleniu odpowiedniego regulaminu i transparentności procesu wyboru arbitrów oraz zasad ich funkcjonowania.

- 9. Dbanie o partycypacyjność i jakość współpracy w procesie ewaluacji** – przez lata używanie partycypacyjnego podejścia do ewaluacji nie spotykało się z zadowalającym poziomem zainteresowania, teraz jednak coraz częściej w ewaluacjach pojawiają się elementy odtwarzania teorii programu, warsztatów strukturyzujących czy paneli eksperckich przed oddaniem raportu. To dobry kierunek, gdyż pozwala zamawiającym znacznie bardziej uczestniczyć w procesie badawczym, jaśniej komunikować potrzeby informacyjne, a obie strony uczy zrozumienia dla każdej z perspektyw. Jednak wciąż brakuje uelastycznienia procesu badawczego, by móc tym lepiej odpowiedzieć na pytania badawcze.
- 10. Przestrzeganie warunków umowy** – niezmiennie trudnością dla wykonawców pozostaje brak elastyczności warunków umowy i możliwości ich demokratycznych negocjacji, nawet przy udziale arbitra. Z istotnych elementów umowy, jakie środowisko ewaluacyjne wspólnie wypracowało w 2008 roku jak na razie zaledwie kilka z ośmiu zaproponowanych postulatów doczekało się sporadycznego bądź częstego stosowania:
1. *przewidywanie możliwości modyfikacji sposobu realizacji badania, na wypadek zaistnienia nieprzewidzianych okoliczności, o ile sprzyja to realizacji celów ewaluacji,*
  2. *wskazanie sposobów rozstrzygnięcia ew. sporów i/lub instytucję arbitrażową,*
  3. *zapisy dot. praw autorskich, w tym ochronę osobistych praw autorskich wykonawców ewaluacji oraz zapewnienie ochrony praw autorskich twórców narzędzi badawczych wykorzystywanych w trakcie ewaluacji,*
  4. *gwarancję niezależności eksperta w kwestiach merytorycznych,*
  5. *precyzyjne określenie zasad i terminów odbioru produktów przewidzianych w umowie,*
  6. *dopuszczenie możliwości składania wyjaśnień ze strony wykonawcy w przypadku pojawienia się problemu nieterminowego lub nienależytego wykonania przedmiotu umowy,*
  7. *ochronę interesów osób, których ewaluacja dotyczy oraz zapewnienie skutecznej anonimizacji danych personalnych i/lub instytucjonalnych,*
  8. *precyzyjne określenie, które dane otrzymywane przez wykonawcę od zlecającego są poufne.*
- Źródło: Standardy ewaluacji PTE

W takich warunkach trudno o relacje partnerskie jak to było przedstawiane w Standardach ewaluacji. Dodać można, że także wyceny kar są kontrowersyjnym punktem, często nieodpowiadającym należycie rzeczywistości. Można także przewrotnie zapytać jeśli stosujemy kary, to jakich nagród, niekoniecznie pieniężnych można się spodziewać od zamawiającego? Czy taką nagrodą za świetnie przeprowadzone badanie byłoby zorganizowanie prezentacji raportu w szerszym gronie interesariuszy lub zgłoszenie do PTE, że taki oto raport należałoby szerzej przedstawić w środowisku w mediach lub Polskim Przeglądzie Ewaluacyjnym?

- 11. Dbanie o jakość relacji ewaluator – badani** – problem władzy w procesie badawczym (*power or empower*) nie jest poruszany w polskich badaniach ewaluacyjnych. A jednak to, na

ile badani mogą i chcą się podzielić pełną informacją, w sposób określony w zaplanowanym badaniu jest często sprawą kluczową dla wniosków. Z kolei niejednokrotnie zdarza się nam, badaczom, zostawić po sobie bardzo złe wrażenie nie informując nawet badanych respondentów o wnioskach z badania.

**12. Zachowanie przejrzystości struktury i języka raportu ewaluacyjnego** – mimo ciekawego badania na temat języka raportów ewaluacyjnych powstałego w 2012 roku, bardzo niewiele w tym obszarze się zmieniło. Brakuje osób dobrze piszących w zespołach badawczych, brakuje kursów i szkoleń z tego zakresu, brakuje często korekty językowej raportu przed jego oddaniem. Rozumienie kluczowych dla raportu pojęć nie jest uzgadniane między zlecającym, a ewaluatorem przed jego zamknięciem. Raporty są często nudne, zbyt długie, ciężko się je czyta nie tylko dyrektorom departamentów, ministrom czy politykom. Przyznać trzeba, że nawet badacze niechętnie wracają do swoich prac, a dla badaczy spoza obszaru ewaluacji są one po prostu niezrozumiałe. Z drugiej strony raport jest podstawą rozliczenia umowy, a więc odpowiedzi na wszystkie pytania badawcze. Zbyt szeroki zakres badań powoduje powstawanie zbyt szerokich, trudnych do czytania raportów ewaluacyjnych, a restrykcyjne tempo narzucone na badaczy nie pozwala im stworzyć dodatkowych form przekazu wniosków. W tym obszarze wspólnymi siłami środowisko musi jak najszybciej wypracować użyteczne usprawnienia, tak by zespoły badawcze były zmotywowane i miały czas stosować wielokrotne przeczytanie tekstu, wzajemne recenzowanie jego fragmentów, odważne skracanie jego ilości, korektę językową specjalisty.

**13. Zapewnienie rzetelności raportu ewaluacyjnego** – problem samozwańczych „firm ewaluacyjnych” bez odpowiedniego przygotowania merytorycznego i zapewnienia odpowiedniej jakości badaniom na poziomie lokalnym, jak międzyregionalnym niewiele się zmienił od czasu kiedy o nim pisałam w 2014 roku<sup>7</sup>. Dobrze znane są konkretne firmy czy kancelarie prawne, które psują rynek i dobre imię ewaluacji powodując niechęć do tego procesu wśród zamawiających. Wiele firm, zbyt małych by znaleźć sprawdzonych ekspertów w danym obszarze, składa oferty na pełen zakres przetargów, nie specjalizując się. To powoduje, że *ekspertem ds. ewaluacji* jest już niemal każdy bez względu na doświadczenie, wiedzę, umiejętności. Nie oskarżam tu jednak jedynie badaczy. Zamawiający wymagają po kilkumiesięcznej równoległej pracy nad wieloma obszarami badania raportów o jakości i pewności wniosków jak w publikacjach naukowych, nad którymi pracuje się latami. Dostarczenie rzetelnych dowodów w ewaluacji, jak pokazują najlepsze praktyki, może odbywać się tylko i wyłącznie dla jasno postawionych tez.

<sup>7</sup> Felcis W. Czynniki wpływające na użyteczność rekomendacji – w poszukiwaniu rozwiązań problemu [w:] *Monitorowanie i ewaluacja polityk publicznych. Doświadczenia przeszłości wyzwaniem dla przyszłości*. UM Województwa Opolskiego, 2015.

- 14. *Formułowanie precyzyjnych wniosków i rekomendacji*** – wyjątkowym, docenianym również zagranicą rozwiązaniem w systemie ewaluacji Polityki Spójności jest Zintegrowany System Zarządzania Wnioskami i Rekomendacjami z Badań Ewaluacyjnych. W dużej mierze dzięki tak szeroko zakrojonemu monitoringowi mamy dowody na to, że wnioski z badań często są niespójne z jego celami, jest ich zbyt dużo, są bardzo rozdrobnione, a wyprowadzane rekomendacje są miałkie, bezużyteczne i nieciekawe. Te problemy jednak często wynikają z zaniedbań na wcześniejszym etapie procesu badawczego – braku odpowiedniego zdiagnozowania potrzeb informacyjnych Zamawiającego na początku badania, których zaspokojenie nie jest sprawdzane na wcześniejszym etapie niż oddanie roboczej wersji raportu końcowego. Dbanie o zrozumiały sposób wyprowadzenia oceny końcowej i transparentność wyprowadzenia wniosków z materiału badawczego i użyteczność rekomendacji to fundamenty procesu ewaluacyjnego.
- 15. *Konsultowanie wniosków i rekomendacji z interesariuszami*** – ten punkt Standardów jest chyba najczęściej pomijanym punktem, jeśli nie brać pod uwagę bezpośrednio zaangażowanych stron – zamawiającego/płatnika i zespołu badawczego. Dyskusja nad ich trafnością, doprecyzowaniem rekomendacji, usunięciem rekomendacji ewidentnie nietrafnych oraz niemożliwych do wdrożenia odbywa się w bardzo wąskim gronie. Protokół rozbieżności opinii jest narzędziem wykorzystywanym w znikomym procencie przypadków. Demokratyczny profesjonalizm, o którym pisze Schwandt<sup>8</sup> - umiejętność uspołeczniania wyników ewaluacji jest czymś, o czym powinniśmy ze sobą częściej umieć rozmawiać.
- 16. *Wykorzystanie rekomendacji*** – narzędzie monitoringu stworzone przez Krajową Jednostkę Ewaluacji pozwala na wykorzystywanie użytecznych rekomendacji przez jednostki na wszystkich szczeblach systemu wdrażania Polityki Spójności. Przydatne byłoby jednak, oprócz wpisywania rekomendacji do systemu, *częstsze organizowanie spotkań z adresatami rekomendacji, w celu wskazania im konkretnych zysków i korzyści, lepszego zrozumienia przez nich istoty proponowanych modyfikacji, wzmocnienia ich akceptacji oraz motywacji do pracy na rzecz zmiany.*

8 Schwandt T. *Reconstructing professional ethics and responsibility: Implications of critical systems thinking* [w:] *Evaluation* 21(4): 462-466, 2015.

## Wnioski

Należy podkreślić, że każda ze stron – zarówno zamawiający, jak i wykonawcy są w stanie lepiej zadbać o wyplewienie patologii i podniesienie standardów ewaluacji.

Po stronie zamawiających należy zwrócić uwagę na niezależność funkcji płatnika od zlecającego badanie oraz konieczność zachowania marginesu elastyczności podpisywanych umów. Warto korzystać z narzędzi takich jak protokoły rozbieżności czy instytucji arbitrażowej w razie pojawiających się konfliktów. Pojawienie się strony trzeciej czy transparentność procesu oddania raportu może być jednak rozwiązaniem ostatecznym. Bez jasnego zakomunikowania potrzeb informacyjnych lub ich warsztatowego odkrycia na początku ewaluacji (z uwzględnieniem przeznaczonego na to budżetu czasowego), a także regularnego powracania do nich w trakcie badania (np. poprzez raporty cząstkowe czy panel konsultujących rekomendacje przed oddaniem raportu), wnioski z ewaluacji mają znacznie mniejsze szanse być wykorzystane.

Z drugiej strony grzechem wielu badaczy funkcjonujących na rynku wydaje się być absolutna bezkrytyczność wobec własnych umiejętności i traktowanie wszelkich warunków zleceń jak wykonywalnych usług. Na szczęście coraz więcej profesjonalistów specjalizuje się w wybranych przez siebie obszarach rezygnując ze zleceń, które są źle rozpisane bądź nie pokrywają się z ich kompetencjami, sygnalizując patologie rynku instytucjom monitorującym rynek. Dzięki debacie na temat profesjonalizacji wiemy, że ewaluator to nie tylko badacz, który bardzo dobrze zna warsztat metodologiczny i teorie ewaluacji, ale także w sposób jasny i wiarygodny dokonuje wartościowania przedmiotu ewaluacji. Wartościowanie jest naturalnym elementem procesu ewaluacji, a umiejętność komunikowania wniosków i rekomendacji jest jednym z fundamentalnych umiejętności ewaluatora. Ewaluator powinien patrzeć nie tylko na przedmiot badania, ale także rozumieć mechanizmy przyczynowe i kontekst w jakim funkcjonuje przedmiot badania. Do tego konieczne jest jednak myślenie systemowe i podstawy socjologii analitycznej. Przede wszystkim jednak ewaluator-profesjonalista analizuje swoje ograniczenia i nieustannie się doksztalca, budując kulturę ewaluacyjną w swoim środowisku i na poziomie krajowym dzieląc się swoimi doświadczeniami i czerpiąc z doświadczeń innych.

Trzecim elementem na jaki należy zwrócić uwagę jest dotkliwy brak szkoleń dla doświadczonych badaczy w naszym kraju. Mimo studiów i kursów z zakresu podstaw ewaluacji rzadko można znaleźć szkolenia podnoszące wiedzę z zakresu konkretnych elementów ewaluacji. Jest to wręcz zaskakujące w dobie internetu, webinarów, e-platform i kursów on-line. Polskie Towarzystwo Ewaluacyjne w zeszłym roku przeniosło swój podstawowy kurs *Kuźni ewaluacji* na e-platformę. Do trendu ogólnokrajowego jednak jeszcze daleko. Jak wynika z powyższego szkicu szczególnie brakuje szkoleń z zakresu etyki, aksjologii czy współpracy z językoznawcami.



## Elementy systemu profesjonalizacji ewaluacji

Kultura to suma zakorzenionych przekonań, idei, wartości i wiedzy, które razem konstituują fundament dla działań społecznych<sup>9</sup>. Kulturę ewaluacyjną zaś charakteryzuje przede wszystkim otwartość na proces uczenia się poprzez omawianie problemów i wspólne wypracowywanie rozwiązań<sup>10</sup>. Jednym zdaniem ewaluator, który nie poddaje własnych umiejętności ewaluacji i nie ulepsza wspólnie z pozostałymi aktorami środowiska, w jakim funkcjonuje, to *szewc, który bez butów chodzi*.

Na podstawie wniosków wyciągniętych z literatury dotyczącej profesjonalizacji ewaluacji, dyskusji toczących się w europejskim, brytyjskim i polskim towarzystwie ewaluacyjnym, a także w europejskiej sieci towarzystw NESE proponuję następujące kroki w rozwoju profesjonalizacji środowiska ewaluatorów i budowaniu kultury ewaluacyjnej.

### 1. Stosowanie standardów ewaluacji w praktyce badawczej.

Na rynku funkcjonuje co najmniej kilka firm, które w swoich ofertach mają zapis, że przestrzegają Standardów ewaluacji PTE. Co to jednak oznacza w praktyce i czy są z tego rozliczane nie jest do końca jasne. W literaturze przedmiotu można się spotkać z opinią, że analiza stosowalności standardów przebiega na poziomie metaewaluacji, badanie takie zazwyczaj jednak opiera się przede wszystkim na badaniach rynkowych lub analizie raportów z badań. Dlatego proponuję zastosowanie trzech elementów:

- a. Analiza ze strony jednostek ewaluacyjnych po stronie zlecających, jak i w firmach badawczych poziomu uwzględnienia standardów ewaluacji w ramach wszystkich etapów procesu badawczego.
- b. Przeprowadzenie metaewaluacji bardziej kompleksowej niż badanie rankingujące jakie podjęło Polskie Towarzystwo Ewaluacyjne w ramach badań zawartych w Katalogu PTE w 2011 roku, jednak z bardziej szczegółowym uwzględnieniem oceny funkcjonowania Standardów ewaluacji PTE na rynku.
- c. Refleksja w środowisku czy i jak uwzględniać standardy ewaluacji w dokumentach przetargowych. Być może rezultatem takiej dyskusji byłby wniosek, że potrzebny jest wspólnie wypracowany kodeks etyczny czy certyfikat dla ewaluatorów polityk realizowanych z funduszy unijnych, a nie tak ogólny i normatywny dokument jak Standardy. Wymaga to jednak wspólnej dyskusji.

### 2. Profil Kompetencji Ewaluatora wypracowany w demokratycznym procesie.

Standardy ewaluacji PTE, jak to zostało już wielokrotnie podkreślone, są dokumentem ogólnym odnoszącym się do procesu, a nie jednostki. Co w praktyce oznacza przestrzeganie standardów ewaluacji dla ewaluatora? Jakie kompetencje w zakresie każdego z 16 punktów

<sup>9</sup> Makins, M., ed. *Collins English Dictionary* London: HarperCollins Publications, 1995.

<sup>10</sup> Loud M. *Setting standards and providing guidelines. The means towards what end?* [w:] *Evaluation* 10(2): 237-245.

standardów miałyby on posiadać? Jak „owskaźnikować” koncepcje zawarte w standardach? Naturalnym rozwinięciem standardów w narzędzie do autorefleksji jest Profil Kompetencji Ewaluatora, o którym pisałam po konferencji w Opolu w 2014 roku. Dokument taki, wypracowany partycypacyjnie w brytyjskim towarzystwie ewaluacyjnym (UKES) i zaadaptowany również w procesach profesjonalizacji w europejskim towarzystwie (EES) może być tylko pewnym wzorem, z pewnością jednak jego użyteczność motywuje do podjęcia podobnych wysiłków w Polsce.

### 3. Wzajemna Ocena Kompetencji Ewaluatora w oparciu o wspólne wartości.

*Voluntary Evaluator Peer Review* Polskie Towarzystwo Ewaluacyjne tłumaczy jako Wzajemną Ocenę Kompetencji Ewaluatora. Recenzowanie wzajemnej pracy bądź kompetencji odbywa się w Czechach i Wielkiej Brytanii (systemy te omawiane były na XI Konferencji Ewaluacyjnej przez reprezentantów tamtejszych towarzystw) i jest obecnie testowane w pilotażu na poziomie europejskim w EES. Tak skonstruowany proces recenzji jest oparty o zasadę autorefleksji i wzajemnego uczenia się. Taka ocena to narzędzie znacznie bardziej demokratyczne niż certyfikacja, która odbywa się w Kanadzie bądź Japonii. To drugie rozwiązanie spotyka się z dużą dozą krytyki i istotnie musiałyby być szczegółowo dopasowane do wąskiego grona zawodowego funkcjonującego w jasno ograniczonym obszarze. Ewaluacja jednak jak wiemy dotyczy szerokiej gamy instytucji, dziedzin i evaluandów. WOKE wydaje się więc być mniej kontrowersyjnym, a wciąż bardzo konkretnym krokiem w stronę profesjonalizacji uczestników rynku.

### 4. Edukacja

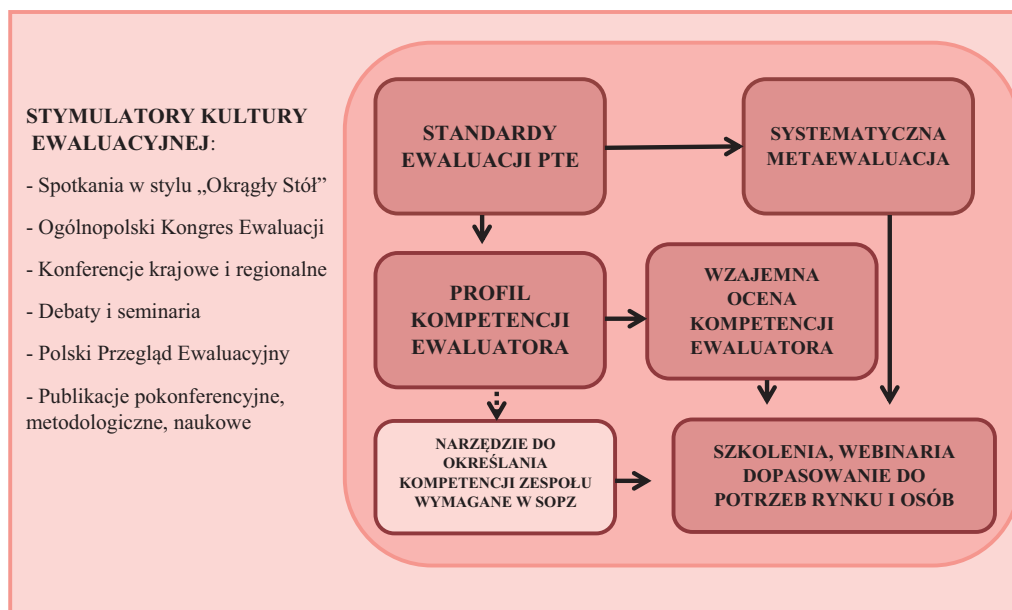
Jak zaznaczone zostało na wstępie tego podrozdziału podstawową wartością kultury ewaluacyjnej jest samodoskonalenie. Kolejnym krokiem jest zatem wypełnienie „białych plam” poprzez szkolenia i seminaria podnoszące kompetencje w diagnozowanych procesie metaewaluacji rynku i kompetencji ewaluatorów w obszarach ich braków.

### 5. Elementy wspierające system profesjonalizacji

Poza partykularnymi szkoleniami potrzebny jest także stały dostęp do wiedzy, najnowszych badań, najświeższych debat – to z kolei wiąże się z dobrym usieciowieniem środowiska, kreatywnością, otwartością i zaangażowaniem poszczególnych jego członków.

To w jaki sposób wpływają na siebie poszczególne elementy przedstawione jest na Rysunku 1.

### 6. Standardy -> zastosowanie do metaewaluacji -> wykorzystywane do szkoleń

**Rysunek 1. Model systemu profesjonalizacji ewaluacji w Polsce**

Źródło: Opracowanie własne.

W artykule o użyteczności rekomendacji<sup>11</sup> znaleźć można trzy propozycje działań wspierających profesjonalizację ewaluacji w Polsce: spotkania w stylu „okrągłego stołu”, gdzie w ramach partnerskich rozmów zamawiający i wykonawcy rozmawiają o możliwych usprawnieniach rynku, monitoring zachowań jawnie przekraczających standardy i certyfikację ewaluatorów jako najwyższy poziom profesjonalizacji. Dwa pierwsze postulaty Polskie Towarzystwo Ewaluacyjne realizuje samodzielnie, a cykliczne Wieczory Ewaluatora podjęte w tym roku w Warszawie, mają szansę rozprzestrzenić się na inne miasta, gdzie ewaluatorzy z różnym doświadczeniem mogliby się wymieniać swoimi spostrzeżeniami i pomysłami na działania w zupełnie nieformalnej atmosferze. Monitoring nie jest usystematyzowanym działaniem, ale kilkakrotnie Polskie Towarzystwo Ewaluacyjne podejmowało kroki by przekonać zlecających do zmiany nierealnych warunków podjęcia badania<sup>12</sup>. Jeśli jednak chodzi o certyfikację wydaje się, że to rozwiązanie możliwe do zastosowania dla wąskiego grona wyspecjalizowanych w określonym schemacie ewaluacji badaczy. Towarzystwo, które zrzesza ewaluatorów zajmujących się tak wieloma dziedzinami i obszarami, stosujących często wykluczające się podejścia ze

11 Felcis W., *Czynniki wpływające na użyteczność rekomendacji – w poszukiwaniu rozwiązań problemu*, [w:] *Monitorowanie i ewaluacja polityk publicznych. Doświadczenia przeszłości wyzwaniem dla przyszłości*, UM Województwa Opolskiego, Opole 2015.

12 Od czasu powstania wspomnianego artykułu nie zdarzyło się by zlecający domagali się podjęcia przez PTE roli arbitra.

względem na konteksty, w jakich pracują, może zaproponować narzędzie refleksji nad własnymi kompetencjami w sposób demokratyczny i wzmacniający. Taki sposób prowadzenia wzajemnej recenzji, spotyka się z krytycznym głosem, iż jest to proces zbyt ogólny i zbyt otwarty. Czy jednak zasłużenie, jeśli wciąż nie umiemy mówić o swoich ograniczeniach, porażkach i inwestować w ciągłe kształcenie by być coraz lepszym specjalistą?

### Zakończenie

Powstałe w 2008 roku pod auspicjami Polskiego Towarzystwa Ewaluacyjnego Standardy ewaluacji były rezultatem wielokrotnych spotkań szerokiego grona osób tworzących środowisko ewaluatorów w Polsce. Efekt ich prac jest doceniany do dzisiaj nie tylko w Polsce, ale także zagranicą jako jeden z najlepszych przykładów tego rodzaju dokumentu. Pozostają one jednak kodeksem normatywnym, zakrojonym na szeroką skalę generalizacji i jak dotąd niezwyfikowanym pod względem wdrożenia.

Po ośmiu latach od powstania Standardów ewaluacji PTE i intensywnym rozwoju międzynarodowej debaty na temat profesjonalizacji ewaluacji wiadomo, że dokument jest dobrym początkiem rozpoczęcia naszych lokalnych dyskusji o rozbudowaniu systemu profesjonalizacji ewaluacji w Polsce.

Polskie Towarzystwo Ewaluacyjne, jako jedyna w Polsce instytucja budująca kulturę ewaluacyjną od ponad 15 lat, edukująca kolejne pokolenia badaczy, tworząca niezwykle szeroką platformę doświadczeń, planuje rozwinąć na ich bazie system profesjonalizacji, nad którym do wspólnej dyskusji zaprasza nie tylko swoich członków. Bez wsparcia Polskiego Towarzystwa Ewaluacyjnego i wspólnego zaangażowania środowiska wracać będziemy wciąż do tych samych tematów i wniosków na rzadkich spotkaniach konferencyjnych, a w codziennej pracy zmagać z konfliktami etycznymi samodzielnie.

## Bibliografia

- Bustelo M.**, *The potential role of standards and guidelines in the development of evaluation culture in Spain*, [w:] *Evaluation* 12(4)/2006, s. 437-453
- Felcis W.**, *Czynniki wpływające na użyteczność rekomendacji – w poszukiwaniu rozwiązań problem*, [w:] *Monitorowanie i ewaluacja polityk publicznych. Doświadczenia przeszłości wyzwaniem dla przyszłości*, UM Województwa Opolskiego, Opole 2015
- Loud M.**, *Setting standards and providing guidelines. The means towards what end?* [w:] *Evaluation* 10(2), s. 237-245
- Makins, M.**, *Collins English Dictionary*, HarperCollins Publications, London 1995
- Polskie Towarzystwo Ewaluacyjne, *Katalog PTE*, Warszawa 2011
- Polskie Towarzystwo Ewaluacyjne, *Standardy ewaluacji PTE*, Warszawa 2008
- Schwandt T.**, *Reconstructing professional ethics and responsibility: Implications of critical systems thinking*, [w:] *Evaluation* 21(4): 462-466, 2015



### WERONIKA FELCIS

Magister socjologii pod kierunkiem profesora Jarosława Górniaka (UJ, 2005). Zajmuje się badaniami ewaluacyjnymi, analizami polityk publicznych oraz ekspertyzami dla sektora publicznego od 2006 roku. Silnie zaangażowana w budowanie profesjonalizacji zawodu ewaluatora w Europie Środkowej. Jest członkiem Europejskiego Towarzystwa Ewaluacyjnego i grupy roboczej wypracowującej Voluntary Evaluator Peer Review, który przeszła jako jedyny ewaluator w Polsce. Prezes Polskiego Towarzystwa Ewaluacyjnego w latach 2014-2016, koordynuje współpracę krajów CEE w zakresie budowy kultury ewaluacyjnej.



## ANDRZEJ REGULSKI

### Wstęp

Połowa 2016 r. to dobry moment na podsumowanie trwającego od 2013 r. długiego procesu prac nad dokumentami programowymi na lata 2014-2020. Jednym z elementów tego procesu było zewnętrzne wsparcie eksperckie dla instytucji zarządzających, początkowo w formie ewaluacji ex-ante, a w kolejnych etapach - w formie wsparcia wyboru i szacowania wskaźników, ocen ex-ante instrumentów finansowych, czy wreszcie ewaluacji systemów wyboru projektów. Niniejsze opracowanie jest pierwszą próbą wyciągnięcia wstępnych wniosków z tych badań z perspektywy wykonawcy.

### Typologia badań (zleceń)

Można szacować, że łączna wartość rynku badań i połączonych z nimi usług doradczych związanych z procesem programowania wyniosła w latach 2013-2016 około 20 mln zł, na co złożyło się około 90-100 zleceń, w większości realizowanych na zasadzie prawa zamówień publicznych. Ewaluacje ex-ante oraz oceny ex-ante instrumentów finansowych przeprowadzono w ramach każdego programu operacyjnego. Nie wszystkie instytucje zarządzające decydowały się na zewnętrzne wsparcie podczas pracy nad wskaźnikami programowymi, z kolei niektóre ewaluacje dotyczące systemu wyboru projektów (szczególnie na poziomie regionalnym) nie zostały w połowie 2016 r. jeszcze nawet rozpoczęte.

Specyfika poszczególnych rodzajów zleceń oraz struktura polskiego rynku ewaluacyjnego miała wpływ na ostateczny wybór wykonawców. W praktyce okazało się, że zdecydowaną większość zleceń zrealizowały trzy konsorcja (zespoły eksperckie). Ten układ konsorcjów dotyczył przede wszystkim ewaluacji ex-ante programów i częściowo znalazł także odzwierciedlenie w ocenach ex-ante instrumentów finansowych. Zasadniczo, dużą koncentrację zleceń można tłumaczyć koniecznością zgromadzenia w zespołach badawczych bardzo specyficznej wiedzy i umiejętności, co z kolei oznaczało istotne bariery wejścia i w konsekwencji przełożyło się na występowanie zauważalnych korzyści skali - szczególnie w przypadku programów regionalnych.

Na cztery omawiane w tym opracowaniu kategorie zamówień można spojrzeć nie tylko z perspektywy wspomnianego już wymiaru formalnego, ale także punktu widzenia roli ewaluatorów (konsultantów) w procesie programowania. Po pierwsze, przesłanką do zamawiania tego typu usług były potrzeby jednostek administracji publicznej związane z dostępem do wiedzy eksperckiej, która nie zawsze była osiągalna w ramach poszczególnych urzędów (mi-



nisterstw). W ten postulat wpisują się wszystkie typy zleceń, jednak w szczególnym stopniu dotyczy on ocen ex-ante instrumentów finansowych, których rzetelna realizacja wymagała nie tylko wiedzy na temat programowania interwencji publicznej, co także dobrej znajomości rynków finansowych. Istotnym powodem do angażowania w proces programowania ewaluatorów (konsultantów) była także ich - przynajmniej potencjalnie - katalizująca rola w animowaniu dyskusji na temat planowanej interwencji, a także konieczność outsourcingu wybranych zadań związanych z procesem programowania poza administrację publiczną.

**Tabela 1. Typologia badań, w których stosowano podejście partycypacyjne**

typ badania	wymóg formalny	zewnętrzna wiedza ekspercka	animowanie dyskusji	„outsourcing” zadań
ewaluacje ex-ante programów operacyjnych	++	++	++	+
oceny ex-ante instrumentów finansowych	++	+++	++	++
wsparcie wyboru i szacowania wskaźników		++		++
ewaluacja systemu wyboru projektów	+	++	+	+

Źródło: Opracowanie własne.

### Ewaluacje ex-ante programów operacyjnych

Ze wszystkich czterech kategorii zleceń, to właśnie ewaluacje ex-ante powinny były mieć największy wpływ na zawartość dokumentów programowych, a tym samym na kształt wdrażanej interwencji. Wynika to z bardzo mocnego osadzenia roli ewaluatora ex-ante w całym procesie programowania w dokumentach Komisji Europejskiej, obligatoryjności przeprowadzenia tego typu badania oraz - najczęściej - partycypacyjnego modelu ewaluacji. Rozwiązanie to zostało określone jako wariant preferowanych w wytycznych Ministerstwa Rozwoju Regionalnego i miało, przynajmniej w teorii, przeciwdziałać problemom występującym we wcześniejszej perspektywie finansowej, kiedy proces ewaluacji nie mógł „nadążyć” za procesem programowania. W przypadku większości programów operacyjnych zastosowano więc rozwiązanie, w którym ewaluatorzy zaczęli swoją pracę już w połowie 2013 r., na bieżąco opiniując kolejne fragmenty powstających dokumentów.

Wydaje się, że to właśnie od tego, w jaki sposób instytucja zarządzająca była w stanie współpracować z ekspertami odpowiedzialnymi za ewaluację, zależała faktyczna użyteczność ich pracy. Szybkie tempo prac nad programami oznaczało, że często nie było czasu na opracowanie formalnej recenzji dokumentu, a znacznie bardziej przydatne okazywały się robocze komentarze i organizowane *ad hoc* spotkania. To na tym etapie była szansa na faktyczną wymianę poglądów pomiędzy ewaluatorami i autorami programu, która docelowo mogła podnieść jakość opracowywanego dokumentu. Efektem eksperckiego wsparcia było najczęściej uporządkowanie logiki interwencji oraz wsparcie w precyzyjnym zdefiniowaniu rezultatów poszczególnych działań (także na poziomie wskaźników – p. sekcja *Wsparcie wyboru i szacowania wskaźników*). Często, wsparcie ewaluatorów polegało nie tyle na bezpośrednim proponowaniu konkretnych rozwiązań, co na animowaniu dyskusji wśród instytucji biorących udział w programowaniu. Warto jednak przy tym podkreślić, że taki tryb współpracy wymagał elastycznego podejścia od wszystkich zaangażowanych stron i nie zawsze wynikał wprost ze specyfikacji zamówienia. Z perspektywy czasu można ocenić, że bardzo trafnym i użytecznym dla zamawiających rozwiązaniem okazała się możliwość zlecenia ewaluatorowi prac w ramach usług doradczych, rozliczanych na podstawie kart czasu pracy. W okresie poprzedzającym programowanie bardzo trudno w wyczerpujący sposób zapisać wszystkie potencjalnie przydatne formy zaangażowania doradców wspierających zespół programujący, tymczasem rozliczenie poprzez godziny doradcze pozwala właśnie na elastyczną realokację zasobów w zależności od potrzeb. Taki wariant stosowano niestety stosunkowo rzadko.

Próba bilansu ewaluacji *ex-ante* programów na lata 2014-2020 wymaga jednak także wskazania pewnych problemów, nad rozwiązaniem których warto zastanowić się w kontekście kolejnych okresów programowania. Najważniejsze wyznania wynikały z przyjętego partycypacyjnego modelu ewaluacji. Drugą stroną sytuacji, w której ewaluator czasami pełni rolę autora dokumentu programowego oznacza, że nie może on jednocześnie być postrzegany jako zewnętrzny recenzent. Jednocześnie, oczekiwania dotyczące zawartości raportu z ewaluacji *ex-ante* sprowadzały ją *de facto* do recenzji, która miała stanowić potwierdzenie dla zapisów programu operacyjnego. Myśląc o pracy nad dokumentami programowymi w przyszłości warto więc zastanowić się, czy lepszym rozwiązaniem nie byłaby rezygnacja z „partycypacyjnej” ewaluacji na rzecz zaangażowania zewnętrznych ekspertów jako konsultantów/doradców. Taka formuła wydaje się być szczególnie użyteczna np. podczas negocjacji programów z Komisją Europejską, ponieważ w przeciwnym razie ewaluator staje się w pewnym sensie „trzecią stroną” negocjacji, co z perspektywy instytucji zarządzającej nie musi być wcale sytuacją korzystną. W przeciwieństwie do ewaluacji wpływu interwencji, niezależność doradców/konsultantów nie powinna być wartością sama w sobie, jest ona zresztą bardzo trudna do zagwarantowania, gdy odbiorcą i finansującym zlecenie jest instytucja odpowiedzialna za dany dokument programowy.

Doświadczenia ewaluacji ex-ante programów na lata 2014-2020 wskazują jeszcze na kilka innych wyzwań. Tak jak zaznaczono wcześniej, elementem programów, na których ewaluacje ex-ante odcisnęły się w największym stopniu, były części projekcyjne, w szczególności opis logiki interwencji. Porównanie poszczególnych programów, przede wszystkim regionalnych, prowadzi jednak do wniosku, że przy często odmiennych sposobach sformułowania oczekiwanych efektów działań, interwencja w praktyce sprowadza się do realizacji typów projektów dopuszczonych na danym poziomie przez linię demarkacyjną. Zbadanie konsekwencji tej sytuacji może być ciekawym zagadnieniem do uwzględnienia w ewaluacjach mid-term oraz ex-post.

**Tabela 2. Ewaluacje ex-ante programów operacyjnych – podsumowanie**

Zalety	problemy/znaki zapytania
<ul style="list-style-type: none"> <li>wyjście poza perspektywę definiowania celów interwencji przez działania – definiowanie oczekiwanych rezultatów</li> <li>animowanie dyskusji o logice interwencji i oczekiwanych rezultatach</li> <li>zewnątrzne spojrzenie na powstający dokument – perspektywa spoza administracji</li> <li>„mózgi do wynajęcia”, niekoniecznie zewnętrzny ewaluator</li> </ul>	<ul style="list-style-type: none"> <li>czy ewaluator może wpływać na cele interwencji (wybór publiczny)?</li> <li>czy określenie cele interwencji (oczekiwane rezultaty) to rzeczywisty wybór publiczny, czy tylko ćwiczenie literackie?</li> <li>problemy z praktycznym zastosowaniem podejścia evidence-based               <ul style="list-style-type: none"> <li>zmieniające się otoczenie (poziom krajowy i europejski)</li> </ul> </li> <li>problem wewnętrznych relacji pomiędzy instytucjami zaangażowanymi w programowanie</li> </ul>

Źródło: Opracowanie własne.

### Wsparcie wyboru i szacowania wskaźników

Na zlecenia związane ze wsparciem doboru i szacowania wskaźników można spojrzeć w podobny sposób, jak na ewaluacje ex-ante. Opisane poniżej spostrzeżenia częściowo pochodzą właśnie z ewaluacji ex-ante, ponieważ w niektórych programach operacyjnych od ewaluatorów oczekiwano nie tylko opiniowania wyboru mierników i ich wartości docelowych, co także bezpośredniej współpracy z zespołem programującym.

Przed przejściem do omówienia zastrzeżeń, warto podkreślić, że zadania związane ze wskaźnikami, w szczególności kwantyfikacja oddziaływania programu, jest dobrym polem do współpracy pomiędzy administracją publiczną oraz zewnętrznymi ekspertami. Autorzy interwencji - co do zasady - znacznie lepiej od osób „z zewnątrz” procesu programowania znają szczegółowe założenia planowanej interwencji, z kolei eksperci są w stanie wesprzeć programu-

jących np. swoją wiedzą statystyczną lub ekonometryczną. Taka forma współpracy w ramach prac nad programami na lata 2014-2020 generalnie sprawdziła się, co potwierdza stosunkowo duża liczba dokumentów programowych (programów operacyjnych oraz szczegółowych opisów osi priorytetowych), w których wskaźniki produktu i rezultatu były szacowane przez zewnętrznych ekspertów.

W ramach tego typu zleceń występował jednak szereg problemów. Częściowo miały one charakter zbliżony do ewaluacji *ex-ante*, szczególnie jeśli prace nad wskaźnikami były formalnie lub nieformalnie włączone do tego procesu. Sytuacja, w której ewaluator (zgodnie z wytycznymi Komisji Europejskiej) powinien w raporcie z ewaluacji potwierdzić zasadność określonych w programie wartości docelowych, będąc jednocześnie autorem oszacowań, mogła powodować pewne zamieszanie. Problem ten dodatkowo potęgowało zaangażowanie w proces programowania różnych instytucji, które wysyłały ekspertom sprzeczne sygnały dotyczące oczekiwanych efektów pracy (np. instytucja zarządzająca oczekuje „ambitnych” oszacowań, a instytucja pośrednicząca - możliwe najbardziej „defensywnych”). To ponownie sugeruje, że dotychczasowy partycypacyjny model ewaluacji *ex-ante* niekoniecznie powinien być stosowany w przyszłości w niezmienionej formie. Szczególnie w przypadku pracy nad wskaźnikami bardziej sprawdza się zaangażowanie ekspertów jako zewnętrznych konsultantów/doradców.

Pozostałe wątpliwości dotyczące wsparcia w wyborze i szacowaniu wartości docelowych wskaźników wychodzą poza teoretyczny i praktyczny wymiar współpracy pomiędzy wykonawcą i zamawiającym. Podobnie jak w okresie programowania 2007-2013, także teraz można było zaobserwować objawy „polowania na wskaźniki”, tj. wyboru mierników, które są dostępne w źródłach statystyki publicznej, nawet jeśli nie zawsze oddają dobrze przyjętą w programie logikę interwencji. Takie postępowanie dodatkowo wzmacniało bardzo szybkie tempo prac nad programami operacyjnymi, przez co w praktyce nie było możliwości zaprojektowania badania, które dopiero zmierzyłoby oddziaływanie interwencji. To zresztą wymagałoby uruchomienia procedury w ramach warunków *ex-ante*, czemu decydenci byli co do zasady niechętni. Choć dostępność danych, szczególnie na poziomie regionalnym, w ostatnich latach znacznie się poprawiła, w ramach średniookresowych przeglądów programów warto szczegółowo zbadać, w których obszarach wskaźniki rezultatu strategicznego faktycznie odpowiadają oczekiwanym efektom działań.

**Tabela 3. Wsparcie wyboru i szacowania wskaźników– podsumowanie**

Zalety	problemy/znaki zapytania
<ul style="list-style-type: none"> <li>• dobry przykład outsourcingu zadania, w którym zewnętrznymi eksperci faktycznie mogą pomóc administracji</li> <li>• współpraca z jednostkami odpowiedzialnymi za monitoring (uzupełnianie kompetencji)</li> </ul>	<ul style="list-style-type: none"> <li>• problem „polowania na wskaźniki” – dostępność danych a faktyczne dostosowanie do logiki interwencji</li> <li>• czym powinna być wartość docelowa: różne bodźce poszczególnych instytucji?</li> <li>• częściowe oderwanie od procesu programowania (w przyszłości to może oznaczać problemy), szczególnie w sytuacji nadmiernego wskaźnikowania</li> <li>• niewielkie zainteresowanie decydentów, tymczasem określenie wartości docelowych wskaźników to de facto definiowanie interwencji w danym obszarze</li> </ul>

Źródło: Opracowanie własne.

### Oceny ex-ante instrumentów finansowych

W ramach pracy nad programami operacyjnymi na lata 2014-2020, obowiązkiem instytucji zarządzających było także zlecenie oceny ex-ante instrumentów finansowych. Dokument ten miał być opracowany przez zewnętrznych ekspertów i określić, optymalną formę finansowania w ramach poszczególnych priorytetów inwestycyjnych. Formuła ta wynikała z zapisów tzw. Rozporządzenia ogólnego i była konsekwencją decyzji na poziomie europejskim, aby polityka spójności w mniejszym niż dotychczas stopniu opierała się na wsparciu dotacyjnym. Zlecenia tego typu miały raczej formę ekspertyzy, a nie ewaluacji, choć często były realizowane w ścisłej współpracy z instytucjami zarządzającymi/pośredniczącymi.

W Polsce, rynek ocen ex-ante instrumentów finansowych zdominowały krajowe firmy od lat obecne na rynku ewaluacyjnym, które w zależności od potrzeb włączały do swoich zespołów osoby posiadające branżowe doświadczenie związane z rynkami finansowymi. Wyjątkiem był jedynie Program Operacyjny Infrastruktura i Środowisko, w którym pierwszy etap oceny ex-ante instrumentów finansowych opracował Bank Światowy, a drugi - Europejski Bank Inwestycyjny. Brak szerszego udziału przedsiębiorstw wyspecjalizowanych w doradztwie finansowym jest interesującą i jest dobrym tematem na pogłębioną analizę.

Pozytywną stroną ocen ex-ante instrumentów finansowych było niewątpliwie bardzo szczegółowe zbadanie logiki interwencji poszczególnych priorytetów inwestycyjnych, znacznie bardziej pogłębione niż miało to miejsce na poziomie ewaluacji ex-ante. Co ciekawie, rozstrzygnięcia na poziomie europejskim doprowadziły do paradoksalnej sytuacji, w której wsparcie dotacyjne de facto nie wymaga uzasadnienia, a wymaga się go jedynie w odniesieniu do instrumentów poza dotacyjnych, które powodują znacznie mniejsze zaburzenia wolnego rynku.

W pracach nad ocenami ex-ante instrumentów finansowych wstąpiły także pewne problemy. Wydaje się, że niekoniecznie trafnym rozwiązaniem była realizacja pełnych ocen na poziomie każdego z programów regionalnych, ponieważ każdy z nich wdrażany jest w bardzo zbliżonym otoczeniu instytucjonalnym, a specyficzne uwarunkowania mogły zostać uwzględnione w ramach jednego, większego ogólnokrajowego projektu badawczego. Niewątpliwie jednak, docelowo nie warto rezygnować z zewnętrznego wsparcia eksperckiego na poziomie prac nad szczegółową strategią inwestycyjną, która powinna już uwzględniać kontekst danego województwa. Dużym wyzwaniem był także brak wielu istotnych punktów odniesienia - np. przedłużające się prace nad oceną ex-ante instrumentów finansowych w ramach PO IŚ oraz ustawą o odnawialnych źródłach energii.

**Tabela 4. Oceny ex-ante instrumentów finansowych – podsumowanie**

Zalety	problemy/znaki zapytania
<ul style="list-style-type: none"> <li>• wybór optymalnych form finansowania dla poszczególnych działań i typów projektów</li> <li>• dużo głębsze przemyślenie logiki działań, w których ostatecznie przewidziano instrumenty finansowe (dodatkowe uzasadnienie wymagane dla instrumentów finansowych, a nie dla dotacji)</li> </ul>	<ul style="list-style-type: none"> <li>• nie wszystkie rozstrzygnięcia wymagają oddzielnej analizy w każdym województwie (problem spójności wniosków)</li> <li>• z góry określone preferencje decydentów, co ma miało przełożenie na przebieg i wyniki badań               <ul style="list-style-type: none"> <li>• duża niepewność rozstrzygnięć na poziomie krajowym (np. IF w programach krajowych) i europejskim (np. konstrukcja instrumentów off-the-shelf)</li> </ul> </li> </ul>

Źródło: Opracowanie własne.

## Ewaluacje systemu wyboru projektów

Jak zaznaczono wcześniej, w połowie 2016 r. ewaluacje systemów wyboru projektów w ramach większości polskich programów operacyjnych nie zostały jeszcze zakończone, a często nawet rozpoczęte. Na podstawie wstępnych i częściowych doświadczeń można jednak pokusić się o wstępne refleksje.

Po pierwsze, wydaje się, że to właśnie ewaluacje systemów wyboru projektów w największym stopniu wymagają podejścia możliwie zbliżonego do modelu partycypacyjnego. Wynika to z tego, że podlegające ewaluacji ex-ante dokumenty programowe zawierają zazwyczaj zapisy na dość dużym poziomie ogólności, przez co dopiero w dokumentach wdrożeniowych decyduje się, na ile przyjęte rozwiązania mają szansę przełożyć się na oczekiwane rezultaty. Jeśli więc zakładamy, że zewnętrzne wsparcie eksperckie w procesie programowania wnosi wartość dodaną, to jest ono potrzebne właśnie w trakcie prac nad systemem wyboru projektów – zarówno w ujęciu ex-ante, jak i mid-term (tj. po przeprowadzeniu pierwszych konkursów).

Po drugie, podobnie jak w przypadku ewaluacji ex-ante, głównym efektem współpracy zamawiającego i wykonawcy nie jest raport końcowy z badania, ale rozwiązania, które uda się wypracować w ramach bieżącej współpracy. Warto rozważyć, czy w przypadku tego typu zleceń produkt w postaci raportu końcowego jest w ogólnie potrzebny, ponieważ jego przygotowanie i uzgodnienie angażuje duże zasoby po stronie zamawiającego i wykonawców. Alternatywnym rozwiązaniem może być (czasami już stosowane) położenie większego nacisku na większą liczbę krótkich opracowań roboczych.

**Tabela 5. Ewaluacje systemu wyboru projektów – podsumowanie**

Zalety	problemy/znaki zapytania
<ul style="list-style-type: none"> <li>• szczególny przypadek partycypacyjnego modelu współpracy – tutaj ma największe uzasadnienie, ponieważ kryteria faktycznie decydują o ukierunkowaniu interwencji</li> <li>• faktyczne wsparcie wtedy, kiedy udaje się nawiązać rzeczywistą współpracę, co obejmuje także włączenie ewaluatora w obieg informacji (ex-ante i ex-post)</li> </ul>	<ul style="list-style-type: none"> <li>• nie ma sensu nadmierna koncentracja na raporcie, który nie jest kluczowym produktem badania – przede wszystkim chodzi o efekty bieżącej współpracy</li> <li>• problem zdefiniowania roli ewaluatora w całym procesie (formalnie zaangażowany przez jedną z zainteresowanych instytucji)</li> </ul>

Źródło: Opracowanie własne.

## Zakończenie

Podsumowując bilans zaangażowania ekspertów/konsultantów w proces programowania w perspektywie finansowej 2014-2020 należy podkreślić, że może być ono realnym wsparciem dla instytucji zarządzających. Warto jednak już teraz rozpocząć dyskusję na temat możliwych modyfikacji dotychczasowych modeli współpracy, w taki sposób, aby ewaluacja ex-ante i inne podobne do niej zlecenia nie były głównie realizacją określonego na poziomie europejskim wymogu formalnego. Wydaje się, że kluczem do sukcesu jest możliwie elastyczne włączenie ekspertów/konsultantów w prace nad dokumentami programowymi, optymalnie od samego początku, kiedy możliwe są jeszcze istotne zmiany. Doświadczenia prac nad programami na lata 2014-2020 sugerują, że ewaluacja ex-ante, której głównym produktem jest dostępny publicznie raport końcowy, niekoniecznie muszą prowadzić do faktycznej realizacji tego postulatu.

## Bibliografia

**Komisja Europejska**, *Guidance document on ex-ante evaluation*, Bruksela 2014

**Komisja Europejska i Europejski Bank Inwestycyjny**, *Ex-ante assessment methodology for financial instruments in the 2014-2020 programming period. General methodology covering all thematic objectives. Quick reference guide*, Bruksela 2014

**Ministerstwo Rozwoju Regionalnego**, *Zalecenia w zakresie ewaluacji ex-ante programów operacyjnych na lata 2014-2020*, Warszawa 2012

**Szlachta, J.** (red), *Ewaluacja ex-ante. Podsumowanie doświadczeń administracji polskiej*, Ministerstwo Rozwoju Regionalnego, Warszawa 2007

*Rozporządzenie Parlamentu Europejskiego i Rady (UE) nr 1303/2013 z dnia 17 grudnia 2013 r. ustanawiające wspólne przepisy dotyczące Europejskiego Funduszu Rozwoju Regionalnego, Europejskiego Funduszu Społecznego, Funduszu Spójności, Europejskiego Funduszu Rolnego na rzecz Rozwoju Obszarów Wiejskich oraz Europejskiego Funduszu Morskiego i Rybackiego oraz ustanawiające przepisy ogólne dotyczące Europejskiego Funduszu Rozwoju Regionalnego, Europejskiego Funduszu Społecznego, Funduszu Spójności i Europejskiego Funduszu Morskiego i Rybackiego oraz uchylające rozporządzenie Rady (WE) nr 1083/2006*





### ANDRZEJ REGULSKI

Od 2013 r. Dyrektor Zarządzający Imapp Sp. z o.o., w przeszłości Dyrektor Programu Ewaluacji Polityk Publicznych w Instytucie Badań Strukturalnych. Z wykształcenia ekonomista, absolwent Szkoły Głównej Handlowej w Warszawie. Dotychczas brał udział w ponad 50 badaniach ewaluacyjnych i innych projektach analitycznych realizowanych na zlecenie administracji rządowej i samorządowej. Specjalizuje się w zagadnieniach związanych z systemem instytucjonalnym polityki rozwoju w Polsce, innowacyjnością oraz energetyką. W ostatnich latach koordynował m.in. ewaluacje ex-ante Umowy Partnerstwa oraz czterech programów operacyjnych na lata 2014-2020 (Program Operacyjny Infrastruktura i Środowisko, regionalne programy operacyjne województw łódzkiego, podkarpackiego i pomorskiego), badania związane z aktualizacjami strategii rozwoju województw (małopolskiego, podkarpackiego, pomorskiego oraz warmińsko-mazurskiego), a także szereg analiz dotyczących budowy systemu wskaźników w dokumentach strategicznych i operacyjnych na poziomie krajowym i regionalnym.

## PIOTR STRONKOWSKI

### Wprowadzenie

Niniejszy artykuł poświęcony jest ocenie stanu ewaluacji polityki spójności w Polsce. Jest on napisany z perspektywy praktyka: ewaluatora z około dziesięcioletnim dorobkiem, i bazuje na doświadczeniach i obserwacjach autora. Jednak, jak na ewaluatora przystało, tam gdzie jest to możliwe artykuł bazuje na dostępnych badaniach i meta-ewaluacjach. Szczególnie wartościowy okazał się tu dorobek dr Tomasza Kupca, który w ramach prac nad swoim doktoratem przeprowadził niezwykle interesujące badania jakości i użyteczności ewaluacji w Polsce.

Artykuł koncentruje się głównie na ewaluacji polityki spójności. Wynika to z prostego powodu – pomimo budzenia się kultury ewaluacji również w innych obszarach polityk publicznych, to jednak cały czas w ramach polityki spójności prowadzonych jest najwięcej ewaluacji.

### Stan ewaluacji w Polsce

Istnieje przekonanie, wyrażane przez przedstawicieli administracji publicznej, krajowej, regionalnej i na poziomie KE, że ewaluacja w Polsce w ostatnich latach rozwijała się bardzo dynamicznie – szczególnie w porównaniu z innymi krajami Europy środkowo-wschodniej. Opinia ta jest częściowo uzasadniona. Porównując stan ewaluacji w roku 2016 do sytuacji np. w roku 2000 czy też 2004, można stwierdzić, że wzrosła ilość ewaluacji. Można też zaryzykować stwierdzenie, że poprawiła się jej jakość (choć w tym przypadku brak jest rzetelnych porównań). Większy też nacisk kładziony jest na wykorzystanie ewaluacji. Wzrósł potencjał ewaluacyjny zarówno firm badawczych, jak i jednostek zajmujących się ewaluacją w administracji publicznej. Upowszechnia się kultura ewaluacji, rozumiana jako potrzeba regularnego oceniania efektów polityk publicznych.

Pomimo tych pozytywnych zjawisk w środowisku ewaluatorów (reprezentujących zarówno wykonawców, jak i zamawiających) spotkać się można stosunkowo często z głosami wyrażającymi rozczarowanie ewaluacją. Nie jest to zjawisko ograniczone tylko do Polski – podobne głosy pojawiają się również w innych krajach, z dojrzałą kulturą ewaluacyjną. Rozczarowanie to dotyczy przede wszystkim jakości (rzetelności) prowadzonych ewaluacji, ich użyteczności i faktycznego wykorzystania w politykach publicznych. Wśród ewaluatorów budzi się rodzaj rozgoryczenia, że efekty ich pracy nie są wykorzystywane do faktycznego doskonalenia polityk publicznych. Pytanie, czy ta ocena jest subiektywna i jest wyrazem swoistej frustracji z niedocenywania dorobku pewnej grupy zawodowej, czy faktycznie wyniki ewaluacji zbyt często lądują na półkach w kolejnych gabinetach, w których nikt z nich nie korzysta.

## Wykorzystanie wyników ewaluacji

Oceniając faktyczne wykorzystanie wyników ewaluacji najlepiej odwołać się do wyników badań, co pozwoli uniknąć ryzyka nadmiernego subiektywizmu. Proponuję bazować na wynikach meta-ewaluacji Tomasza Kupca, jednego z niewielu zrealizowanych w Polsce systematycznych badań wykorzystania ewaluacji zrealizowanych w ramach RPO. W roku 2014 T. Kupiec wykonał analizę wykorzystania rekomendacji z ewaluacji wykonanych w latach 2007-2012 w trzech, celowo dobranych województwach. W tych trzech województwach zrealizowano łącznie 44 ewaluacje RPO. Spośród nich 35 zawierało rekomendacje (ok. 12,5 rekomendacji na raport). Jednak, jak wynika z analizy, zaledwie 40% rekomendacji zostało wdrożonych. Co więcej, szczegółowa analiza tych rekomendacji, które zostały wdrożone, pokazuje, że dotyczyły one głównie kwestii technicznych, organizacyjnych, usprawnienia sposobu wdrażania. Co ciekawe, wiele z nich dotyczyło promocji i informacji. Ich waga została oceniona przez badacza jako niska lub średnia.

To, że adresaci rekomendacji wdrażali głównie te o charakterze technicznym czy też organizacyjnym nie wynika z ich oporu przed wdrażaniem rekomendacji o charakterze strategicznym, np. podziału alokacji na poszczególne typy interwencji, znaczących modyfikacji w kształcie interwencji itd. Okazuje się, że takich rekomendacji niemal nie było w analizowanych badaniach. Natomiast wiele z nich, jak wynika z analizy T. Kupca, było na poziomie banału, trywialności, oczywistości.

W praktyce oznacza to, że ewaluacje faktycznie nie miały wpływu na kluczowe decyzje dotyczące podstawowych założeń interwencji, definiowania problemów i sposobów ich rozwiązania, celów, priorytetów czy też alokacji finansowej. Można oczywiście przypuszczać, że w badanym okresie tego typu decyzje nie były podejmowane. Okazuje się jednak, że było wręcz przeciwnie. W tym czasie we wszystkich trzech województwach dokonano wielu zmian w zapisach SzOOP. Ponadto, co najmniej raz zmieniono regionalne programy operacyjne, co wiąże się zwykle z istotnymi modyfikacjami (np. alokacji). Wszystkie te zmiany wynikały zapewne z refleksji zarządzających programami, że istniejące rozwiązania z jakiegoś powodu wymagają poprawienia, doskonalenia. Jednak tej refleksji nie wspomagały w większości przypadków wyniki ewaluacji.

Zdarzały się również przykłady niewłaściwego wykorzystania ewaluacji, kiedy to wnioski i rekomendacje były traktowane jako uzasadnienie dla zmian, choć faktycznie nie uzasadniały one wprowadzonych zmian. Przykładowo, jedna z IZ twierdziła, że podjęta decyzja bazowała na ewaluacji, chociaż była ona dokładnie odwrotna do rekomendacji z badania.

W literaturze podkreśla się, że ewaluacja służy nie tylko dostarczaniu konkretnych propozycji zmian, ale powinna wspierać proces uczenia się organizacji, czyli krytycznego namysłu nad interwencją publiczną i wprowadzania kolejnych udoskonaleń. W praktyce okazuje się, że raporty ewaluacyjne w trzech badanych województwach w niewielkim stopniu wspierają tego

typu procesy i w praktyce rzadko są wykorzystywane w trakcie dyskusji nad interwencją. Wynika to jednak z tego, że ewaluacje w niewielkim stopniu dostarczają nowej wiedzy czy też generują nowe spojrzenie na problem, pobudzają do refleksji, zachęcają do krytycznego spojrzenia na dotychczasowe działania.

Najważniejszym jednak problemem okazuje się, że w opinii przedstawicieli instytucji zarządzających – głównych odbiorców ewaluacji, raporty z badań dostarczają mało istotnych i mało użytecznych wniosków z ewaluacji. Badani nie byli w stanie przywołać żadnych wartościowych wniosków, które byłyby wykorzystywane jako argument w dyskusji<sup>1</sup>.

### Przyczyny słabego wykorzystania ewaluacji

Pojawia się w tej sytuacji pytanie, dlaczego ewaluacje są w tak niewielkim stopniu wykorzystywane do doskonalenia polityk publicznych. Odpowiedzią jest niska użyteczność (jakość) badań ewaluacyjnych. Tutaj jako punkt odniesienia służyć może kolejna publikacja Tomasza Kupca, który przeanalizował 30% badań zrealizowanych na poziomie RPO w latach 2007-2013. Warto w tym miejscu przytoczyć kilka kluczowych wniosków i obserwacji z tej analizy. Kupiec zwrócił uwagę na dwa zasadnicze problemy:

- niespójność rekomendacji z celami badań i niska istotność rekomendacji,
- niedostateczna kompletność raportów z badań.

Odnosząc się do tych dwóch punktów można stwierdzić, że tylko 20% ewaluacji zawiera rekomendacje w pełni zgodne z celami badania, a w 30% badań rekomendacje były całkowicie niezgodne z celami badawczymi. Co prawda 77% rekomendacji ocenianych jest jako spójnych z wnioskami badań (dlatego można je ocenić jako wiarygodne). Jedna ocena ich jakości wypada zdecydowanie słabiej. Zaledwie 37% rekomendacji ocenianych jest jako istotne, tj. powiązane z wnioskami, czytelne i możliwe do wdrożenia, niosące pewien potencjał do wywołania znaczących, pozytywnych zmian, rozwiązania problemu. Tylko 7% rekomendacji ma charakter strategiczny (dotyczy kierunków wsparcia, zakresu działań, typów projektów i alokacji finansowych). Potwierdza to, że rekomendacje często są mało użyteczne.

Drugim, poważnym problemem jest niedostateczna kompletność raportów z badań. W tym przypadku Tomasz Kupiec przeanalizował, w jakim stopniu raporty dostarczają odpowiedzi na pytania badawcze. Okazuje się, że zdecydowana większość, bo aż 72% badań ewaluacyjnych nie zawiera odpowiedzi na wszystkie pytania badawcze. Często są to odpowiedzi pozorowane – luźne dywagacje na dany temat lub opinie badanych – pochodzące z badań ilościowych lub jakościowych – które mogą stanowić cenną informację przy odpowiedzi na pytania badawcze, jednak zwykle nie powinny być traktowane jako bezpośrednia odpowiedź na pytanie badawcze.

1 Kupiec T., *Program evaluation use and its mechanisms: The case of Cohesion Policy in Polish regional administration*, Zarządzanie Publiczne, nr 3 (33), 2015, s. 67-83.

Oczywiście, największe trudności stwarza poszukiwanie odpowiedzi na pytania dotyczące wpływu RPO na rozwój społeczno-gospodarczy, szacowanie oddziaływań interwencji, ocena stopnia zaspokojenia potrzeb społecznych. Paradoksalnie – kwestie te stanowią samo sedno ewaluacji. Jednak w wielu przypadkach ewaluacje dostarczają odpowiedzi pozornych lub całkowicie nietrafnych.

Niska jakość badań ewaluacyjnych wynika z wielu czynników. Część z nich wynika ze specyfiki interwencji. Przykładowo, skala zrealizowanych projektów może być zbyt mała, aby oddziaływanie mogło być uchwycone dostępnymi metodami badawczymi. Inne przyczyny wiążą się z charakterystyką zamawianych badań. Przykładowo, badania mogły być realizowane zbyt wcześnie, aby możliwe było uchwycenie zmiany (choć oczywiście zdarzają się też badania realizowane zbyt późno).

Często okazuje się również, że zaplanowany budżet i wymagane metody badawcze uniemożliwiają uchwycenie oddziaływania. Rzetelne, rygorystyczne zbadanie wpływu interwencji w wielu przypadkach wymaga bardziej zaawansowanych planów badawczych i metod badawczych, w tym przede wszystkim zastosowania badań kontrfaktycznych, z uwzględnieniem pomiarów przed i po interwencji. W praktyce jednak budżet, sposób zaplanowania badania i czas jego realizacji nie pozwalają na zastosowanie tego typu podejść. W związku z tym autorzy raportu formułują wnioski na temat oddziaływania interwencji na podstawie dostępnych informacji, a te często ograniczają się do opinii samych zainteresowanych: projektodawców czy też odbiorców wsparcia. Oczywiście wartość poznawcza tego typu informacji jest raczej niska. Oznacza to, że zamawiający często projektują badania ewaluacyjne zawierające ambitne cele i pytania badawcze, jednak z budżetem i zakresem metod, które pozwalają na pozorowanie odpowiedzi na te ambitne pytania badawcze.

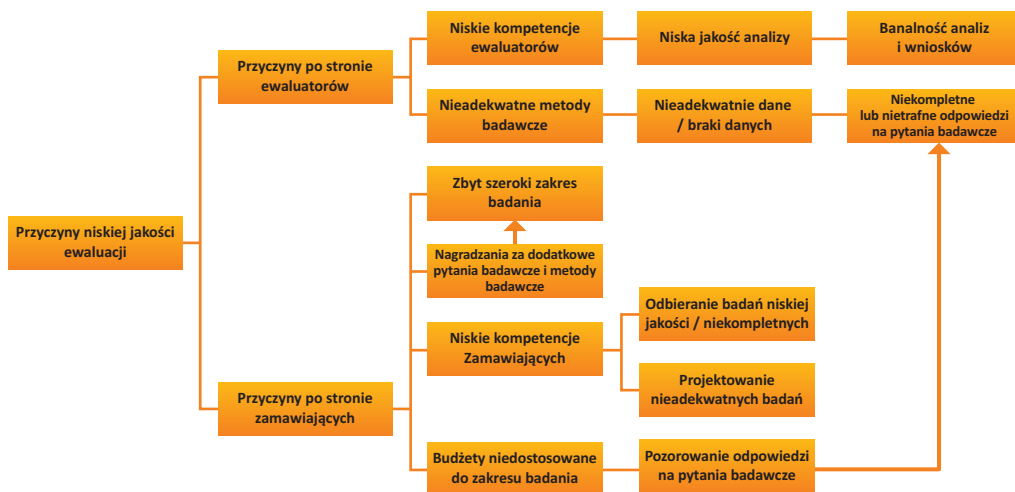
Z powyższym wiąże się jeszcze jeden powszechny błąd przy projektowaniu badań ewaluacyjnych, to jest zbyt duża liczba pytań badawczych. Jak wynika z analizy próby badań – średnio 36 pytań na badanie. Przy czym czasami odpowiedzi są wymagane w dodatkowych przekrojach, a niektóre pytania są bardzo szerokie. W praktyce powoduje to, że ograniczone zasoby są jeszcze bardziej rozproszone i przeznaczone na poszukiwanie odpowiedzi na pytania mało istotne. Co więcej, w niektórych przypadkach formułowane są pytania, które w praktyce nie są rozstrzygalne – ze względu na ich charakter, dostępne metody i budżet badania.

Zjawisko to jest jeszcze pogłębiane przez stosunkowo częstą praktykę premiowania przy wyborze ofert dodatkowych pytań badawczych. Powoduje to, że liczba celów i pytań rośnie znacząco – co oczywiście rzadko przekłada się pozytywnie na jakość badania – częściej – wręcz przeciwnie – powoduje dalsze rozproszenie ograniczonych z natury zasobów<sup>2</sup>.

2 Kupiec T., *Użyteczność ewaluacji jako narzędzia zarządzania regionalnymi programami operacyjnym*, Studia Regionalne i Lokalne Nr 2(56)/2014.

W praktyce okazuje się, że ewaluacja często przypomina grę pozorów: zamawiający pozorują, że robią ewaluację wpływu, a wykonawcy pozorują odpowiedzi na ambitne pytania badawcze. Trudno tu wskazać jednego winnego. Raczej można powiedzieć, że wszyscy aktorzy trochę zawinili, chociaż należy podkreślić, że największy wpływ na jakość badań ewaluacyjnych mają zamawiający, którzy podejmują decyzję kiedy jakie badania będą realizowane, jakie budżety na nie zostaną przeznaczone i jakie będą wymagania wobec wykonawców.

**Rysunek 1. Przyczyny niskiej jakości badań ewaluacyjnych**



Źródło: Opracowanie własne.

**Błędne przekonanie, że ewaluacja służy do rozwiązywania problemów**

Powyżej omówione zostały przyczyny niskiego wykorzystania wniosków i rekomendacji z ewaluacji. Wskazaliśmy na niską jakość i użyteczność formułowanych rekomendacji. Warto jednak zadać sobie pytanie, czy ewaluacja, tak jak jest ona pojmowana szczególnie w obszarze polityki spójności, faktycznie jest w stanie dostarczyć wysokiej jakości rekomendacji dotyczących polityk publicznych. Wydaje się, że pewne fundamentalne cechy ewaluacji znacząco utrudniają osiągnięcie takiego stanu. Wynika to z uproszczającego założenia, że formułowanie trafnych i istotnych rekomendacji jest immanentną cechą ewaluacji. W praktyce jednak mamy tu do czynienia z dwoma oddzielnymi procesami: ewaluacji i rozwiązywania problemów.

**Ewaluacja** służy przede wszystkim sprawdzeniu czy coś działa czy nie działa i dostarczeniu wyjaśnień, dlaczego ewentualnie nie działa (lub działa niedostatecznie dobrze). Krótko mówiąc – dobrze zrobiona ewaluacja powinna nam dostarczyć rzetelnej wiedzy o tym, jak jest (w odniesieniu do pewnego poziomu normatywnego) oraz dlaczego tak jest.

**Rozwiązywanie problemów (wypracowywanie rozwiązań)** jest jednak zupełnie oddzielnym procesem, wymagającym innych kompetencji, narzędzi analitycznych i dodatkowych zasobów (czasu i pracy). Jest to ustrukturyzowany proces analizowania różnych dostępnych alternatyw dla rozwiązania problemu, w trakcie której każdy element problemu / decyzji powinien zostać poddany analizie oddzielnie, systematycznie i w wystarczającym stopniu.

W praktyce ewaluacyjnej istnieje niewypowiedziane głośno założenie, że analiza i ocena polityki publicznej jest wystarczająca, aby sformułować trafne rekomendacje. Tym samym zakłada się, że rekomendacje, czyli propozycje rozwiązania problemu, pojawiają się niejako „przy okazji” poszukiwania odpowiedzi na pytania badawcze. Jest to podejście błędne, powodujące że faktycznie formułowane rekomendacje obciążone są licznymi błędami.

Mechanizmy, które powodują tego typu błędy są świetnie opisane w literaturze dotyczącej psychologii podejmowania decyzji, np.: znakomitej publikacji Daniela Kahnemana. Warto jednak uświadomić sobie, jakie błędy są często popełniane, co prowadzi do formułowania nietrafnych rekomendacji.

- a. Rozpoczynanie analizy problemu od sformułowania wniosków. Przykładowo, na samym początku procesu podejmowana jest decyzja, na czym powinno polegać rozwiązanie danego problemu, a następnie poszukuje się najlepszych rozwiązań, aby je wdrożyć. W praktyce więc proces analityczny ulega odwróceniu: na samym początku nie są rozważane alternatywne sposoby podejścia do problemu.
- b. Proponując rozwiązanie problemu, zwykle następuje koncentracja na jednej alternatywie, która intuicyjnie wydaje się najbardziej właściwa, nie przykładając się odpowiednio do analizy innych, alternatywnych rozwiązań.
- c. Co więcej, rozwiązania które są intuicyjnie preferowane, często wydają się satysfakcjonujące i wystarczające dla rozwiązania danego problemu. W efekcie analiza możliwych rozwiązań nie jest pogłębiania i zatrzymuje się na pierwszych pomysłach przychodzących do głowy, które często nie są tymi najlepszymi.
- d. Często brakuje dostatecznej koncentracji na procesie poszukiwania najlepszego rozwiązania – zamiast tego następuje skupienie się na argumentach, dowodach i wnioskach, wymianianiu się nimi i przywoływaniu ich w trakcie dyskusji. Utrudnia to dojście do właściwych wniosków.
- e. W wielu przypadkach ewaluatorzy nie są wystarczająco przygotowani do strukturyzowania analiz<sup>3</sup>.

W świetle powyższych zagrożeń, warto przyjrzeć się, jak w praktyce tworzone są rekomendacje, w jakich odbywa się to warunkach i na jakich zasadach i w jakim stopniu prawdopodobne jest, że powyższe błędy mogą się pojawić. Obserwacje te bazują na osobistych doświadczeniach i oczywiście są ich pewnym uogólnieniem.

<sup>3</sup> Jones M.D., *The thinker's toolkit*; Crown Business, New York 1995.



Po pierwsze, rekomendacje zwykle są tworzone w ostatniej fazie prac nad raportem końcowym z badania, co samo w sobie nie jest niepokojące, jednak często są one opracowywane w ostatniej chwili i czasu na bardziej ustrukturyzowaną pracę pozostaje stosunkowo niewiele.

Opracowywane są również przez badaczy, którzy proponują rozwiązania zidentyfikowanych przez siebie problemów na podstawie swojej intuicji, wiedzy doświadczenia i przekonań jak należy rozwiązać dany problem. Szybko pojawiają się pierwsze, bazujące na intuicji pomysły i praca koncentruje się na nich. Rzadko zdarza się, aby prowadzona była analiza alternatywnych rozwiązań – brakuje na to czasu i ewaluatorzy nie dostrzegają takiej potrzeby. Co prawda w wielu ewaluacjach w ostatniej fazie badania stosowane są panele / warsztaty, które mogą służyć poprawie trafności rekomendacji (dzięki wymianie opinii, uwzględnieniu dodatkowych racji). W praktyce jednak często służą dyskusji, a nie systematycznej analizie prowadzonej zgodnie z regułami sztuki. Stosunkowo rzadko wykorzystywane są narzędzia analityczne, wspierające proces analizy i podejmowania decyzji. Rzadko na panelach porównywane są alternatywne podejścia do rozwiązania problemu.

W praktyce więc warunki tworzenia rekomendacji sprzyjają pojawieniu się opisanych powyżej błędów, co wpływa na ograniczoną trafność i istotność formułowanych rekomendacji. Dominujące problemy, które zostały uwidocznione:

- brak wydzielenia fazy formułowania rekomendacji,
- zbyt krótki czas na wypracowanie rekomendacji i brak zasobów przeznaczonych na to zadanie w badaniach ewaluacyjnych,
- ograniczone stosowanie technik ustrukturyzowanej analizy problemów.

Powyższym problemom towarzyszy również zbyt wąskie patrzanie na kompetencje ewaluatora. Dominuje postrzeganie ewaluatora przede wszystkim jako badacza. Jest to powszechne zjawisko i nie dotyczy tylko Polski – dyskusje o ewaluacji często koncentrują się na podejściach, metodach i technikach badawczych, służących jak najrzetelniejszej ocenie polityk publicznych. Takie podejście jest widoczne na przykład w pracach Europejskiego Towarzystwa Ewaluacyjnego w opracowanym przez niego ramach kompetencji ewaluatora (*evaluation capabilities framework*).

Wydaje się jednak, że aby zapewnić wysoką jakość formułowanych rekomendacji, niezbędne jest również uwzględnienie dwóch dodatkowych zestawów kompetencji: eksperta i konsultanta. Poniżej krótka charakterystyka każdej z tych kompetencji

- a. Badacz – potrafi zoperacjonalizować problem i pytania badawcze, skonstruować narzędzia, poddać analizie wyniki badania;
- b. Ekspert – ma dużą wiedzę o tematyce badania (rozwoju regionalnym, edukacji itd.);
- c. Konsultant – stara się zrozumieć potrzeby klienta, systematycznie przeanalizować problem i zaproponować najlepsze rozwiązania.



Wśród ewaluatorów dominują kompetencje badawcze. Zwracając uwagę na jakość ewaluacji szczególny nacisk jest kładziony na poprawność metodologiczną badań, dobór próby, wnioskowania itd. Widoczne jest to zarówno wśród zamawiających, jak i wśród samych ewaluatorów. Duży nacisk na kompetencje badawcze jest absolutnie niezbędny i konieczny, lecz dalece niewystarczający.

Potrzeba kompetencji eksperckich jest dostrzegana przez zamawiających i w niektórych ewaluacjach zamawiający starają się zapewnić tego typu kompetencje w zespole. Potrzeba eksperckiej wiedzy wynika z dużego poziomu skomplikowania poszczególnych polityk publicznych, na których efekty wpływa wiele, wzajemnie powiązanych czynników, często o charakterze szczegółowym. I choć przeciętny ewaluator jest w stanie dosyć szybko zrozumieć ogólne założenia danej interwencji, to bez doświadczenia w danej dziedzinie istnieje ryzyko formułowania bardzo ogólnych i mało odkrywczych wniosków i rekomendacji. W praktyce eksperckość przez samych ewaluatorów (ale i zamawiających – co widać w wymaganiach) traktowana jest czasami powierzchownie (brak wystarczającego zrozumienia tematyki) lub akademicko (wymagany tytuł naukowy).

Stosunkowo najrzadziej ewaluator jest postrzegany jako konsultant – czyli osoba, która potrafi ustrukturyzować problem (rozłożyć go na czynniki pierwsze), a następnie poddać systematycznej analizie różne, alternatywne rozwiązania. Kompetencje konsultanta wymagają też uważnego wsłuchania się w potrzeby klienta i ich zrozumienie, tak aby proponowane rozwiązanie w najpełniejszym stopniu odpowiadało na jego potrzeby. W przypadku ewaluacji polityk publicznych oczywiście należy uwzględnić szeroką definicję klientów (uwzględniając perspektywę interesariuszy). Dążąc do wzmocnienia jakości wniosków z badań i formułowanych rekomendacji, niezbędne wydaje się wzmocnianie w zespołach ewaluacyjnych nie tylko kompetencji badawczych, ale również eksperta i konsultanta.

## Podsumowanie

Powyżej przedstawiono syntetyczną próbę ujęcia problemu jakości ewaluacji i jej użyteczności. Jak wykazano, wnioski i rekomendacje z ewaluacji mają często charakter techniczny, a rzadko dotyczą kwestii ważnych z punktu widzenia polityk publicznych. Często też są formułowane na dużym poziomie ogólności i w sumie są banalne. Prowadzi to jednak do upowszechniania się ewaluacji jako działania pozornego – prowadzenia procesów które wydają się zgodne z zasadami prowadzenia ewaluacji, ale w praktyce nie mającymi dużego wpływu na rzeczywistość podejmowane decyzje. Ta pozorność dotyczy zarówno zamawiających, którzy zlecają ewaluacje o ambitnych celach, często z bardzo rozbudowaną listą pytań, które jednak nie mogą być osiągnięte przy proponowanych sposobach realizacji, budżetach i czasie na jego zrealizowanie. Dotyczy to również wykonawców, którzy nie będąc w stanie osiągnąć celów formułowanych przez zamawiających i odpowiedzieć rzetelnie na pytania, dostarczają badań pozorujących odpowiedzi na wszystkie pytania badawcze.

Takie mechanizmy prowadzą do psucia się rynku – pojawia się coraz więcej firm, które doskonale zrozumiały ten mechanizm i dostarczają niedrogich badań niskiej jakości, pozorujących odpowiedzi na pytania badawcze. Prowadzi to do zaostrzenia się konkurencji cenowej, która zmusza firmy z ambicjami do ścinania kosztów i obniżania jakości.

Pojawia się pytanie, czy możliwe jest ograniczenie tych negatywnych zjawisk. Oczywiście, wymaga to jednak wprowadzenia szeregu zmian do praktyki ewaluacji i dużego zaangażowania zamawiających, którzy faktycznie regulują ten rynek. Poniżej znajduje się lista kierunkowych propozycji działania:

- Większy nacisk na kompetencje ewaluatorów:  
Zaostrzenie wymagań wobec firm badawczych/ ewaluatorów, z jasnym wyodrębnieniem badaczy na pozycji senior i na pozycji junior.  
Określanie wymagań jakościowych i ich konsekwentne stosowanie na etapie odbioru raportów z badań.
- Ograniczenie liczby pytań badawczych:  
Koncentracja na kluczowych kwestiach badawczych.  
Wstępna weryfikacja pytań pod kątem ich rozstrzygalności.  
Rezygnacja z kryteriów premiujących dodatkowe pytania badawcze.
- Wymuszenie lepszego rozpoznania tematu i strukturyzacji problemów badawczych:  
Wydłużenie fazy wstępnej badania.  
Zastąpienie raportu metodologicznego raportem wstępnym.  
W raporcie wstępnym powinny znaleźć się pierwsze ustalenia badawcze oraz analiza teorii zmiany interwencji.  
Częstsze stosowanie raportów cząstkowych, podsumowujących kolejne etapy procesu ewaluacji.
- Wyodrębnienie fazy rozwiązywania problemów zdiagnozowanych na etapie badań:  
Wyodrębnienie w procesie ewaluacji fazy przygotowania rekomendacji.  
Oczekiwanie w ofertach prezentacji metod analizy problemów i sposobu dochodzenia do rekomendacji.
- Urealnienie budżetów badań ewaluacyjnych, ze szczególnym naciskiem na koszty: przygotowania koncepcji badania, analizy materiału i przygotowania rzetelnego raportu.

## Bibliografia

European Evaluation Society, *The EES Evaluation Capabilities Framework* <http://www.europeanevaluation.org/sites/default/files/EES%20EVALUATION%20CAPABILITIES%20FRAMEWORK.pdf>

**Jones M.D.**, *The thinker's toolkit*; Crown Business, New York 1995

**Kupiec T.**, *Program evaluation use and its mechanisms: The case of Cohesion Policy in Polish regional administration*, *Zarządzanie Publiczne*, nr 3 (33), 2015, s. 67-83

**Kupiec T.**, *Co możemy zrobić by zwiększyć poziom wykorzystania ewaluacji? – wnioski z badania instytucji zarządzających regionalnymi programami operacyjnymi 2007-2013.*; [w:] *Wykorzystanie wiedzy przez administrację – mechanizmy nadzoru obywatelskiego*; Fundacja Idea Rozwoju, Warszawa 2015

**Kupiec T.**, *Użyteczność ewaluacji jako narzędzia zarządzania regionalnymi programami operacyjnym*, *Studia Regionalne i Lokalne* Nr 2(56)/2014



### PIOTR STRONKOWSKI

Piotr Stronkowski jest niezależnym ewaluatorem, zajmującym się głównie tematyką rynku pracy, edukacji, integracji społecznej i trzeciego sektora. Posiada szerokie, wieloletnie doświadczenie badawcze, zdobyte w Coffey International Development (Senior Manager Evaluation and Research) i PSDB (dyrektor ds. badań). Przez wiele lat pracował również w administracji publicznej. W latach 2004-2008 zajmował stanowisko Zastępcy Dyrektora Departamentu Zarządzania Europejskim Funduszem Społecznym w Ministerstwie Pracy i Polityki Społecznej oraz Ministerstwie Rozwoju Regionalnego. Jest absolwentem Polityki Społecznej na Uniwersytecie Warszawskim oraz autorem licznych publikacji. Jest również współpracownikiem i ekspertem licznych instytucji, takich jak: Collegium Civitas, UNDP, OECD, KSAP, MPiPS, MRR, Fundacja Bałtockiego itd.



# DIAGNOZA PROBLEMÓW I ŚRODKÓW ZARADCZYCH W PERSPEKTYWIE 2014-2020 NA PRZYKŁADZIE DOŚWIADCZEŃ MAŁOPOLSKIEJ JEDNOSTKI EWALUACYJNEJ

## AGNIESZKA ZAWRZYKRAJ I BARBARA PORĘBSKA-MACIOŁEK

### Wstęp

Ewaluacja jest wokół nas, towarzyszy nam na każdym etapie życia. W życiu codziennym zbieramy dane, analizujemy je, wyciągamy wnioski i następnie podejmujemy na ich podstawie decyzje. Czy są słuszne? Zwykle jesteśmy w stanie to ocenić z perspektywy upływającego czasu. Oczywiście nie zawsze łatwo jest podjąć decyzje, szczególnie wtedy, gdy musimy najpierw przyznać się do popełnionego błędu. Jednak ważne jest to, że dążymy do poprawy i jesteśmy otwarci na czasem znaczące zmiany.

Rozważania nad samą ewaluacją przyniosły następujące refleksje:

- **E**waluacja to edukacja - dzięki ewaluacji uczymy się o badanym zjawisku;
- **e**Waluacja to wiedza - dzięki ewaluacji więcej wiemy;
- **ew**Aluacja to aktualność - dzięki ewaluacji pozyskujemy aktualne informacje;
- **ewa**Luacja to ludzie - ewaluację robią ludzie dla ludzi;
- **ewal**Uacja to ufność - w ewaluacji liczy się wzajemnie zaufanie w relacji Wykonawca – Respondent, Wykonawca – Zamawiający – Odbiorca ostateczny;
- **ewalu**Acja to analiza - ewaluacja opiera się na analizie problemu, zagadnienia;
- **ewalua**Cja to cel - ewaluacja jest prowadzona w określonym, jasno zdefiniowanym celu;
- **ewaluac**Ja to jakość - jakość powinna być jednym z przymiotów rzetelnie prowadzonej ewaluacji;
- **ewaluacj**A to ankieta - podstawowym narzędziem ewaluacyjnym jest ankieta.

Zamysłem niniejszego opracowania jest pokazanie głównych problemów oraz kierunków działań naprawczych, jakie należy podjąć w zakresie ewaluacji w perspektywie finansowej 2014-2020. Większość z opisanych problemów nie należy do nowych, stanowią pewną kontynuację wyzwań powstałych jeszcze w okresie wdrażania funduszy unijnych w latach 2007-2013. Co warto zaznaczyć, zdiagnozowane zostały na bazie kilkuletnich doświadczeń w realizacji badań ewaluacyjnych, z perspektywy zamawiającego.

Należy również nadmienić, że kwestie poruszane w niniejszym opracowaniu nie odnoszą się do konkretnej jednostki ewaluacyjnej, z wyłączeniem tych treści, gdzie została wprost wskazana Małopolska Jednostka Ewaluacyjna.

Małopolska Jednostka Ewaluacyjna, oficjalnie nosząca nazwę Jednostka Ewaluacyjna Instytucji Zarządzającej Regionalnym Programem Operacyjnym Województwa Małopolskiego na lata 2014-2020, powstała pierwszego stycznia 2015 roku jako odpowiedź na zapotrzebowanie

na utworzenie wspólnej jednostki ewaluacyjnej dla działań finansowanych zarówno z EFS, jak i EFRR, w ramach jednego programu operacyjnego, o którym mowa wyżej. Małopolska Jednostka Ewaluacyjna mieści się w Departamencie Zarządzania Programami Operacyjnymi Urzędu Marszałkowskiego Województwa Małopolskiego, w Zespole ds. Monitoringu, Sprawozdawczości i Ewaluacji. Obecnie w jej skład wchodzi trzy stanowiska ds. ewaluacji. Osoby zajmujące się ewaluacją, posiadają doświadczenie w prowadzeniu ewaluacji w perspektywie 2007-2013 oraz wiedzę zdobytą na licznych szkoleniach, warsztatach i studiach podyplomowych z zakresu ewaluacji.

W toku prac nad Planem Ewaluacji Regionalnego Programu Operacyjnego Województwa Małopolskiego na lata 2014-2020, Małopolska Jednostka Ewaluacyjna zdiagnozowała szereg problemów w zakresie ewaluacji, z którymi należy zmierzyć się w obecnej perspektywie finansowej, a mianowicie:

### 1. Lepsza jakość ewaluacji

Problem, jaki został zdiagnozowany jeszcze w perspektywie finansowej 2007-2013, to niska jakość badań. Na rynku polskim powstaje coraz więcej firm realizujących badania, w tym badania ewaluacyjne, jako odpowiedź na zapotrzebowanie na tego typu usługi. Jednak firmy te zwykle mają niewielkie doświadczenie w realizacji badań ewaluacyjnych, i w związku z tym spotkać się można z niezrozumieniem potrzeb zamawiającego. Wykonawcy ewaluacji często nie znają specyfiki funkcjonowania danej instytucji, projektu, czy programu. Stan wiedzy prezentują już w ofertach przetargowych, zawierając w nich minimum z minimum. Przejawia się to w ujmowaniu w ofertach tylko takich zapisów, które są niezbędne z punktu widzenia spełnienia wymogów stawianych przez zamawiającego.

To, że niewielkie doświadczenie przekłada się na realizację badania, jest sprawą oczywistą. Ponadto równie oczywiste jest to, że wiarygodność wyników takiego badania może budzić duże zastrzeżenia, i co za tym idzie - może wywoływać duży opór wobec rekomendacji, a jeszcze większy wobec ich wdrażania.

Wykonawcy składają oferty w kilku przetargach jednocześnie. *De facto* realizują dwa lub więcej (sic!) badań w tym samym czasie, i finalnie nie realizują ich na odpowiednim poziomie. Przypadek ten oczywiście nie dotyczy dużych firm, mających kilka zespołów badawczych. Jednak takie należą do wyjątków.

Niewielkie doświadczenie oraz nakładanie się kilku badań, znacznie wpływa na jakość raportów. Ich niska jakość generuje duże zaangażowanie czasowe i kadrowe zamawiającego. Nierzadko sytuacja taka sprowadza się do kopiowania uwag i komentarzy zamawiającego, i wprowadzania ich treści do raportów jako odpowiedź na zgłoszoną reklamację. Niestety wykonawcy wykorzystują fakt, że termin przekazania raportu jest zdeterminowany czasowymi ograniczeniami wykorzystania wyników badania, które rzadko kiedy mają uniwersalny, ponadczasowy charakter. Są raczej operacyjne, reagujące na bieżące potrzeby zamawiającego.

Jednak już dzisiaj zamawiający dysponuje pewnymi środkami, które może wdrożyć celem poprawy jakości realizowanych badań. Należą do nich m.in.:

- a) **podniesienie wymagań wobec wykonawców w zakresie posiadanego doświadczenia.** Nie wystarczy już wykazanie się doświadczeniem w realizacji badań w zakresie szeroko pojętych funduszy unijnych. Ich tematyka powinna być skonkretyzowana i adekwatna do zakresu zamówienia.
  - b) **częste spotkania i kontakty z wykonawcą w trakcie realizacji badania.** Obecnie w ewaluacji coraz częściej stosuje się podejście partycypacyjne (*ang. participatory approach*), opierające się na współpracy wykonawcy i zamawiającego w toku badania.
  - c) **raportowanie z przebiegu prac.** Forma raportowania może być różnaita, w zależności od preferencji zamawiającego. W przypadku Małopolskiej Jednostki Ewaluacyjnej wykorzystuje ona formę cotygodniowych, codwutygodniowych sprawozdań. W ramach sprawozdań wykonawca przedstawia czynności, jakie zostały przez niego podjęte w danym okresie sprawozdawczym, m.in. zestawienie zrealizowanych wywiadów, oraz działania, jakie chce podjąć w dalszej kolejności. Jest to również narzędzie pomagające wykonawcy w uporządkowywaniu prac badawczych. Sprawozdania są także dobrym narzędziem do sygnalizowania wszelkich pojawiających się w toku badania problemów. Daje to możliwość sprawnego zareagowania i podjęcia odpowiednich środków zaradczych. Małopolska Jednostka Ewaluacyjna pracuje nad wystandaryzowaniem tego narzędzia.
  - d) **właściwe zaplanowanie badania w czasie.** Badania muszą zostać zaplanowane w odpowiednim momencie, tak, by wyniki i rekomendacje był możliwe do wykorzystania w celu zareagowania na bieżące potrzeby i wdrożenia. Nierzadko zdarza się tak, że z powodu późnej realizacji badania, wyniki mogą zostać niewykorzystane (problem, potrzeba informacyjna musiały zostać skonkretyzowane i rozwiązane czy też spełnione wcześniej). Pomocne w tym zakresie mogą być opracowywane przez poszczególne jednostki ewaluacyjne Plany Ewaluacji, zawierające listę planowanych do realizacji badań w perspektywie finansowej 2014-2020. Dzięki ich wcześniejszemu zdiagnozowaniu powstaje możliwość sprawnego uskuteczniania wdrażanego programu, odpowiadania na realne potrzeby, jak i wypełniania zobowiązań wynikających z wymogów Komisji Europejskiej.
- Co więcej, badania muszą być rozplanowane tak, by czas przeznaczony na ich realizację był wystarczający z punktu widzenia potrzeb zamawiającego i wykonawcy, a także uwarunkowań zewnętrznych umożliwiających zebranie danych dotyczących grup docelowych. Służyć temu może opracowanie i przyjęcie racjonalnego harmonogramu badania. Jego racjonalność powinna przejawiać się w co najmniej kilku kwestiach:
- przeznaczenie odpowiedniego czasu na przygotowanie przez wykonawcę raportów częściowego, metodologicznego i końcowego,
  - odpowiedni czas na ocenę ich jakości przez zamawiającego,



- optymalny czas na opracowanie narzędzi badawczych,
  - właściwe rozplanowanie w czasie wywiadów terenowych (bynajmniej nie służy temu odgórne narzucanie przez zamawiającego sztywnych terminów ich realizacji).
- e) **większy nacisk Zamawiającego na wysoką jakość raportów końcowych.** Praktyką powinno być przyjmowanie raportów o wysokiej jakości. W tym celu zamawiający powinien wcielić w życie wyżej wymienione środki poprawiające jakość realizowanych badań, co więcej zmienić nastawienie właściwie do całego procesu ewaluacji - nie traktować ewaluacji jako „zło konieczne”, a narzędzie do uskuteczniania realizowanych procesów, wdrażanych polityk.
- f) **wyższe kary za niewłaściwą realizację zapisów umowy i ich egzekwowanie.** Każda umowa zawiera zapisy chroniące zamawiającego przed nierzetelnym wykonawcą. Jednak często są one niewykorzystywane w sytuacji jednoznacznie wskazującej na wystąpienie nieprawidłowości w realizowanej umowie. Zapisy te powinny być egzekwowane, w szczególności w sytuacjach otrzymywania raportów niskiej jakości. Rozpatrując jednak kwestie kar jako elementu egzekwowania, przede wszystkim należy mieć na uwadze dążenie do zakończenia badania z sukcesem, w czym ma pomóc m.in. stosowanie podejścia partycypacyjnego w realizacji badania.

## 2. Ewaluacja zewnętrzna vs (?) ewaluacja wewnętrzna

Reasumując powyższe wywody na temat ewaluacji realizowanej przez firmy badawcze, jej jakość coraz częściej rozczarowuje. W poszukiwaniu alternatywnych rozwiązań warto zastrzymać się przy, dotychczas w niewielkim stopniu wykorzystywanej, ewaluacji wewnętrznej. Zgodnie z definicją zawartą w *Wytycznych w zakresie ewaluacji polityki spójności na lata 2014-2020* (dalej: *Wytyczne*) ewaluacja wewnętrzna to *ocena przeprowadzana przez ewaluatora funkcjonującego w ramach instytucji odpowiedzialnej za realizację interwencji publicznej np. programu, z wykorzystaniem zasobów wewnętrznych*<sup>1</sup>.

Jednak prowadzenie ewaluacji wewnętrznej zdeterminowane jest kilkoma czynnikami, do których należą:

- a) odpowiednie zasoby kadrowe,
- b) wystarczająca wiedza merytoryczna, jaką powinni posiadać pracownicy jednostki ewaluacyjnej, by taką ewaluację wykonać,
- c) odpowiednie zaplecze techniczne m.in. w postaci specjalistycznego oprogramowania, dostępu do baz telefonicznych, etc., gdyż cel badania niejednokrotnie generuje konieczność prowadzenia badań terenowych. Pomocna w tym zakresie może okazać się ewaluacja mieszana, w obecnej perspektywie finansowej zwana hybrydową. Zgodnie z wyżej przytoczo-

1 Wytyczne w zakresie ewaluacji polityki spójności na lata 2014-2020 [https://www.funduszeuropejskie.gov.pl/media/9318/wytyczne\\_090915\\_final.pdf](https://www.funduszeuropejskie.gov.pl/media/9318/wytyczne_090915_final.pdf) (28.09.2016).

nymi *Wytycznymi* jest to ewaluacja, w ramach której działania badawcze w ramach jednego badania wykonywane są zarówno przez ewaluatora wewnętrznego, jak i zewnętrznego. Realizacja badań terenowych spoczywa wtedy na, wyłonionych w drodze przetargu, firmach zewnętrznych, które dostarczają materiał niezbędny do analiz, prowadzonych przez pracowników jednostki ewaluacyjnej. Nadal jednak istnieje bariera w postaci braku posiadania specjalistycznego oprogramowania koniecznego do przeanalizowania zebranych danych, którego zakup jest często zbyt kosztowny.

Niemniej jednak wartym podkreślenia jest to, że dzięki prowadzeniu ewaluacji wewnętrznych jednostka ewaluacyjna nabywa wiedzę z zakresu konkretnych technik niezbędnych do przeprowadzenia badań, konstruowania narzędzi badawczych, opracowywania odpowiedniej metodologii, itp. W rezultacie tworzone są podwaliny brokeringu wiedzy<sup>2</sup>.

Mając na uwadze powyższe, należy dążyć do co najmniej:

- a) wzrostu zatrudnienia osób posiadających również wiedzę z zakresu prowadzenia analiz, badań terenowych, etc. lub/i szkolenia osób aktualnie zajmujących się ewaluacją wdrażając narzędzia o charakterze warsztatowym, czy nawet samych staży u wykonawców w celu wymiany i nabycia nowych kompetencji,
- b) właściwego wyposażenia stanowiska pracy.

### 3. Stosowanie metod kontrfaktycznych w ewaluacji

Patrząc na doświadczenia z wdrażania poprzedniej perspektywy finansowej oraz w związku z koniecznością oszczędniejszego wydatkowania środków publicznych (coraz mniej środków na finansowanie wsparcia), pojawiają się obawy o skuteczność wykorzystania funduszy strukturalnych UE. W obecnej perspektywie finansowej Komisja Europejska (KE) zachęca więc państwa członkowskie do zwiększenia wysiłków w celu dostarczania wiarygodnych dowodów na temat efektów realizacji programów z funduszy UE przy uwzględnieniu efektów jakie mogłyby zostać osiągnięte w przypadku braku wsparcia UE. Takich dowodów dostarczy nam ewaluacja wpływu oparta na sytuacji kontrfaktycznej tj. porównanie osiągniętych rezultatów z szacowanymi rezultatami, które mogłyby się pojawić pomimo braku interwencji<sup>3</sup>. Ta krótka charakterystyka ewaluacji z zastosowaniem metod kontrfaktycznych, tak mocno promowanych przez Komisję Europejską w obecnej perspektywie finansowej, wydaje się prosta i zrozumiała. Jednakże w praktyce pojawia się wiele problemów tak z ich zrozumieniem, ale przede wszystkim z ich zastosowaniem. Główną barierą

2 Szerzej o Brokeringu wiedzy: Olejniczak K., Kupiec T., Raimondo E., *Brokerzy wiedzy. Nowe spojrzenie na rolę jednostek ewaluacyjnych*, [w:] (Re)ewaluacja 2. Wiedza w działaniu. Red. A. HABER, K. OLEJNICZAK. PARP, Warszawa 2014, s. 68.

3 Projektowanie i zlecenie ewaluacji wpływu opartych na sytuacji kontrfaktycznej. Praktyczne wytyczne dla Instytucji Zarządzających EFS. [http://www.archiwum.ewaluacja.gov.pl/Dokumenty\\_ewaluacyjne/Documents/Przewodnik\\_KE\\_po\\_metodach\\_kontrfaktycznych\\_PL.pdf](http://www.archiwum.ewaluacja.gov.pl/Dokumenty_ewaluacyjne/Documents/Przewodnik_KE_po_metodach_kontrfaktycznych_PL.pdf) (28.09.2016).

stosowania metod kontrfaktycznych jest po prostu niewielka wiedza nt. metod kontrfaktycznych zamawiających oraz małe doświadczenie wykonawców w realizacji badań z ich wykorzystaniem.

Pierwszy problem pojawia się już etapie planowania oraz projektowania założeń do badania ewaluacyjnego, tj. próba odpowiedzi na pytanie „W jakich badaniach metody kontrfaktyczne stosować?”, i w konsekwencji, czy nasze badanie należy do takich zaliczyć. Odpowiedź na to pytanie rodzi następne, a mianowicie, jakie w danym badaniu zastosować podejście do ewaluacji wpływu, czy porównanie „przed i po”, czy zaś wystarczy porównanie beneficjentów i podmiotów niekorzystających ze wsparcia. To określenie odpowiedniego podejścia jest kluczowe w celu zdiagnozowania grup kontrolnych, które będziemy porównywać, aby móc odpowiedzieć na pytanie o skuteczność podejmowanych przez nas interwencji<sup>4</sup>. Określając grupy kontrolne należy też diagnozować, z jakich źródeł pozyskamy dane na ich temat. Wiąże się z tym kolejny problem dotyczący m.in. dostępności danych (m.in. problem z udostępnieniem danych przez ZUS), nieaktualności publikowanych danych, np. przez GUS, czy też niekompletnych systemów gromadzenia danych takich jak SL, ale przede wszystkim regionalny system (w przypadku Małopolski jest to e-RPO), z którego dane „zasysane” do SL nadal nie są kompletne, a dostęp do możliwości generowania raportów mają właściwie tylko uprawnieni administratorzy. Aby odpowiednio wcześniej zabezpieczyć dostęp do danych KE zaleciła właściwym Instytucjom Zarządzającym sporządzenie Planów Ewaluacji dla poszczególnych programów operacyjnych, które mają zawierać wszystkie możliwe do zidentyfikowania badania ewaluacyjne planowane do realizacji w trakcie trwania programu. Zawarte w Planach Ewaluacji opisy poszczególnych badań zawierają m.in. główne pytania badawcze, obszary problemowe oraz ogólny zarys metodologii wraz z zakresem niezbędnych danych i informacją skąd zostaną pozyskane.

Wskazane powyżej trudności powodują, że zamawiający badanie ewaluacyjne, oparte o metody kontrfaktyczne, mają trudności w przygotowaniu dobrego Opisu Przedmiotu Zamówienia (OPZ), a następnie zrealizowaniu badania, które oszacuje stopień oddziaływania programu. Pomysłem na rozwiązanie problemu z OPZ jest próba „przerzucenia” przygotowania szczegółowych założeń badawczych na wykonawców. Pomocnym tutaj może być przeprowadzenie dialogu technicznego, w trakcie którego zamawiający zwraca się do ekspertów lub wykonawców o doradztwo w zakresie przygotowania OPZ oraz specyfikacji istotnych warunków zamówienia (m.in. poprzez dialog techniczny Małopolska Jednostka Ewaluacyjna ma możliwość wykorzystania wiedzy ze świata nauki i potencjał ośrodków naukowych, jakimi dysponuje Kraków).

Kolejne pytanie, jakie się pojawia w przypadku stosowania metod kontrfaktycznych brzmi: *Jak wybrać dobrą ofertę?* Trudnością, z jaką można wtedy się zetknąć, jest brak do-

4 Martini A., *Wprowadzenie do podstawowych pytań stawianych w ewaluacji oddziaływania interwencji: Czy interwencja coś zmieniła? A jeśli tak, to dla kogo?*, [w:] Ocena wpływu i prognozowanie efektów w badaniach ewaluacyjnych. Red. Haber A., Trzciniński R., PARP, Warszawa 2011, s. 23.

świadczenia i niewystarczająca wiedza zamawiających na temat zastosowania metod kontrfaktycznych, co może znacznie utrudnić dokonywanie ocen adekwatności i trafności zaproponowanych przez wykonawców konkretnych metod kontrfaktycznych w ramach złożonych ofert. Dlatego warto poszerzać swoją wiedzę m.in. poprzez uczestnictwo w specjalistycznych warsztatach, szkoleniach dotyczących metod kontrfaktycznych, również „szytych na miarę” potrzeb osób zlecających tego typu badania. Godnym polecenia jest pomysł stworzenia szkoły PONAD (Polityki Publiczne Oparte NA Dowodach), wspólnej inicjatywy Centrum Ewaluacji i Analiz Polityk Publicznych (CEAPP) oraz Fundacji IDEA Rozwoju pod patronatem Polskiego Towarzystwa Ewaluacyjnego, która w ramach praktycznych zajęć ma przygotować osoby zamawiające (i nie tylko) do realizacji ewaluacji z wykorzystaniem analiz kontrfaktycznych.

Przed wszystkim jednak najlepiej uczy się poprzez zdobyte doświadczenie a zdobędziemy je tylko realizując tego typu badania.

#### **4. Mniej szczegółowy Opis Przedmiotu Zamówienia**

Kolejny zdiagnozowany problem to zbyt szczegółowe OPZ, tendencja do zawierania w nich „sztampowych”, najbardziej popularnych metod i technik badawczych. Jak istotnym elementem w procesie zlecenia badań ewaluacyjnych jest należyte sformułowanie opisu przedmiotu zamówienia i jak wiele trudności sprawia jego przygotowanie, wie każdy zamawiający. Przyjętą praktyką jest aby OPZ były bardzo szczegółowe, gdyż dzięki temu wykonawca będzie w stanie zidentyfikować to, czego zamawiający od niego oczekuje. Co więcej będzie miał gwarancje, że otrzymamy produkt lub usługę spełnił jego oczekiwania. O ile praktyka ta sprawdza się w przypadku zakupu produktów czy robót budowlanych, o tyle w badaniach ewaluacyjnych może się okazać przeszkodą. Patrząc na dotychczasowe doświadczenie, można odnieść wrażenie, że zbyt szczegółowe OPZ wcale nie gwarantują otrzymania wyników realizowanych badań na jakościowo odpowiednim poziomie. Można zaobserwować, że często dobrze przygotowane, zawierające maksimum informacji OPZ skłaniają firmy do składania „ubogich”, ograniczających się do skopiowania treści OPZ, oferty. Często są to „nieprofesjonalne” firmy lub z małym doświadczeniem, które nie mają wiele do zaoferowania, a konsekwencją tego jest trudna wzajemna współpraca. Ta trudna współpraca jest głównym powodem utraty zaufania do potencjalnych wykonawców, co bezpośrednio przekłada się na zwyczaj opracowywania bardzo szczegółowych OPZ, w których pozostawiamy niewiele miejsca na inwencję wykonawcy i możliwość wyjścia „poza schemat”. Broniąc niejako schematycznych propozycji zapisywanych w SOPZ, wynikają one głównie z ograniczonej specjalistycznej wiedzy zamawiających w zakresie niestandardowych technik badawczych, którą to wiedzą powinni jednak dysponować wykonawcy. Zastosowanie niestandardowych technik, indywidualnych rozwiązań zaproponowanych przez wykonawcę na pewno podniesie jakość badań. Innymi słowy wykorzystanie mniej znanych, nowych rozwiązań czy technik badawczych, może przynieść pozytywne efekty.

Natomiast stosowanie szczegółowych zapisów wraz z popularnymi i powszechnie stosowanymi technikami badawczymi jest przyczyną zgłaszania się do przetargów firm bez dobrego zaplecza badawczego, które próbują konkurować poprzez niższe ceny. A jak powszechnie wiadomo produkt tani to niekoniecznie produkt dobrej jakości. Powstaje pytanie Jak sobie z tym radzić? Warto zacząć od budowania wzajemnego zaufania. **Dajmy szansę wykonawcy** na zgłaszanie swoich, niestandardowych rozwiązań, zważywszy, że ustawa o prawie zamówień publicznych daje nam możliwość stosowania pozacenowych kryteriów wyboru ofert. Odrębną kwestią jest oczywiście jak te pozacenowe kryteria formułować, aby były mierzalne i dawały nam szansę na wybór najciekawszych ofert. Niemniej jednak warto się nad nimi pochylić i sformułować je tak, aby dawały szansę na „wykazanie” się potencjalnemu wykonawcy. To on będzie odpowiedzialny za realizację badania transferując swoją wiedzę, umiejętności i zaplecze.

## 5. Ewaluacja to nie wytykanie błędów – zmiana nastawienia do ewaluacji

Równie ważnym problemem, choć o nieco odmiennym charakterze, niż opisywane dotychczas, jest brak przekonania decydentów, ostatecznych odbiorców ewaluacji co do zasadności i celowości jej stosowania.

Ewaluacja jest traktowana jako krytyka podejmowanych działań, wytykanie błędów systemu zarządzania i wdrażania funduszy unijnych, a jej realizacja jest obowiązkiem narzuconym przez Unię Europejską (ewaluacja to „zło konieczne”). W związku z tym często napotkać można opór wobec wdrażania rekomendacji, szczególnie tych, które faktycznie reagują na zdiagnozowane w toku badania uchybienia. Jednak jest to realizowane w jasno określonym celu, by usprawnić, ulepszyć. Zgodnie z *Wytycznymi ewaluacja to badanie społeczno-ekonomiczne, którego celem jest oszacowanie w odniesieniu do jasno sformułowanych kryteriów (najczęściej skuteczności, efektywności, użyteczności, trafności i trwałości) jakości oraz efektów wdrażania interwencji publicznych*<sup>5</sup>. Spotkać się również można z traktowaniem ewaluacji jako kontroli. Nic bardziej mylnego. Kontrola to *ogół działań (wszelkie kontrole dokumentacji, kontrole na miejscu) mających na celu uzyskanie racjonalnego zapewnienia, że projekty dofinansowywane z funduszy unijnych realizowane są prawidłowo, a w konsekwencji wydatki poświadczane do KE są wiarygodne i wolne od nieprawidłowości*<sup>6</sup>. Ewaluacja jest zatem nastawiona na badanie przede wszystkim jakości realizowanych przedsięwzięć, kontrola – na diagnozowaniu nieprawidłowości funkcjonowania systemu.

5 Wytyczne w zakresie ewaluacji polityki spójności na lata 2014-2020. [https://www.funduszeuropejskie.gov.pl/media/9318/wytyczne\\_090915\\_final.pdf](https://www.funduszeuropejskie.gov.pl/media/9318/wytyczne_090915_final.pdf) (28.09.2016).

6 Szymański M., *Kontrola projektów w ramach polityki spójności oraz propozycje na okres programowania 2014–2020*. [w:] *Kontrola państwowa*. Numer 3(344), maj-czerwiec 2012, s. 11.

Jednostka ewaluacyjna powinna dążyć do:

- zmiany postrzegania przez decydentów, ostatecznych odbiorców ewaluacji i jej przeznaczenia poprzez odpowiednie ich edukowanie (częste spotkania, przekazywanie związanych informacji z wynikami badania, tworzenie grup roboczych z udziałem decydentów na potrzeby konkretnej ewaluacji),
- uświadamiania, iż celem ewaluacji jest niwelowanie błędów systemu, by był bardziej skuteczny, wydajny i efektywny, a sama ewaluacja może pomóc w kreowaniu bardziej racjonalnej polityki regionalnej (ewaluacja gwarantuje racjonalne uzasadnienie realizowanej polityki, weryfikuje zakładane cele z rzeczywistymi osiągnięciami w celu poprawy efektywności i skuteczności tej polityki).

## 6. Przetargi powyżej progów unijnych

Na pytanie, czy wybór wykonawcy badania ewaluacyjnego w procedurze przetargu nieograniczonego powyżej progów unijnych będzie stanowił problem, jak na razie ciężko jest odpowiedzieć. Wynika to z braku doświadczenia w tym zakresie, ponieważ Małopolska Jednostka Ewaluacyjna dotychczas nie przeprowadzała tego typu procedury. Niepokoi jednak fakt, że w obecnym podejściu do ewaluacji pojawiła się tendencja do łączenia badań ewaluacyjnych w ramach planów zamówień publicznych tworzonych dla projektów Pomocy Technicznej, co rodzi ryzyko „wejścia” w przetargi powyżej progów unijnych. Skąd obawy przed realizacją tego typu przetargów? Główna obawa to czas realizacji takiego zamówienia, gdyż termin składania ofert standardowo w przetargu nieograniczonym to minimum 40 dni (poniżej progów to 7 lub 14 dni), natomiast termin związania ofertą to dni 60 (w postępowaniach poniżej progów termin ten wynosi 30 dni). Terminy te niosą ze sobą spore ryzyko znacznego wydłużenia procedury przetargowej, a co za tym idzie - zagrożenia terminowej realizacji badań. W obecnej perspektywie finansowej każda z jednostek ewaluacyjnych ma zaplanowanych od kilkunastu do kilkudziesięciu badań ewaluacyjnych, a mowa tu tylko o tych zaplanowanych w ramach Planów Ewaluacji, pozostają jeszcze te nieprzewidziane i niezaplanowane. Przy takiej ilości badań czas zaczyna odgrywać znaczącą rolę i lepiej dążyć do uproszczenia i skrócenia procedur niżeli do ich wydłużania. W tej sytuacji należy stosować racjonalne podejście do łączenia badań ewaluacyjnych w planach zamówień publicznych, tj. traktować poszczególne badania ewaluacyjne jako odrębne byty, gdyż każde z nich jest specyficzne i nie sposób wrzucać ich do jednego worka z napisem „Ewaluacja”. Należy cały czas uświadamiać i przekonywać decydentów, osoby odpowiedzialne za tworzenie projektów PT oraz realizację planów zamówień o wyjątkowości i specyficznym charakterze ewaluacji.

Ta krótka charakterystyka głównych problemów zdiagnozowanych na etapie realizacji badań ewaluacyjnych oczywiście nie wyczerpuje tematu wszystkich trudności, z jakimi zlecający badania borykają się w swojej codziennej pracy. Jednakże należą one do tych najczęstszych.

Ważne jest aby traktować je jako swoistego rodzaju wyzwania, którym na pewno chce sprostać Małopolska Jednostka Ewaluacyjna.

## Bibliografia

**Martini A.**, *Wprowadzenie do podstawowych pytań stawianych w ewaluacji oddziaływania interwencji: Czy interwencja coś zmieniła? A jeśli tak, to dla kogo?* [w:] *Ocena wpływu i prognozowanie efektów w badaniach ewaluacyjnych*. (red.) Haber A., Trzeciński R., PARP, Warszawa 2011

Projektowanie i zlecanie ewaluacji wpływu opartych na sytuacji kontrfaktycznej. Praktyczne wytyczne dla Instytucji Zarządzających EFS. [http://www.archiwum.ewaluacja.gov.pl/Dokumenty\\_ewaluacyjne/Documents/Przewodnik\\_KE\\_po\\_metodach\\_kontrfaktycznych\\_PL.pdf](http://www.archiwum.ewaluacja.gov.pl/Dokumenty_ewaluacyjne/Documents/Przewodnik_KE_po_metodach_kontrfaktycznych_PL.pdf) (28.09.2016)

Wytyczne w zakresie ewaluacji polityki spójności na lata 2014-2020. [https://www.funduszeu-ropiejskie.gov.pl/media/9318/wytyczne\\_090915\\_final.pdf](https://www.funduszeu-ropiejskie.gov.pl/media/9318/wytyczne_090915_final.pdf) (28.09.2016)

**Szymański M.**, *Kontrola projektów w ramach polityki spójności oraz propozycje na okres programowania 2014–2020*. [w:] *Kontrola państwa*. Numer 3(344), maj – czerwiec 2012





### BARBARA PORĘBSKA-MACIOŁEK

Absolwentka studiów magisterskich i studiów doktoranckich Wydziału Nauk Społecznych Uniwersytetu Śląskiego oraz Podyplomowego Studium Ewaluacji Polityki Społecznych Uniwersytetu Warszawskiego. Doświadczenie w zakresie badań ewaluacyjnych zdobyła w Departamencie Polityki Regionalnej Urzędu Marszałkowskiego Województwa Małopolskiego w Krakowie, pełniąc rolę Instytucji Pośredniczącej PO KL, pracując na stanowisku Inspektora do spraw ewaluacji, obecnie zaś na stanowisku Koordynatora do spraw ewaluacji w Departamencie Zarządzania Programami Operacyjnymi. Od początku wdrażania PO KL w regionie zaangażowana w budowanie potencjału ewaluacyjnego oraz systemu ewaluacji. Brała udział w licznych szkoleniach z zakresu ewaluacji, badaniach ewaluacyjnych, w tym realizowanych wspólnie z OECD, Ministerstwem Infrastruktury i Rozwoju oraz jednostkami naukowymi. Członek Grupy Sterującej Ewaluacją PO KL oraz Zespołu Sterującego Ewaluacją Polityki Spójności, działającego przy Ministerstwie Rozwoju. Członek Polskiego Towarzystwa Ewaluacyjnego. Autorka artykułu „Rola ewaluacji w zarządzaniu regionem”, współautorka Rocznika Ewaluacji oraz publikacji pt. „Ogólnopolska sieć współpracy w ramach badań regionalnych rynków pracy i ewaluacji Programu Operacyjnego Kapitał Ludzki”. Autorka publikacji naukowych poświęconych tematyce samorządu lokalnego w Federacji Rosyjskiej.



### **AGNIESZKA ZAWRZYKRAJ**

Absolwentka studiów magisterskich Uniwersytetu Ekonomicznego w Krakowie oraz Akademii Ewaluacji organizowanej przez EUROREG Centrum Studiów Regionalnych i Lokalnych Uniwersytetu Warszawskiego. Doświadczenie w zakresie badań ewaluacyjnych zdobyła w Departamencie Zarządzania Programami Operacyjnymi Urzędu Marszałkowskiego Województwa Małopolskiego w Krakowie, pełniącego rolę Instytucji Zarządzającej Programami Regionalnymi na stanowisku Inspektora do spraw ewaluacji. Członek Zespołu Sterującego Ewaluacją Polityki Spójności, działającego przy Ministerstwie Rozwoju. Od 2009 roku związana z zarządzaniem regionalnymi programami operacyjnymi, w tym odpowiedzialna m.in. za monitorowanie i sporządzanie sprawozdań z realizacji Małopolskiego Regionalnego Programu Operacyjnego na lata 2007-2013. Posiada uprawnienia trenerskie potwierdzone certyfikatem ukończenia kursu Trenera Zarządzania w szkole trenerów MATRIK.

## MONIKA LEMKE

### Wstęp

Obecna perspektywa finansowa 2014-2020 stawia przed ewaluacją i realizującymi ją Jednostkami Ewaluacyjnymi wiele wyzwań. Jednym z najważniejszych, obok dostarczania wiarygodnych informacji na temat rzeczywistych efektów interwencji i pomocy w podejmowaniu trafnych decyzji, jest dążenie do użyteczności i wysokiej jakości realizowanych badań ewaluacyjnych. Dbanie o jakość jest trudną sztuką, o której trzeba pamiętać i o którą warto zabiegać nie w momencie opracowywania przez wykonawcę raportu ewaluacyjnego, ale znacznie wcześniej – już podczas przygotowywania założeń do badań ewaluacyjnych, czyli tworzenia koncepcji badania i opracowywania kryteriów wyboru wykonawcy. W przeciwnym wypadku istnieje znacznie wyższe ryzyko otrzymania produktu, który nie będzie spełniał naszych oczekiwań i będzie dla nas nieużyteczny.

Dlaczego jakość w ewaluacji staje się tak ważna i co można zrobić aby ją nie tylko poprawić, ale też nieustannie podwyższać na każdym etapie realizowanych przez nas badań? Jakie korzyści dla zamawiającego przyniesie nasza zwiększona troska o jakość? Próbą odpowiedzi na te pytania jest ten artykuł oparty przede wszystkim na doświadczeniach Jednostki Ewaluacyjnej Województwa Zachodniopomorskiego, będących efektem poszukiwania nowych rozwiązań służących podniesieniu jakości w ewaluacji i zdobytych podczas realizacji najnowszych badań ewaluacyjnych.

Artykuł koncentruje się na dwóch aspektach związanych z wyborem wykonawcy: cenie oferty oraz jej jakości i kryteriach premiujących wybór dobrych merytorycznie propozycji.

### Kryterium ceny i jego rola w „kreowaniu” jakości

Cena jest, obok jakości, jednym z dwóch najważniejszych elementów każdej usługi, w tym także ewaluacji.

Jak pokazuje dotychczasowa praktyka powszechne jest stosowanie tradycyjnego, standardowego wzoru obliczania punktów za cenę, tj.:

$$C = \frac{C_{min}}{C_{of}} * L_{max}$$

gdzie:

$C_{min}$  – oferta z najniższą ceną

$C_{of}$  – cena badanej oferty

$L_{max}$  – waga, maksymalna liczba punktów w kryterium.

Przeważnie przyjmowanym stosunkiem kryteriów ceny do kryteriów jakościowych jest relacja 40:60. Przy tej najczęściej stosowanej wadze, wykorzystywanie powyższego wzoru powoduje, że prawie niemożliwe jest wybranie oferty z jedną z najwyższych cen, a jednocześnie proponującej nowe lub rzadko stosowane, niestandardowe rozwiązania czy metody i techniki badawcze. Ponadto przy jego zastosowaniu, nawet jeśli przewagę mają kryteria jakościowe, różnice w punktach są tak duże, że wygrywa oferta niekoniecznie dobra jakościowo, ale za to z niższą ceną.

Co można zrobić, aby zmniejszyć rolę ceny, w zamian za zwiększenie roli jakości? Istnieją różne sposoby wpływania na ceny ofert składanych przez wykonawców. Poniżej przedstawię kilka przykładów.

Pierwszym z nich jest zmiana wagi ceny i kryteriów jakościowych. Stosunkiem, przy którym różnice w punktach za cenę nie będą znaczące jest relacja 20:80, a nawet 10:90. Taka relacja umożliwia wybranie najlepszej jakościowo oferty, tym bardziej, że nowa ustawa Prawo zamówień publicznych dopuszcza nawet całkowitą rezygnację z kryterium ceny.

Kolejnym sposobem jest odejście od wskazanego wyżej, tradycyjnie stosowanego wzoru. Zamiast niego można zastosować jedną z dwóch możliwości: wzór logarytmiczny lub proporcjonalny. Obydwa wzory (przedstawione poniżej) zostały wykorzystane w badaniach ewaluacyjnych zlecanych przez Jednostkę Ewaluacyjną Województwa Zachodniopomorskiego.

W obydwu przypadkach konieczne jest podanie budżetu przeznaczonego na badanie, gdyż stanowi on punkt odniesienia i porównania cen złożonych ofert, a także zabezpiecza przed otrzymaniem zerowej czy ujemnej liczby punktów. Ponadto podanie budżetu uniemożliwia wybranie oferty z ceną przewyższającą ten przewidziany na badanie.

Wzór logarytmiczny jest następujący:

$$Lp = \frac{\ln(1 + (B - Cof))}{\ln(1 + (B - Cmin))} * Lmax$$

$Lp$  – liczba punktów oferty ocenianej

$\ln$  – funkcja logarytmu naturalnego

$B$  – budżet przeznaczony na badanie

$Cof$  – cena oferty ocenianej

$Cmin$  – 70% wartości zamówienia lub cena najtańszej oferty spośród tych, które zostaną (zgodnie z prawem zamówień publicznych) wykazane przez wykonawców jako realne

$Lmax$  – maksymalna liczba punktów jakie można uzyskać w ramach kryterium.

Przykłady zastosowania powyższego wzoru pokazują poniższe tabele.

**Tabela 1. Kryterium ceny obliczonej wg. logarytmu naturalnego – przykład 1 (waga ceny 40%)**

Budżet	C min	C1	C2	C3	C4	C5	C6
275 000,00	193 233	252 150	241 695	193 233	250 000	219 622,65	215 250
<b>Liczba punktów</b>		<b>35,49</b>	<b>36,82</b>	<b>40,00</b>	<b>35,81</b>	<b>38,62</b>	<b>38,89</b>
<b>Ranking ceny</b>		6	4	1	5	3	2

Źródło: Opracowanie własne.

**Tabela 2. Kryterium ceny obliczonej wg. logarytmu naturalnego – przykład 2 (waga ceny 40%)**

Budżet	C min	C1	C2	C3	C4
330 000	246 949,56	299 259	290 895	246 949,56	264 450
<b>Liczba punktów</b>		<b>36,49</b>	<b>37,34</b>	<b>40,00</b>	<b>39,16</b>
<b>Ranking ceny</b>		4	3	1	2

Źródło: Opracowanie własne.

Powyższe przykłady pokazują, jak wzór logarytmiczny zmniejsza („spłaszcza”) do kilku punktów różnice w punktach za cenę, co powoduje, że jej rola się zmniejsza na korzyść kryteriów jakościowych.

Natomiast wzór proporcjonalny jest następujący:

$$Lp = \frac{B - Cof}{B - Cmin} * Lmax$$

Lp – liczba punktów oferty ocenianej

B – budżet przeznaczony na badanie

Cof – cena oferty ocenianej

Cmin – 70% wartości zamówienia lub cena najtańszej oferty spośród tych, które zostaną (zgodnie z prawem zamówień publicznych) wykazane przez wykonawców jako realne

Lmax – maksymalna liczba punktów jakie można uzyskać w ramach kryterium.

Przykład zastosowania powyższego wzoru przedstawia poniższa tabela.

**Tabela 3. Kryterium ceny obliczonej wg. wzoru proporcjonalnego (waga ceny 40%)**

Budżet	C min	C1	C2	C3	C4	C5	C6
282 900,00	196 800	196 800	258 000	198 030	233 700	215 446,80	245 877
<b>Liczba punktów</b>		<b>40,00</b>	<b>11,57</b>	<b>39,43</b>	<b>22,86</b>	<b>31,43</b>	<b>17,20</b>
<b>Ranking ceny</b>		1	6	2	4	3	5

Źródło: Opracowanie własne.

Powyższy przykład pokazuje jak wzór proporcjonalny „premiuje” oferty najtańsze i zbliżone do najtańszych. W tym wypadku różnice między ofertami wynoszą kilkadziesiąt punktów. Trzeba jednak pamiętać, że wzór ten został zastosowany przy 40% wadze dla ceny. Różnice w punktach za cenę zmniejszą się do kilku lub kilkunastu punktów, gdy waga ceny będzie wynosiła odpowiednio 10 lub 20%. Pokazują to poniższe tabele.

**Tabela 4. Kryterium ceny obliczonej wg. wzoru proporcjonalnego (waga ceny 20%)**

Budżet	C min	C1	C2	C3	C4	C5	C6
282 900,00	196 800	196 800	258 000	198 030	233 700	215 446,80	245 877
<b>Liczba punktów</b>		<b>20,00</b>	<b>5,78</b>	<b>19,71</b>	<b>11,43</b>	<b>15,67</b>	<b>8,60</b>
<b>Ranking ceny</b>		1	6	2	4	3	5

Źródło: Opracowanie własne.

**Tabela 5. Kryterium ceny obliczonej wg. wzoru proporcjonalnego (waga ceny 10%)**

Budżet	C min	C1	C2	C3	C4	C5	C6
282 900,00	196 800	196 800	258 000	198 030	233 700	215 446,80	245 877
<b>Liczba punktów</b>		<b>10,00</b>	<b>2,89</b>	<b>9,86</b>	<b>5,71</b>	<b>7,83</b>	<b>4,30</b>
<b>Ranking ceny</b>		1	6	2	4	3	5

Źródło: Opracowanie własne.

Przy zmniejszonej wadze ceny, także ten wzór „pomaga” w premiowaniu jakości i zwiększeniu roli kryteriów merytorycznych. Jednocześnie powyższe dwa przykłady wskazują, że warto zastanowić się nad zastosowaniem połączenia dwóch rozwiązań: zmniejszenia wagi ceny nawet do 10% i zastosowania wzoru innego niż standardowy: proporcjonalnego lub logarytmicznego.

### Kryteria jakościowe i ich rola w „kreowaniu” jakości

W ewaluacji jakość wpływa na użyteczność, a tym samym na stopień wykorzystania jej wyników ewaluacji oraz decyduje o rzetelności i wiarygodności wyników, które mogą pomóc w podejmowaniu decyzji dotyczących nie tylko usprawnień w zarządzaniu i wdrażaniu programów regionalnych, ale także decyzji strategicznych związanych z kreowaniem polityki regionalnej.

Aby ewaluacja osiągnęła tak wysoką jakość konieczne jest stosowanie zaawansowanych metod badawczych oraz nowych, niestandardowych, niekiedy innowacyjnych rozwiązań i merytorycznych kryteriów jakościowych, już na etapie wyboru wykonawcy badań ewaluacyjnych. Jest to wyzwanie dla obu stron procesu ewaluacyjnego – zamawiających i wykonawców, gdyż wymaga otwartości w poszukiwaniu i stosowaniu nowych metod i technik badawczych oraz analiz, a nie tylko tradycyjnych typu CAWI, CATI czy IDI. Jednak decydującą rolę w tym zakresie mają zamawiający, gdyż to oni kreują rynek badań ewaluacyjnych i poprzez stawiane wymagania i kryteria wyboru ofert wpływają na stosowanie przez wykonawców innych niż dotychczasowe rozwiązań i sposobów działania.

W jaki sposób można wpływać na jakość na etapie wyboru wykonawcy? Jednym z możliwych rozwiązań jest wprowadzenie w zapytaniach ofertowych, czy w specyfikacjach istotnych warunków zamówienia zapisów promujących jakość. Przykładem takiego zapisu jest zapis o konieczności otrzymania w każdym z podkryteriów merytorycznych określonej przez zamawiającego minimalnej liczby punktów - *„Zamawiający - chcąc zagwarantować wysoką jakość i profesjonalizm realizacji zadania – wymaga, by w każdym z podkryteriów oferta wykonawcy uzyskała minimum 50% punktów możliwych do otrzymania. (...) Jeżeli oferta uzyska mniej niż 50% możliwych do otrzymania punktów w danym podkryterium, zostanie odrzucona jako niespełniająca wymagań Zamawiającego na podstawie art. 89 ust.1 pkt 2 ustawy PZP („jej treść nie odpowiada treści istotnych warunków zamówienia”).”*

Takie podejście wymaga od potencjalnych wykonawców bardziej uważnego czytania wymagań zawartych w zapytaniach ofertowych czy specyfikacjach istotnych warunków zamówienia oraz staranniejszego przygotowania ofert i zawartych w nich koncepcji badawczych. Zastosowanie tego rozwiązania umożliwiło Jednostce Ewaluacyjnej Województwa Zachodniopomorskiego odrzucenie słabych niespełniających oczekiwań zamawiającego ofert. Tym samym proponowane rozwiązanie wpływa na podniesienie jakości, na której nam zamawiającym najbardziej zależy.

Kolejną propozycją jest stosowanie kryteriów typu: trafność, spójność. Kryteria te, dobrze opisane, pomagają sprawdzić, jak wykonawca rozumiał badanie i jego cele oraz specyfikę. Są trudniejsze do oceny, ale dzięki nim zamawiający może prześledzić tok rozumowania wykonawcy i sprawdzić, czy jest ono zgodne z jego oczekiwaniami i wymaganiami. Ponadto kryteria te pomagają w różnicowaniu wykonawców, ale też wymagają od zamawiającego odpowiedniego wyboru członków komisji oceniającej (przetargowej) oferty w zakresie merytorycznej znajomości tematu objętego badaniem ewaluacyjnym.

Jednostka Ewaluacyjna Województwa Zachodniopomorskiego już od kilku lat stosuje kryteria trafności i spójności do oceny proponowanych przez wykonawców koncepcji czy metodyki badania. Przykładowe kryteria są przedstawione w poniższych tabelach.

**Tabela 6. Kryteria merytoryczne dotyczące metodyki badania – przykład 1**

Kryteria merytoryczne	Max ilość punktów	Min ilość punktów
spójność koncepcji badania: czy wykonawca przejrzyście wskazał sposób, w jaki poszczególne elementy badania łączą się ze sobą umożliwiając realizację celów i zakresu badania	15	7,5
trafność oparcia koncepcji badawczej na teorii badawczej (badań ewaluacyjnych lub społeczno-ekonomicznych) lub koncepcji naukowej (dotyczącej przedmiotu badania popartej odniesieniem/ odniesieniami do naukowych publikacji tematycznych)	15	7,5
trafność zaproponowanych metod i technik badawczych w kontekście realizacji celu badania, pozyskania rzetelnego (umożliwiającego dokonywanie analiz ilościowych i jakościowych) materiału badawczego i udzielenia odpowiedzi na postawione pytania ewaluacyjne	10	5
trafność dodatkowych pytań badawczych uwzględniających specyfikę badania oraz specyfikę grupy badawczej.	10	5
trafność zidentyfikowania potencjalnych obszarów ryzyka mogących wystąpić w trakcie realizacji badania i środka/-ów go minimalizującego/ych w kontekście celów i zakresu badania oraz wpływających na poprawę rzetelności i precyzji badania	10	5
<b>Suma punktów za kryteria merytoryczne</b>	<b>60</b>	<b>30</b>

Źródło: Opracowanie własne.

**Tabela 7. Kryteria merytoryczne dotyczące metodyki badania – przykład 2**

Kryteria merytoryczne	Max ilość punktów	Min ilość punktów
spójność koncepcji badania – czy wykonawca przejrzyście wskazał sposób, w jaki poszczególne elementy badania łączą się ze sobą umożliwiając realizację celów i zakresu badania. Ponadto koncepcja badania ma być autorską modyfikacją zapisów zawartych w SOPZ, a nie ich wiernym powtórzeniem	15	7,5
trafność oparcia koncepcji badawczej na teorii badawczej (badań ewaluacyjnych lub społeczno-ekonomicznych) lub koncepcji naukowej (dotyczącej przedmiotu badania popartej odniesieniem/odniesieniami do naukowych publikacji tematycznych)	10	5
trafność zaproponowanych w ramach planu badawczego metod i technik badawczych w kontekście realizacji celu badania, pozyskania rzetelnego (umożliwiającego dokonywanie analiz ilościowych i jakościowych) materiału badawczego i udzielenia odpowiedzi na postawione pytania ewaluacyjne.	10	5

trafność zaproponowanych w ramach planu badawczego metod i technik analizy danych. Metody i techniki analizy powinny być opisane i uzasadnione.	10	5
trafność dodatkowych pytań badawczych uwzględniających specyfikę badania oraz specyfikę grupy badawczej. Dodane pytania muszą się odnosić do celów i zakresu badania, a także muszą zostać uzasadnione przez Wykonawcę. Pytania powinny pogłębiać zakres badania tzn. wprowadzać nowe zagadnienia, pogłębiające przedmiot badania, które nie zostały ujęte przez Zamawiającego i będą adekwatne do celów badania będącego przedmiotem zamówienia oraz będą istotne z punktu widzenia użyteczności wyników.	10	5
Trafność zidentyfikowania potencjalnego ryzyka mogącego wystąpić w trakcie realizacji badania i środka środków go minimalizującego minimalizujących w kontekście celów i zakresu badania oraz wpływających na poprawę rzetelności i precyzji badania - do 1 pkt za każde zidentyfikowane potencjalne ryzyko, łącznie max 5 pkt. Ocenie będzie podlegać pierwsze 5 zidentyfikowanych potencjalnych ryzyk wskazanych przez Wykonawcę.	5	2,5
<b>Suma punktów za kryteria merytoryczne</b>	<b>60</b>	<b>30</b>

Źródło: Opracowanie własne.

Powyższe przykłady różnią się między sobą. Jest to istotne, ponieważ kryteria jakościowe trzeba dobierać pod kątem tematyki i specyfiki badania. Szczególnie ważne wydają się być kryteria dotyczące trafności metod i technik analizy danych i ich interpretacji. To właśnie analiza i interpretacja zebranego materiału badawczego oraz wyciągnięte na jego podstawie wnioski i wypływające z nich rekomendacje w znacznym stopniu decydują o jakości badań ewaluacyjnych i rzetelności oraz wiarygodności i użyteczności prezentowanych w nich wyników.

Ostatnią propozycją jest wprowadzenie wymagania dotyczącego przedstawienia przez wykonawcę w ofercie tzw. planu badawczego. Plan badawczy i jego trafność jest jednym z merytorycznych kryteriów oceny ofert zastosowanym przez Jednostkę Ewaluacyjną Województwa Zachodniopomorskiego. Został on zdefiniowany jako zestaw metod i technik badawczych stosowanych na etapie obserwacji, analizy i oceny zmierzających do realizacji celów badania i wypracowania odpowiedzi na postawione pytania badawcze. Niestety ta definicja planu badawczego okazała się mało precyzyjna, ponieważ została odczytana jako zachęta do typowego przedstawienia zestawu metod i technik badania i analizy stanowiącego trzon późniejszego raportu metodologicznego. Tymczasem plan badawczy nie jest raportem metodologicznym.

W ostatnich latach w środowisku ewaluacyjnym przyjęto, że rzetelnym badaniem dostarczającym mocnych dowodów jest takie, w którym plan badawczy i metody są dobrze dopasowane do pytań badawczych. Jest to tzw. platynowy standard, który w praktyce oznacza podejście macierzowe, pokazujące optymalne konfiguracje pytań i planów badawczych. Podejście to przedstawia poniższa tabela.



**Tabela 8. Relacja między pytaniami badawczymi a metodyką (planem badawczym)**

	Wiedza kontekstowa – jaka jest sytuacja, jakie problemy, trend zmian?	Wiedza o efektach – czy interwencja dała efekt?	Wiedza o efektach – czy interwencja może zadziałać w innym kontekście?	Wiedza o mechanizmach – dlaczego wystąpiły zmiany?	Wiedza operacyjna – jak wdrażać płynnie?
Meta-analiza i systematyczny przegląd źródeł	++	-	+++	+	-
Eksperymentalny plan badawczy	-	+++	-	-	-
Studium statystyczne	+	++	+++	-	-
Gra symulacyjna	-	-	+	+++	+
Ewaluacja oparta na teorii	-	++	++	+++	+
Podejście partycypacyjne	++	+	+	++	+++
Studium przypadku	+	+	-	+++	++
Opisowy plan badawczy	+++	-	-	+	++

„+++” oznacza bardzo wysoką użyteczność, „-” oznacza brak użyteczności danego planu badawczego przy poszukiwaniu wiedzy danego typu

Źródło: K. Olejniczak, T. Kupiec, E. Raimondo, Brokerzy wiedzy. Nowe spojrzenie na rolę jednostek ewaluacyjnych.

Upraszczając, pytania badające relacje przyczynowo-skutkowe (oceniające wpływ interwencji na daną grupę docelową) są najlepiej adresowane podejściami eksperymentalnymi i ilościowymi opartymi na analizie kontrfaktycznej. Badania wyjaśniające mechanizm – dlaczego interwencja działa – opierają się na mieszanych jakościowo-ilościowych podejściach badawczych. Odpowiedzi na pytania zarządcze wymagają interakcji i studiów przypadków.

Plan badawczy jest projektem badania. Wiedząc co chcemy zbadać, wiemy jakie metody i techniki będą odpowiednie. Stanowi on logiczne ramy do zbierania i analizy danych. Jest całościową logiką badania – poszukiwania i udzielania rzetelnej i wiarygodnej odpowiedzi na pytania badawcze. Co istotne – różne plany badawcze mogą wykorzystywać te same metody i techniki. Tym co odróżnia plany badawcze nie są poszczególne metody i techniki, ale logika podejścia do badania. Koncepcja planów badawczych jest obecnie nowym, obiecującym kierunkiem w badaniach ewaluacyjnych. Plany badawcze wzmacniają jakość ewaluacji dzięki zachowaniu standardów naukowych oraz jednocześnie obejmują całe bogactwo metod badawczych, dając przy tym przejrzyste i względnie proste ramy do oceny dopasowania metodyki do potrzeb wiedzy.

Wyróżnia się osiem głównych planów badawczych charakteryzujących się następującymi cechami, przedstawionymi w poniższej tabeli.

**Tabela 9. Plany badawcze i ich charakterystyka**

Plan badawczy	Cechy charakterystyczne
Meta-analiza i systematyczny przegląd źródeł	Logika tego planu opiera się na formułowaniu wniosków i oceny na podstawie wcześniejszych badań i ewaluacji dotyczących analogicznego przedmiotu lub tematu.
Eksperymentalny plan badawczy	Logika tego planu opiera się na schemacie właściwym dla badań eksperymentalnych oraz quasi eksperymentalnych i jest podporządkowana identyfikacji efektów danej interwencji poprzez porównanie obiektów (np. osób należących do tej samej populacji), dla których objęcie lub nieobjęcie interwencji stanowi jedyną cechę różnicującą.
Studium statystyczne	Plan badawczy opiera się na zastosowaniu narzędzi statystyki opisowej, w celu określenia relacji między interwencją a oczekiwanymi zmianami.
Gra symulacyjna	Logika planu badawczego opiera się na wyjaśnianiu działania interwencji poprzez odtworzenie i symulację w warunkach laboratoryjnych jej rzeczywistych mechanizmów.
Ewaluacja oparta na teorii	Według tego planu, interwencja jest zestawem szczegółowych założeń przyczynowo-skutkowych, których zasadność można zweryfikować poszukując dowodów na ich wystąpienie w rzeczywistości.
Podejście partycypacyjne	Logiką tego planu jest zbieranie różnych perspektyw aktorów skoncentrowanych wokół danego zagadnienia, ich wyważenie, zrozumienie i wypracowanie wspólnego rozwiązania problemu.
Studium przypadku	Studium przypadku polega na dogłębnym przestudiowaniu aspektów, stanu lub procesów jednego przypadku.
Opisowy plan badawczy	Logika tego planu opiera się na opisywaniu zjawisk, ich skali, przebiegu, dynamiki zmiany w czasie, a także zmian opinii, przy wykorzystaniu bezpośredniej obserwacji lub danych zastanych.

Źródło: EGO s.c. i PGS (2015) Brokerzy wiedzy – gra symulacyjna, Warszawa: [www.knowledgebrokers.edu.pl](http://www.knowledgebrokers.edu.pl).

Wymaganie dotyczące stosowania planów badawczych w ofertach przedstawiających koncepcję realizacji określonego badania ewaluacyjnego powoduje, że wykonawcy muszą wyjść poza dotychczasowe schematy przygotowywania ofert i być bardziej kreatywnymi w ich przygotowywaniu. Plany badawcze wymuszają stosowanie zindywidualizowanego podejścia do każdego badania dopasowanego do potrzeb zamawiającego.

### **Doświadczenie wykonawcy a jakość badania**

Osobnym zagadnieniem, także mającym wpływ na jakość jest doświadczenie wykonawcy. Problemem w badaniach ewaluacyjnych, które najczęściej zlecane są na podstawie przepisów prawa zamówień publicznych, jest to że są one usługami priorytetowymi. Oznacza to, że zamawiający nie mogą przyznawać punktów za doświadczenie. Mogą jedynie stawiać wymagania do-

tyczące posiadania doświadczenia przez wykonawcę i członków jego zespołu badawczego. Dlatego też, patrząc na dynamiczny wzrost liczby firm wykonujących badania ewaluacyjne, istotne jest zwiększenie wymagań dotyczących liczby i wartości zrealizowanych badań ewaluacyjnych, tak by ograniczyć krąg wykonawców do tych, którzy posiadają większe doświadczenie w realizacji badań ewaluacyjnych. Wtedy zamawiający zwiększy szansę, że wykonawcy z większym doświadczeniem będą wykonywać bardziej rzetelne, wiarygodne i wyższej jakości badania ewaluacyjne.

### **Podsumowanie**

Przedstawione zagadnienia koncentrują się na etapie wyboru wykonawcy. Odrębną kwestią jest kontrola jakości i jej wyegzekwowanie na etapie realizacji badania oraz opracowywania rekomendacji i raportu końcowego, a także podczas odbierania przez zamawiającego wykonanego dzieła. Tutaj ważne jest określenie wymagań jakościowych dotyczących raportu końcowego i ich konsekwentne stosowanie podczas odbioru raportu z badania.

Można, podobnie jak w przypadku interwencji współfinansowanych ze środków unijnych, rozważyć wprowadzenie zasady koncentracji, zwłaszcza w odniesieniu do pytań badawczych. Chodzi tu przede wszystkim o ograniczenie ich liczby i koncentrację na kluczowych zagadnieniach badawczych.

Warto w przypadku stosowania planów badawczych zastanowić się nad rezygnacją z raportu metodologicznego, który jest powtórzeniem oferty złożonej przez wykonawcę. Natomiast czas poświęcony na jego przygotowywanie wykorzystać na skonstruowanie narzędzi badawczych i przeprowadzenie dogłębnej analizy dostępnych danych i dokumentów metodą desk research. Zamiast niego warto stosować raport wstępny, w którym powinny znaleźć się pierwsze ustalenia badawcze oraz analiza teorii zmiany interwencji. Ponadto warto zacząć częściej wymagać od wykonawców przygotowywania raportów częściowych, aby na bieżąco kontrolować postępy w realizacji badania i móc odpowiednio wcześniej zareagować na pojawiające się problemy i trudności w realizacji.

Istotny jest też harmonogram badania i zapewnienie dłuższego czasu na wstępną fazę badania oraz wyodrębnienie fazy przygotowania rekomendacji, które często decydują o końcowym odbiorze badania przez kadrę zarządzającą i decydentów.

Zasygnalizowane w tym miejscu zagadnienia są istotne dla zapewnienia jakości w całym procesie ewaluacji, ale są tematem na odrębny artykuł.

Jednostka Ewaluacyjna Województwa Zachodniopomorskiego, podobnie jak wszyscy, nadal będzie poszukiwać sposobów oceny i poprawy jakości. Ma też nadzieję, że ten artykuł stanie się inspiracją do dalszych twórczych poszukiwań w tym obszarze.

## Bibliografia

EGO s.c. i PGS, *Brokerzy wiedzy - gra symulacyjna*, www.knowledgebrokers.edu.pl, Warszawa 2015

**Koba D.**, *Kryteria oceny ofert w postępowaniach o udzielenie zamówienia publicznego. Materiał szkoleniowy*, Koba Consulting Group, Warszawa 2014

**Olejniczak K., Kupiec T., Raimondo E.**, *Brokerzy wiedzy. Nowe spojrzenie na rolę jednostek ewaluacyjnych*, [w:] *(R)ewaluacja 2. Wiedza w działaniu*, (red.) Haber A., Olejniczak K., PARP, Warszawa 2014

Ustawa Prawo zamówień publicznych z dnia 29 stycznia 2004 r.

*Specyfikacja istotnych warunków zamówienia dla postępowania prowadzonego w trybie przetargu nieograniczonego o wartości poniżej wyrażonej w złotych równowartości kwoty 209.000 euro na wykonanie badania ewaluacyjnego pn. „Ewaluacja podsumowująca postęp rzeczowy i rezultaty RPO WZ 2007-2013.”*, Biuletyn Informacji Publicznej: <http://www.bip.wzp.pl/artukul/15042016-r-przetarg-nieograniczony-o-wartosci-ponizej-209000-euro-na-wykonanie-badania>, 2016

*Specyfikacja istotnych warunków zamówienia dla postępowania prowadzonego w trybie przetargu nieograniczonego o wartości poniżej wyrażonej w złotych równowartości kwoty 207.000 euro na wykonanie badania ewaluacyjnego pn. „Ocena wsparcia w obszarze rewitalizacji w ramach Regionalnego Programu Operacyjnego Województwa Zachodniopomorskiego na lata 2007-2013 oraz identyfikacja potencjału i potrzeb regionu w zakresie rewitalizacji.”*, Biuletyn Informacji Publicznej: <http://bip.rbip.wzp.pl/artukul/30062015-r-przetarg-nieograniczony-o-wartosci-ponizej-207000-euro-na-wykonanie-badania>, 2015



### MONIKA LEMKE

Absolwentka Katolickiego Uniwersytetu Lubelskiego, Wyższej Szkoły Integracji Europejskiej oraz studiów podyplomowych: Zarządzanie Projektami Celowymi Finansowanymi ze środków UE Uniwersytetu Szczecińskiego, Ewaluacja programów finansowanych ze środków UE Akademii Leona Koźmińskiego, Zamówienia publiczne Wyższej Szkoły Bankowej, Zarządzanie rozwojem lokalnym i regionalnym Zachodniopomorskiej Szkoły Biznesu. Pracownik Urzędu Marszałkowskiego Województwa Zachodniopomorskiego. Od 2007 roku w Jednostce Ewaluacyjnej Regionalnego Programu Operacyjnego Województwa Zachodniopomorskiego. W zakresie ewaluacji zajmuje się zlecaniem badań na zewnątrz, współpracą z wykonawcami i nadzorem nad realizacją.

## JOLANTA RUDNICKA

### Wprowadzenie

Niniejszy artykuł jest rozwinięciem prezentacji przedstawionej na III międzyregionalnej konferencji ewaluacyjnej pn. Nowe wyzwania dla ewaluacji programów operacyjnych w perspektywie 2014-2020 (Toruń, 21-22 czerwca 2016 r.).

Siedmioletnie doświadczenia w zakresie prowadzenia procesu ewaluacji (zamawiania badań, również ich realizacji, współpracy z adresatami wyników, uczestniczeniu w tworzeniu kultury ewaluacyjnej w strukturze administracji publicznej itp.) zebrane przez Jednostkę Ewaluacyjną Regionalnego Programu Operacyjnego Województwa Kujawsko-Pomorskiego (JE RPO WK-P) mogą uprawniać już do podzielenia się spostrzeżeniami w tym zakresie.

Świadomość o potrzebie prowadzenia badań ewaluacyjnych dla wdrażanych programów operacyjnych przez administrację samorządową na szczeblu regionalnym jest na tyle duża, że można przypuszczać, że byłaby realizowana nawet gdyby nie była narzucona zapisami aktów prawnych<sup>1</sup>.

Warto jednak podkreślić, że w perspektywie 2014-2020 rola ewaluacji i wymagania w tym zakresie zostały znacząco wzmocnione. Podjęto kroki w kierunku standaryzacji pewnych elementów w procesie ewaluacji polityki spójności<sup>2</sup>.

### Umiejscowienie Jednostki Ewaluacyjnej w strukturze Instytucji Zarządzającej RPO WK-P

Jednostka Ewaluacyjna RPO WK-P (JE) została wyodrębniona w 2009 r. w formie wydzielonego, nowego biura w Wydziale Zarządzania RPO, w którym wcześniej oprócz ewaluacji – zajmowano się również programowaniem. Była to dość znacząca zmiana, bo zadania związane z ewaluacją programu zostały wydzielone, jednak cały czas JE pozostawała w strukturze Instytucji Zarządzającej RPO WK-P (IZ RPO WK-P). Miało to swoje zalety, gdyż przepływ informacji odnośnie programowania interwencji był bezpośredni, a w przypadku pojawienia się jakichkolwiek problemów w zarządzaniu czy wdrażaniu programu podejmowana była szybka decy-

1 Rozporządzenie Rady (WE) 1083/2006 ustanawiające przepisy ogólne dot. Europejskiego Funduszu Rozwoju Regionalnego, Europejskiego Funduszu Społecznego oraz Funduszu Spójności i uchylające rozporządzenie (WE) nr 1260/1999.

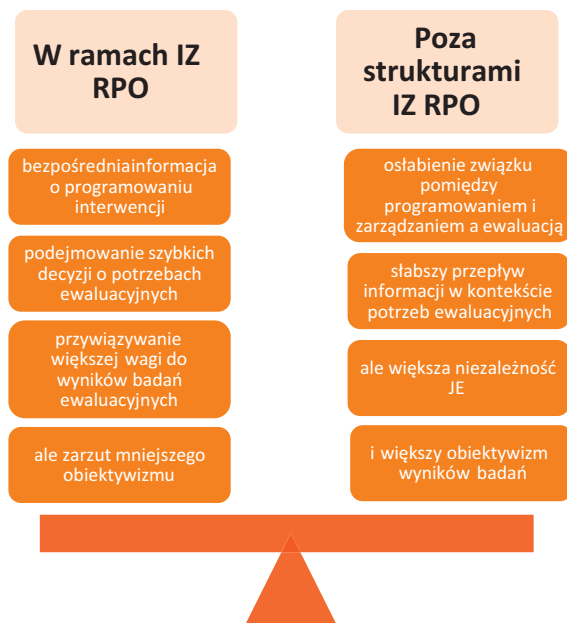
Rozporządzenie Parlamentu Europejskiego i Rady(UE) nr 1303/2013 ustanawiające wspólne przepisy dotyczące Europejskiego Funduszu Rozwoju Regionalnego, Europejskiego Funduszu Społecznego oraz Funduszu Spójności i uchylającego rozporządzenie (WE) nr 1083/2006.

2 Mówią o tym m.in. Wytyczne w zakresie ewaluacji polityki spójności na lata 2014-2020 (nr MlIR/H/2014-2020/27(1)09/2015) z dnia 22 września 2015.

zja o potrzebie przeprowadzenia procedury ewaluacji. Zdobywano w ten sposób informacje, które pomagały w rozwiązaniu problemów i jednocześnie można było zauważyć, że odbiorcy ewaluacji przywiązują większą wagę do wyników badań. W tym czasie, kiedy JE stanowiła jedną z komórek IZ RPO, zrealizowano dwie ewaluacje wewnętrzne, które dostarczyły rzetelnych informacji w krótkim czasie i pozwoliły na podjęcie przez IZ RPO środków zaradczych i zmiany podejścia w przygotowywaniu kolejnych konkursów. Umieszczenie jednak JE w strukturze IZ RPO obarczone było zarzutem mniejszego obiektywizmu działań i tym samym wyników badań. Pomimo tego warto podkreślić, że bliskość JE i zarządzania programem ma wiele zalet (nie można jednak przenosić tego stwierdzenia na wszystkie jednostki w kraju, gdyż w pewnym stopniu uzależnione jest to też od indywidualnej współpracy i podejścia pracowników).

Wraz z nową perspektywą zmieniło się umiejscowienie JE (znalazła się ona poza strukturami IZ RPO), a także rozszerzono jej zadania, choćby z uwagi na dwufunduszowość regionalnego programu operacyjnego (Europejski Fundusz Rozwoju Regionalnego, Europejski Fundusz Społeczny). Krótkie doświadczenie funkcjonowania w nowej strukturze pokazuje osłabienie związku pomiędzy programowaniem i zarządzaniem a ewaluacją oraz słabszy przepływ informacji w kontekście potrzeb ewaluacyjnych. Na pewno JE zyskała większą niezależność, a tym samym nie powinno być zarzutów dotyczących małego obiektywizmu wyników badań. Ocena umiejscowienia JE w strukturze IZ RPO pokazuje ryc.1.

**Rycina 1. Umiejscowienie JE w strukturze IZ RPO – słabe i mocne strony każdej z lokalizacji**



Źródło: Opracowanie własne.

### Dorobek badań ewaluacyjnych JE

W okresie od 2009 do 2016 roku Jednostka Ewaluacyjna przeprowadziła w sumie 23 badania, na realizację których wydano kwotę prawie 1,84 mln zł (por. ryc. 2). Procedury w ramach których zlecone zostały badania to:

- 14 postępowań przetargowych,
- 7 postępowań tzw. podprogowych (5 zamówień poniżej 14 tys. euro i po 2014 r. 2 zamówienia poniżej 30 tys. euro).

Ponadto wykonano 2 ewaluacje wewnętrzne.

### Rycina 2. Statystyka przeprowadzonych badań ewaluacyjnych w latach 2009-2016

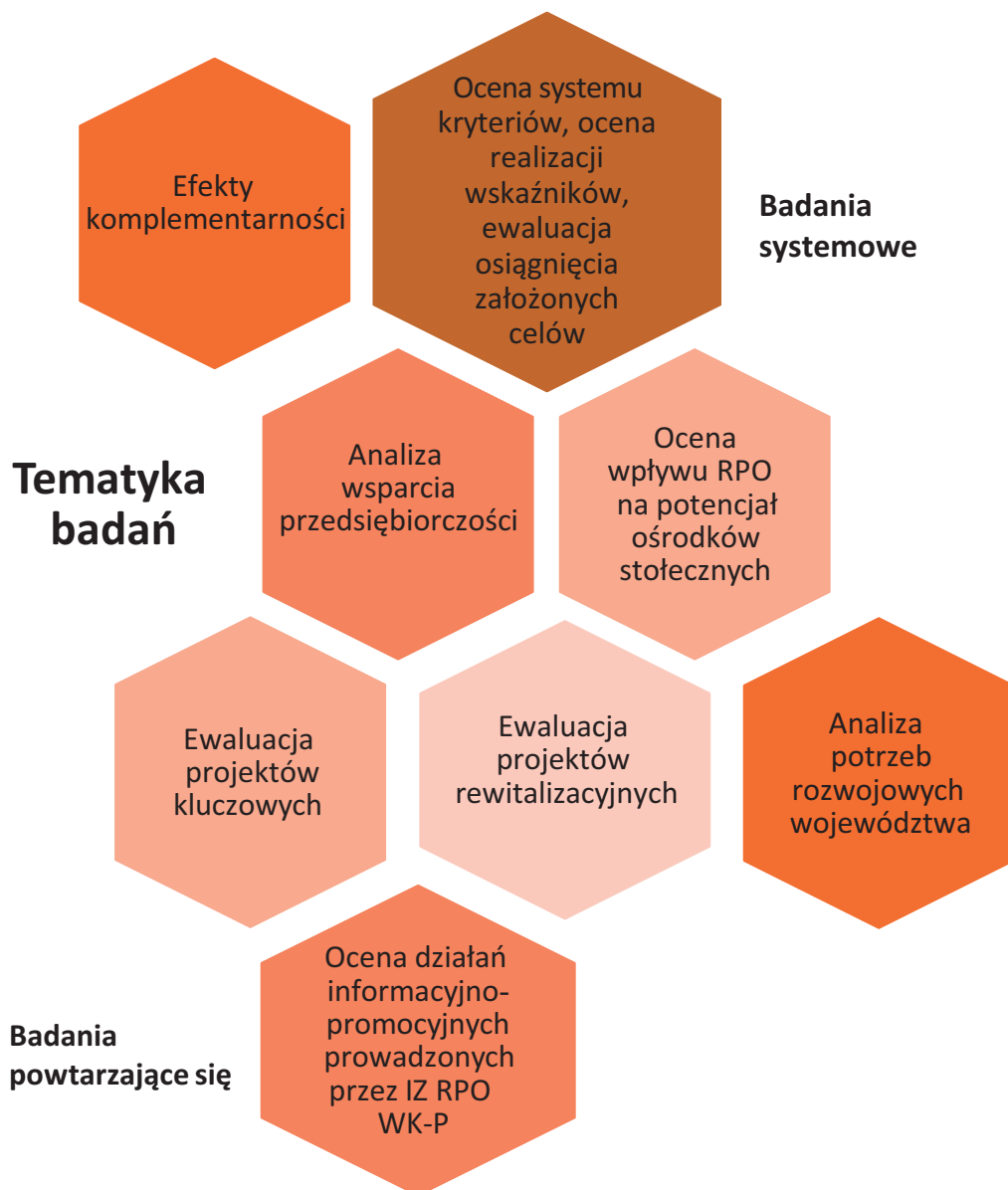


Źródło: Opracowanie własne.

Biorąc pod uwagę tematykę prowadzonych badań była ona na tyle różnorodna, że pozwoliła na zebranie kompleksowej wiedzy o realizacji regionalnego programu operacyjnego, począwszy od oceny systemu kryteriów, oceny realizacji wskaźników, ewaluacji osiągnięcia założonych celów programu, przez ocenę działań informacyjno-promocyjnych prowadzonych przez IZ RPO WK-P, po analizę wsparcia przedsiębiorczości, efekty komplementarności działań, ocenę projektów kluczowych, a także analizę potrzeb rozwojowych województwa. Cenne doświadczenia zebrane zostały podczas prowadzenia ewaluacji ex ante w modelu partycypacyjnym dla nowego programu operacyjnego (por. ryc. 3).



Rycina 3. Tematyka wybranych badań ewaluacyjnych

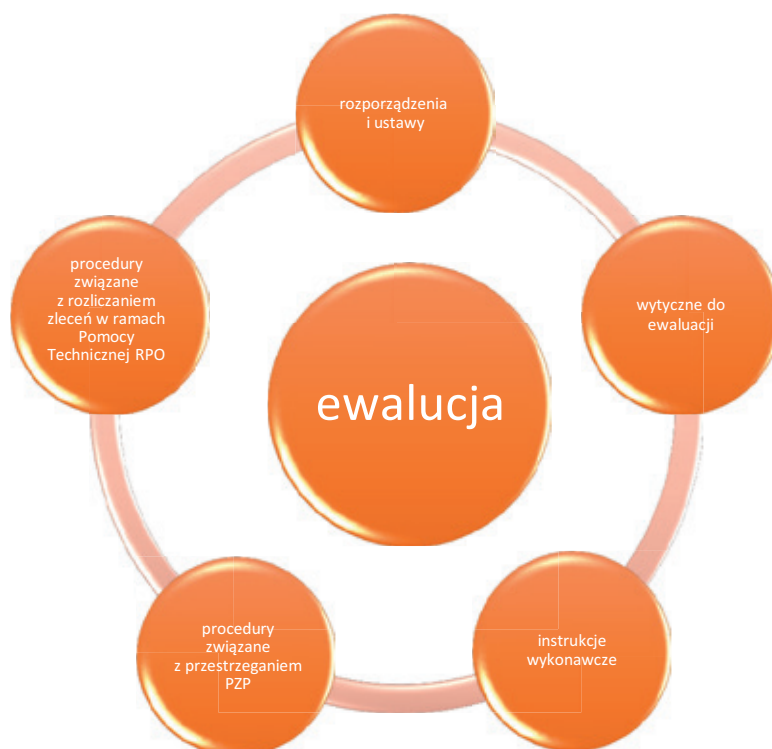


Źródło: Opracowanie własne.

## Ewaluacja w gąszczu procedur urzędowych

Zarówno w perspektywie 2007-2013 jak i 2014-2020 istnieje szereg zapisów prawnych mówiących o tym kiedy wymagane jest przeprowadzenie ewaluacji, jak powinien przebiegać jej proces i kto jest za niego odpowiedzialny (por. ryc. 4).

**Rycina 4. Regulacje prawne dla ewaluacji**



Źródło: Opracowanie własne.

Jednostki ewaluacyjne zlecające przeprowadzanie badań muszą zwracać uwagę z jednej strony na zapisy aktów prawnych dotyczących postępowań o udzielanie zamówień publicznych, z drugiej zaś strony na obwarowania związane z prawodawstwem unijnym i wydatkowaniem środków pochodzących z RPO WK-P.

Zestawienie najistotniejszych regulacji prawnych obowiązujących dla perspektyw finansowych UE przedstawia tabela 1.

**Tabela 1. Regulacje prawne dla ewaluacji programów operacyjnych w perspektywie 2007-2013 oraz 2014-2020**

2007-2013	2014-2020
<b>Rozporządzenia</b>	
<p><b>ROZPORZĄDZENIE RADY (WE) nr 1083/2006</b> z dnia 11 lipca 2006 ustanawiające przepisy ogólne dotyczące Europejskiego Funduszu Rozwoju Regionalnego, Europejskiego Funduszu Społecznego oraz Funduszu Spójności i uchylające rozporządzenie(WE) nr 1260/1999</p> <p>Wymóg przygotowania oceny nakłada na Polskę art. 33 ust.1 lit. c; art. 47, 48, 49, 60 lit c i e oraz art. 65 lit. c, d i f, art. 68 pkt.3</p>	<p><b>Rozporządzenie Rady UE nr 1303/2013</b> ustanawiające wspólne przepisy dotyczące Europejskiego Funduszu Rozwoju Regionalnego, Europejskiego Funduszu Społecznego oraz Funduszu Spójności i uchylające rozporządzenie (WE) nr 1083/2006</p> <p>Zapisy dotyczące ewaluacji zawarte zostały w art.4 ust.5, art.5 ust.3 lit. d; art7; art.15 ust. 1 ppkt. a lit. ii; art.16 ust.1; art. 26 ust.4; art. 27 ust.4; art.28 ust.2 ; art. 29 ust.1; art. 33 ust.1 lit. f; art. 34 ust.3 lit. g; art.35 ust.1 lit. d; art.37 ust.2 i 3, art. 39 ust.4 lit. a; art.49 ust.4; art.50 ust.2,4 i 5; art.53 ust.1, art. 54; art.55; art. 56; art.56; art.57; art.58 ust.1 lit. e, g i h; art.96 ust.2 lit. a, ust.5 lit. c; art. 106 ust.8 lit. c; art. 110 ust.1 lit. b, ust.2 lit. c; art.111 ust.4 lit. a i c; art. 114; art. 125 lit. d</p>
	<p><b>ROZPORZĄDZENIE PARLAMENTU EUROPEJSKIEGO I RADY (UE) NR 1304/2013</b> z dnia 17 grudnia 2013 r. w sprawie Europejskiego Funduszu Społecznego i uchylające rozporządzenie Rady (WE) nr 1081/2006</p> <p>Zapisy dotyczące ewaluacji zawarte zostały w art.7, art. 9 ust.1 i 4, art.19, załącznikach 1 i 2.</p>
<b>Ustawy</b>	
<p><b>Ustawa z dnia 6 grudnia 2006 r. o zasadach prowadzenia polityki rozwoju</b> <i>Dz.U. Nr 84 z 2009r. poz.712</i></p> <p>Zapisy dotyczące ewaluacji zostały zawarte w art. 5 pkt.11, art. 12a pkt. 1 ppkt.3, art. 15 pkt.6,7,8, art. 17 pkt.1, art.26 pkt.1 ppkt. 12</p>	<p><b>Ustawa z dnia 6 grudnia 2006 r. o zasadach prowadzenia polityki rozwoju</b> <i>Dz.U. z 2015 r. poz.1890</i></p> <p>Zapisy dotyczące ewaluacji zostały zawarte w art. 3a pkt.4a; art. 6b pkt.1,2,3; art. 10a pkt.1i 2; art. 14i ppkt.1; art. 14l pkt.1, 2 ppkt.3; art.15 pkt.6, art. 17 pkt.1 ppkt.4 i 6, pkt.1a; art. 26 pkt.1 ppkt.12.</p>
	<p><b>Ustawa z dnia 11 lipca 2014 r. o zasadach realizacji programów w zakresie polityki spójności finansowanych w perspektywie finansowej 2014-2020 (Dz. U. z 2016 r. poz. 217),</b></p> <p>Zapisy dotyczące ewaluacji zostały zawarte w rozdziale 8 art.26 pkt.1</p>

2007-2013	2014-2020
<b>Wytyczne</b>	
<p><b>Wytyczne MRR nr 6 w zakresie ewaluacji programów operacyjnych na lata 2007-2013 z dnia 30 maja 2007 r.</b></p>	<p><b>Wytyczne MliR w zakresie ewaluacji polityki spójności 2014-2020 z 22 września -2015 r.</b></p>
<p><b>Wytyczne MRR w zakresie informacji i promocji z dnia 10 czerwca 2010 r.</b></p> <p>Zapisy dotyczące ewaluacji zostały zawarte w rozdziale 8 pkt.5 i 6.</p>	<p><b>Wytyczne MliR w zakresie informacji i promocji programów operacyjnych polityki spójności na lata 2014-2020 z dnia 30 kwietnia 2015 r.</b></p> <p>Zapisy dotyczące ewaluacji zostały zawarte w rozdziale 8 podrozdział 8.2.</p>
<p><b>Wytyczne MRR w zakresie sprawozdawczości z dnia 30 sierpnia 2011 r. (zmienione dnia 30 marca 2016 r.)</b></p> <p>Zapisy dotyczące ewaluacji zostały zawarte w rozdziale 3; rozdz. 7 podrozdz.7.1 pkt.1 oraz w załącznikach nr 2a i 2 c do wytycznych.</p>	<p><b>Wytyczne MliR w zakresie sprawozdawczości na lata 2014-2020 z dnia 8 maja 2015 r.</b></p> <p>Zapisy dotyczące ewaluacji zostały zawarte w rozdziale 4 pkt.2; rozdz. 6 podrozdz. 6.2 pkt.1-7, tabeli 2 pkt.4 i 17.</p>
<p><b>Wytyczne MRR nr 3 dotyczące komitetów monitorujących z dnia 6 września 2014 r.</b></p> <p>Zapisy dotyczące ewaluacji zostały zawarte w rozdziale 6 pkt.3.</p>	<p><b>Wytyczne MliR w zakresie komitetów monitorujących na lata 2014 – 2020 z dnia 21 stycznia 2015 r.</b></p> <p>Zapisy dotyczące ewaluacji zostały zawarte w rozdziale 5 pkt.2 ppkt. d, pkt.3 ppkt. b, pkt.5 ppkt. c; rozdz.8 pkt.2 ppkt.p.</p>
<p><b>Wytyczne MRR dotyczące dokonywania przeglądu i renegocjacji z Komisją Europejską programów operacyjnych w ramach Narodowych Strategicznych Ram Odniesienia na lata 2007-2013 z dnia 1 września 2008 r.</b></p> <p>Zapisy dotyczące ewaluacji zostały zawarte w rozdziale 3 pkt.3 ppkt.c; pkt.5</p>	<p><b>Wytyczne MIR w zakresie monitorowania postępu rzeczowego realizacji programów operacyjnych na lata 2014-2020 z dnia 22 kwietnia 2015r.</b></p> <p>Zapisy dotyczące ewaluacji zostały zawarte w rozdziale 1 podrozdz. 1.2 pkt.2; rozdziale 2 podrozdz.2.1.3 pkt. 3 i 4; podrozdz. 2.2.1 pkt.2; rozdziale 3 podrozdz. 3.1.2 pkt.2 i 6; podrozdz. 3.3.1 pkt. 2; podrozdz.3.3.4 pkt. 3 i 6 oraz załącznik nr 6.</p>
	<p><b>Wytyczne MIR w zakresie realizacji zasady partnerstwa na lata 2014-2020 z dnia 28 października 2015 r.</b></p> <p>Zapisy dotyczące ewaluacji zostały zawarte w rozdziale 3 podrozdz.3.1 pkt.5 a, e i f; podrozdz.3.2 pkt.1; rozdz.4 pkt.1 i 3; rozdz.5 pkt.1, 8 ppkt. f; Rozdz. 6 podrozdz.6.3 pkt.1; rozdz.7 pkt.1 rozdz.8 podrozdz.8.1 pkt. pkt.2 ppkt. b; podrozdz.8.2 pkt.1 ppkt. d i g; podrozdz.8.3 pkt.1, 2 i 3; sekcja 8.3.4 oraz załącznik 1 do wytycznych.</p>

2007-2013	2014-2020
	<p><b>Wytyczne MIR w zakresie realizacji przedsięwzięć z udziałem środków Europejskiego Funduszu Społecznego w obszarze rynku pracy na lata 2014-2020 z dnia 22 lipca 2015 r.</b></p> <p>Zapisy dotyczące ewaluacji zostały zawarte w rozdziale 3 podrozdz. 3.1 pkt.3 i 6.</p>
	<p><b>Wytyczne MIR w zakresie realizacji zasady równości szans i niedyskryminacji, w tym dostępności dla osób z niepełnosprawnościami oraz zasady równości szans kobiet i mężczyzn w ramach funduszy unijnych na lata 2014-2020 z dnia 8 maja 2015 r.</b></p> <p><i>Zapisy dotyczące ewaluacji zostały zawarte w Rozdziale 5 Podrozdział 5.2 Sekcja 5.2.3 pkt.2 ppkt.d, Sekcja 5.2.4 pkt. 1-3; Rozdziale 6 Podrozdział 6.2 sekcja 6.2.3 pkt.1-4.</i></p>
	<p><b>Wytyczne MIR w zakresie realizacji przedsięwzięć z udziałem środków Europejskiego Funduszu Społecznego w obszarze edukacji na lata 2014-2020 z dnia 2 czerwca 2015 r.</b></p> <p>Zapisy dotyczące ewaluacji zostały zawarte w rozdziale 3 podrozdział 3.2 pkt.6 ppkt. c; rozdział 6 podrozdział 6.3 pkt.5 ppkt. f.</p>
	<p><b>Wytyczne w zakresie realizacji przedsięwzięć z udziałem środków Europejskiego Funduszu Społecznego w obszarze przystosowania przedsiębiorców i pracowników do zmian na lata 2014-2020 z dnia 10 kwietnia 2015 r.</b></p> <p>Zapisy dotyczące ewaluacji zostały zawarte w rozdziale 4 podrozdział 4.1 pkt.20 i 21; podrozdział 4.2 pkt.9.</p>
<p><b>Indicative Guidelines On Evaluation Methods: Monitoring And Evaluation Indicators Working Document No. 2</b> (Wytyczne metod ewaluacji: monitoring i ewaluacja wskaźników )</p>	<p><b>Monitoring and evaluation of European Cohesion Policy. European Social Fund. Guidance document.</b> (Wytyczne KE w zakresie prowadzenia monitorowania i oceny EFS)</p>
<p><b>Indicative Guidelines On Evaluation Methods: Ex Ante Evaluation Working Document No.1</b> (Wytyczne metod ewaluacji : ewaluacja ex ante)</p>	<p><b>Guide on ex-ante evaluation</b> (Wytyczne do ewaluacji ex ante)</p>
<p><b>Indicative Guidelines On Evaluation Methods: Evaluation during The Programming Period Working Document No.5</b> (Wytyczne metod ewaluacji: ewaluacja w procesie programowania)</p>	<p><b>Guidance document on evaluation plans</b> (Wytyczne dot. tworzenia planów ewaluacji)</p>

2007-2013	2014-2020
<b>Plany ewaluacji</b>	
<p><b>Plan ewaluacji Regionalnego Programu Operacyjnego</b> przyjęty uchwałą nr 86/1113/07 przez Zarząd Województwa w dniu 27 grudnia 2007 roku z późn. zm. i przez Komitet Monitorujący RPO WK-P na lata 2007-2013 w dniu 29 lutego 2008 r. uchwałą nr 8/2008 roku.</p>	<p><b>Plan Ewaluacji Regionalnego Programu Operacyjnego Województwa Kujawsko-Pomorskiego na lata 2014-2020</b> przyjęty uchwałą Nr 35/2015 Komitetu Monitorującego Regionalny Program Operacyjny Województwa Kujawsko-Pomorskiego z dnia 15 grudnia 2015 roku</p>
<b>Instrukcja Wykonawcza Instytucji Zarządzającej Regionalnym Programem Operacyjnym Województwa Kujawsko-Pomorskiego</b>	
<p><b>Instrukcja Wykonawcza RPO WK-P 2007-2013 Wersja XIV-</b> Zmiany przyjęte Uchwałą nr 16/484/15 Zarządu Województwa Kujawsko-Pomorskiego z dnia 22 kwietnia 2015 r.</p>	<p><b>Instrukcja Wykonawcza Instytucji Zarządzającej Regionalnym Programem Operacyjnym Województwa Kujawsko-Pomorskiego na lata 2014-2020 wersja III, luty 2016 r.</b> Załącznik do uchwały nr 8/231/16 Zarządu Województwa Kujawsko-Pomorskiego z dnia 24 lutego 2016 r.</p>
<b>Akty prawne powiązane</b>	
<p>Ustawa z dnia 27 sierpnia 2009 r. o finansach publicznych (Dz. U. z 2013 r. poz. 885 z późn. zm.);</p>	<p>Ustawa z dnia 27 sierpnia 2009 r. o finansach publicznych (Dz. U. z 2013 r., poz. 885 z późn. zm.),</p>
<p>Ustawa z dnia 29 stycznia 2004 r. prawo zamówień publicznych (Dz. U. z 2013 r. poz. 907 z późn. zm.);</p>	<p>Ustawa z dnia 29 stycznia 2004 r. - Prawo zamówień publicznych (Dz. U. z 2015 r., poz. 2164),</p>
<p>Ustawa z dnia 29 sierpnia 1997 r. o ochronie danych osobowych (Dz.U. 2014 poz. 1182 z późn.zm.)</p>	<p>Ustawa z dnia 29 sierpnia 1997 r. o ochronie danych osobowych (Dz. U. z 2015 r. poz. 2135 z późn. zm.)</p>
<p>Ustawa z dnia 29 sierpnia 1997 r. o ochronie danych osobowych (Dz.U. 2014 poz. 1182 z późn.zm.)</p>	<p>Ustawa z dnia 6 września 2001 r. o dostępie do informacji publicznej (Dz. U. z 2015 r. poz. 2058 z późn. zm.)</p>
<p>Uchwała NR 57/1131/11 Zarządu Województwa Kujawsko-Pomorskiego z dnia 21 września 2011 r. w sprawie określenia zasad i trybu udzielania zamówień publicznych w zakresie wykonywania zadań Województwa Kujawsko-Pomorskiego nie zastrzeżonych do kompetencji wojewódzkich jednostek organizacyjnych i wojewódzkich osób prawnych</p>	<p>Uchwała NR 6/166/16 Zarządu Województwa Kujawsko-Pomorskiego z dnia 10 lutego 2016 r. w sprawie określenia zasad i trybu udzielania zamówień publicznych w zakresie wykonywania zadań Województwa Kujawsko-Pomorskiego nie zastrzeżonych do kompetencji wojewódzkich jednostek organizacyjnych i wojewódzkich osób prawnych</p>
<p>Wytyczne IZ RPO WK-P na lata 2007-2013 w sprawie udzielania zamówień współfinansowanych ze środków Europejskiego Funduszu Rozwoju Regionalnego, w stosunku do których beneficjenci nie są zobowiązani do stosowania ustawy Prawo Zamówień Publicznych;</p>	

Źródło: Opracowanie własne.

Rozszerzanie obowiązku badań ewaluacyjnych na gruncie polityk krajowych świadczy o rosnącej świadomości polskich decydentów, a także o potrzebie ustalenia obiektywnych metodologicznie kryteriów ewaluacji dokumentów programowych. Potwierdzeniem tego jest fakt, że w dniu 24 maja 2014 roku weszła w życie nowelizacja Ustawy o Zasadach Prowadzenia Polityki Rozwoju, która nałożyła wymóg realizacji ewaluacji ex ante dla wszystkich strategii rządowych oraz dla programów krajowych o budżecie powyżej 300 mln zł. Dodatkowo wskazano, iż minister właściwy ds. rozwoju regionalnego odpowiedzialny jest za koordynację i realizację ewaluacji polityki rozwoju w Polsce, a także za ustalanie standardów i zaleceń dla ewaluacji. Określono, że w ramach urzędu obsługującego ministra zadania te pełni Krajowa Jednostka Ewaluacyjna.

Perspektywa 2014-2020 w porównaniu z poprzednim okresem programowania przyniosła jeszcze więcej zapisów, m.in. w formie wytycznych odnośnie konieczności prowadzenia ewaluacji, bądź korzystania z jej wyników, dlatego jednostki ewaluacyjne zobligowane są do wypełniania wszystkich zaleceń.

### Plany ewaluacji jako kluczowy element w pracy JE

Trudno wyobrazić sobie, aby we współczesnym świecie funkcjonować bez ustalonego planu działania. Dla sprawnego zarządzania procesem ewaluacji zarówno w perspektywie 2007-2013 (roczne plany) jak i 2014-2020 (jeden wieloletni plan) sporządzane są plany ewaluacji. Pozwalają one na odpowiednią organizację pracy w Jednostce Ewaluacyjnej, wymuszają poniekąd dokonywanie dostatecznie wcześniej przeglądu potrzeb ewaluacyjnych w IZ RPO, a także są argumentem dla gromadzenia danych i poszukiwania wyzwań badawczych (por. ryc. 5).

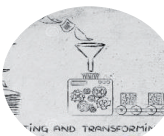
### Rycina 5. Najważniejsze zadania jakie pełnią Plany ewaluacji



porządkują i organizują prace w JE



wymuszają odpowiednio wcześniej dokonanie przeglądu potrzeb ewaluacyjnych w IZ



są argumentem dla gromadzenia danych i poszukiwania wyzwań badawczych

Źródło: Opracowanie własne.

W okresie 2007-2013 funkcjonowały dwa rodzaje planów: strategiczne i operacyjne. Plany strategiczne obejmowały cały okres wdrażania – przedstawiając organizację systemu ewaluacji, jak również najważniejsze obszary badań. Plany operacyjne natomiast definiowały konkretne tematy badań, a także najważniejsze działania (np. konferencje, szkolenia) jakie zostaną podjęte w danym roku kalendarzowym. W okresie programowania 2014-2020 w sposób znaczący zwiększyły się wymagania Komisji Europejskiej dotyczące procesu planowania ewaluacji. Plany ewaluacji stały się obowiązkiem wynikającym z zapisów rozporządzenia<sup>3</sup>, a ich zakres i sposób przygotowania oraz zatwierdzania został szczegółowo doprecyzowany w wytycznych ministerialnych<sup>4</sup>.

Każdy plan zawiera zestaw badań do realizacji w określonym czasie, a w przypadku Planu ewaluacji dla RPO WK-P na lata 2014-2020 - oprócz listy badań i opisu zastosowanej metodologii - znajdują się informacje dotyczące organizacji procesu ewaluacji i budowy potencjału ewaluacyjnego, w tym plan finansowy, sposoby rozpowszechniania i wykorzystania wyników ewaluacji, planowane szkolenia, a także monitoring realizacji planu oraz zasady jego aktualizacji, w przypadku pojawienia się potrzeby zbadania określonego problemu czy zjawiska.

### **Przykład kompleksowego badania**

Jednym z zaplanowanych do realizacji badań dla RPO WK-P 2007-2013 była ewaluacja pn. *Wpływ projektów realizowanych w ramach RPO WK-P na lata 2007-2013 na zatrudnienie w województwie kujawsko-pomorskim*. Było to duże badanie o charakterze ewaluacji ex post, przeprowadzone w terminie od kwietnia do połowy sierpnia 2015 r. przez Fundację Idea Rozwoju. Główną przesłanką do realizacji tego badania była konieczność przeliczenia wielkości osiągniętej wartości dla wskaźnika *Liczba bezpośrednio utworzonych nowych etatów (EPC) – w podziale na płeć*, która oszacowana została na 3 200, z równym udziałem mężczyzn i kobiet.

Z uwagi na to, że system gromadzenia danych nie posiadał takich możliwości, aby monitorować ten wskaźnik, już na początkowym etapie realizacji programu zaplanowano przeprowadzenie przedmiotowego badania. Uzyskane informacje dotyczące ww. wskaźnika wykorzystane zostały bezpośrednio w sprawozdawczości programu.

3 Rozporządzenie Parlamentu Europejskiego i Rady(UE) nr 1303/2013 ustanawiające wspólne przepisy dotyczące Europejskiego Funduszu Rozwoju Regionalnego, Europejskiego Funduszu Społecznego oraz Funduszu Spójności i uchylające rozporządzenie (WE) nr 1083/2006.

4 Wytyczne w zakresie ewaluacji polityki spójności na lata 2014-2020, MliR, 22 września 2015 r.



Celem głównym badania ewaluacyjnego było określenie wpływu projektów realizowanych w ramach RPO WK-P na lata 2007-2013 na zatrudnienie w województwie kujawsko-pomorskim poprzez analizę ich skuteczności i efektywności. Badanie realizowało ponadto trzy cele szczegółowe, tj.:

- Diagnozę stanu rynku pracy w województwie kujawsko-pomorskim.
- Ocenę realizacji wskaźników celu strategicznego i wskaźników horyzontalnych pod kątem osiągnięcia wartości docelowych.
- Analizę i ocenę wpływu projektów realizowanych w ramach RPO WK-P 2007-2013 na wzrost lub utrzymanie zatrudnienia w województwie kujawsko-pomorskim.

W badaniu wykorzystano zróżnicowanie liczebne próby (n) oraz zastosowano zasadę triangulacji wykorzystując następujące metody i techniki badawcze:

1. Analiza danych zastanych (*desk research*).
2. Badanie ilościowe projektów zrealizowanych w ramach RPO WK-P 2007-2013 (n = 886).
3. Wywiad grupowy (FGI) z beneficjentami RPO WK-P 2007-2013 (n=8).
4. Indywidualne wywiady pogłębione (IDI) z przedstawicielami IZ RPO WK-P 2007-2013 zajmującymi się koordynowaniem projektów, w ramach których utworzono miejsca pracy (n=7), z przedstawicielami Wojewódzkiego Urzędu Pracy (n=3), z przedstawicielami Prezydium Sejmiku Gospodarczego Województwa Kujawsko-Pomorskiego i przedsiębiorstw (n=5) oraz krajowymi i regionalnymi ekspertami rynku pracy (n=2).
5. Analiza SWOT.
6. Metaanaliza badań ewaluacyjnych, których celem było m.in. tworzenie miejsc pracy.
7. Spotkanie warsztatowe.
8. Studia przypadku (*case study*).

Głównym aspektem podlegającym badaniu było określenie skali miejsc pracy utworzonych w sposób bezpośredni oraz pośredni w związku z realizacją RPO WK-P.

Na podstawie zrealizowanych badań terenowych określono liczbę bezpośrednio utworzonych nowych etatów (EPC) - szacowaną wg stanu na czerwiec 2015 roku oraz planowaną do osiągnięcia na koniec realizacji RPO WK-P 2007-2013 (por. tab. 2).

**Tabela 2. Liczba utworzonych miejsc pracy, dzięki wsparciu z RPO WK-P 2007-2013**

Oś priorytetowa (OP)	Szacowana liczba utworzonych miejsc pracy (wg badania CAWI)	Liczba nowych etatów planowanych do utworzenia do końca realizacji Programu (wg danych z dokumentacji projektowej)
OP I - Rozwój infrastruktury technicznej	0	0
OP II - Zachowanie i racjonalne użytkowanie środowiska	49,20	52,00
OP III - Rozwój infrastruktury społecznej	459,00	373,66
OP IV - Rozwój infrastruktury społeczeństwa informacyjnego	94,00	88,80
OP V - Wzmocnienie konkurencyjności przedsiębiorstw	3 409,20 w tym 9 etatów badawczych	3 751,33 w tym 3 etaty badawcze
OP VI - Wsparcie rozwoju turystyki	244,90	307,23
OP VII - Wspieranie przemian w miastach i w obszarach wymagających odnowy	73,90	166,66
<b>Ogółem</b>	<b>4 330,20</b> <b>w tym 9 etatów badawczych</b>	<b>4 954,68</b> <b>w tym 3 etaty badawcze</b>

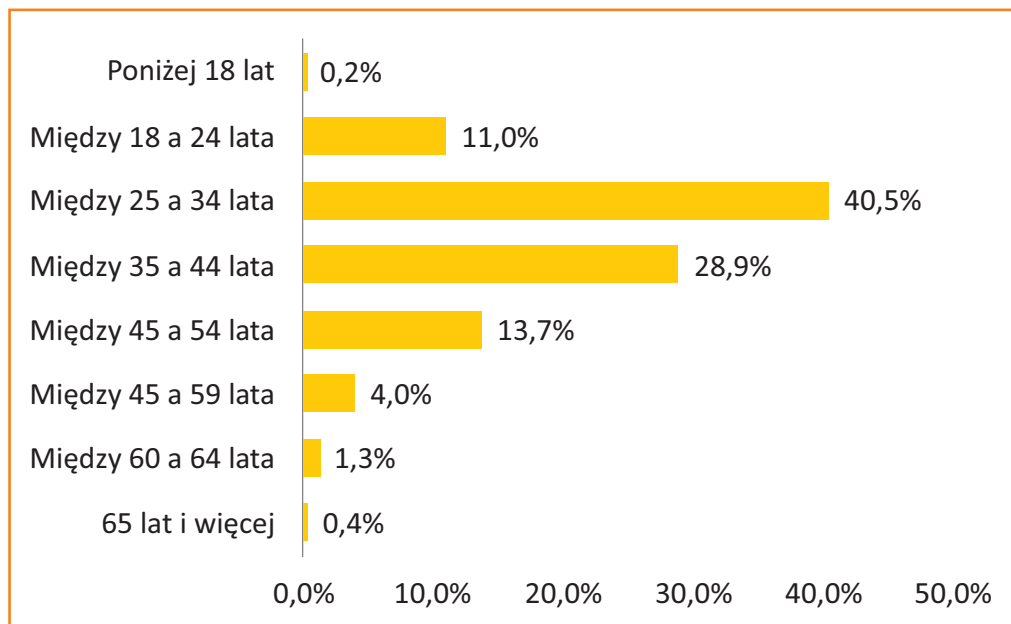
Źródło: Opracowanie własne Fundacji IDEA Rozwoju na podstawie wyników badania CAWI.

Zgodnie z wynikami ankietowania beneficjentów do momentu realizacji badania utworzonych zostało 4330,2 etatów, co oznacza, iż renegocjowana w 2011 roku wartość wskaźnika: *Liczba bezpośrednio utworzonych nowych etatów (EPC)* – w podziale na płeć, która oszacowana została na 3 200 została osiągnięta (w pierwszej wersji dokumentu ww. wskaźnik wynosił 8 388 etatów). Podział powstałych miejsc pracy ze względu na płeć przedstawia się następująco: mężczyźni (57%), kobiety (43%). W momencie, w którym realizowano badanie, niektóre projekty nie były zakończone. Dlatego docelowa liczba nowopowstałych miejsc pracy została oszacowana na poziomie 4 955 etatów.

Należy podkreślić, że w ramach przedmiotowego badania oprócz określenia stopnia realizacji wskaźnika strategicznego istotnym elementem były również analizy w podziale na podstawowe cechy nowoutworzonych miejsc pracy tj. wiek (ryc. 6), płeć (ryc. 7) i wykształcenie pracownika (ryc. 8) oraz stanowisko (ryc. 9) i sektor ekonomicznego działania beneficjenta (ryc. 10). Dokonanie szczegółowej analizy struktury miejsc pracy powstałych w wyniku realizacji RPO WK-P jest cennym elementem z punktu widzenia całościowego spojrzenia na zagadnienia związane z zatrudnieniem oraz skutecznego programowania polityki rozwoju w tym zakresie.

Biorąc pod uwagę wiek zatrudnionych w ramach RPO WK-P można stwierdzić, że największy odsetek stanowią osoby w przedziale 25-34 lata. Jest to grupa wiekowa, w której zgodnie z danymi statystycznymi występuje największy poziom bezrobocia w regionie, co pozwala wnioskować, że interwencja w ramach Programu była dobrze skorelowana z potrzebami województwa.

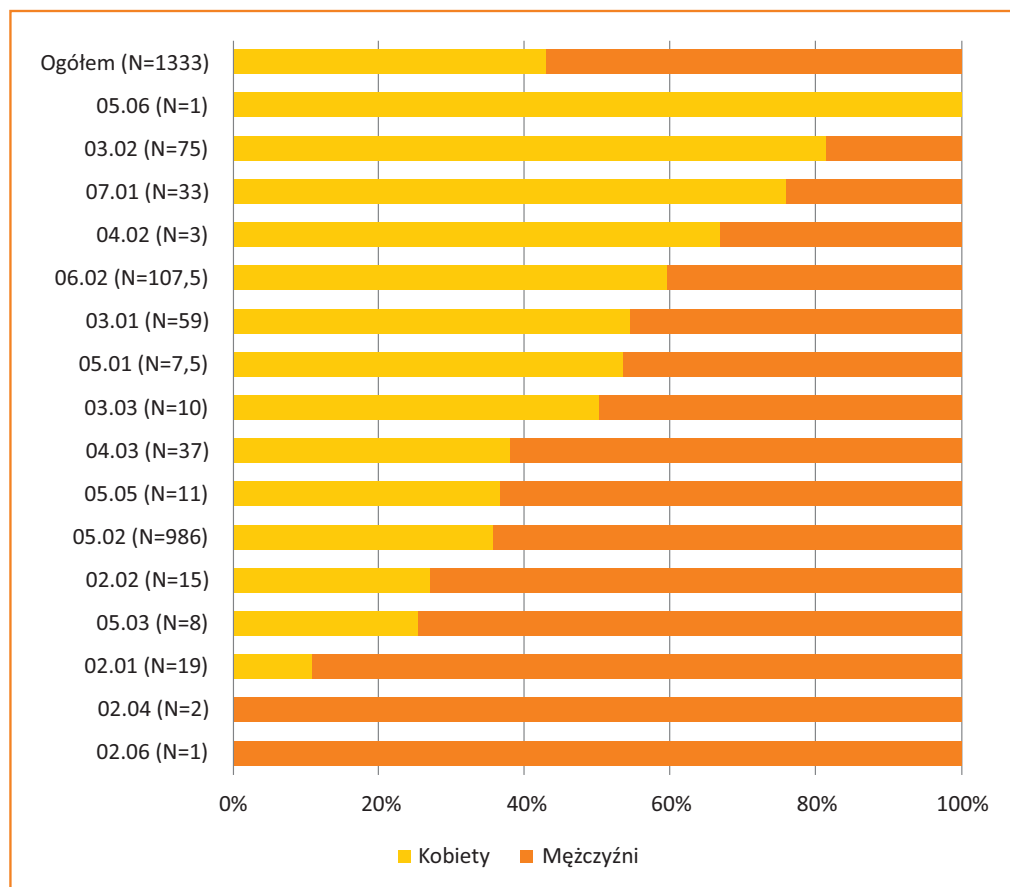
**Rycina 6. Rozkład nowoutworzonych miejsc pracy według wieku zatrudnionych (n=1312)**



Źródło: Opracowanie własne Fundacji IDEA Rozwoju na podstawie wyników badania CAWI.

Jak już wcześniej wspomniano, rozkład nowozatrudnionych w wyniku realizacji RPO WK-P wskazuje, iż 43% zatrudnionych stanowią kobiety, natomiast 57% mężczyźni. O ile patrząc na rozkład płci w ujęciu ogólnym różnice między kobietami, a mężczyznami nie są znaczące, jednak w poszczególnych osiach programu zauważalne są zdecydowane dysproporcje. W osiach/działaniach, w których beneficjentami byli przedsiębiorcy zdecydowanie częściej zatrudniano mężczyzn, podczas gdy kobiety częściej znajdowały pracę w ramach projektów realizowanych przez JST, jednostki opieki zdrowotnej, instytucje kultury oraz nauki.

**Rycina 7. Rozkład nowoutworzonych miejsc pracy według płci zatrudnionych w poszczególnych podziałaniach programu**

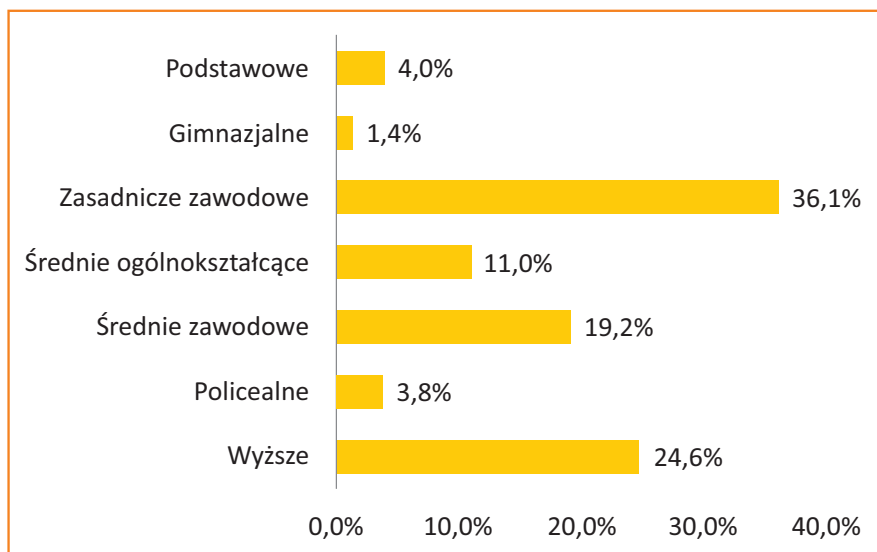


Źródło: Opracowanie własne Fundacji IDEA Rozwoju na podstawie wyników badania CAWI.

Zgodnie z rozkładem nowoutworzonych miejsc pracy według wykształcenia zatrudnionych należy stwierdzić, że największą grupę stanowią osoby z wykształceniem zasadniczym zawodowym, na drugim miejscu znalazły się osoby posiadające wykształcenie wyższe. Porównując wyniki analiz z danymi ze statystyki publicznej stwierdzono, że wykształcenie osób zatrudnionych w ramach Programu jest zbliżone do struktury wykształcenia wszystkich osób pracujących w regionie.

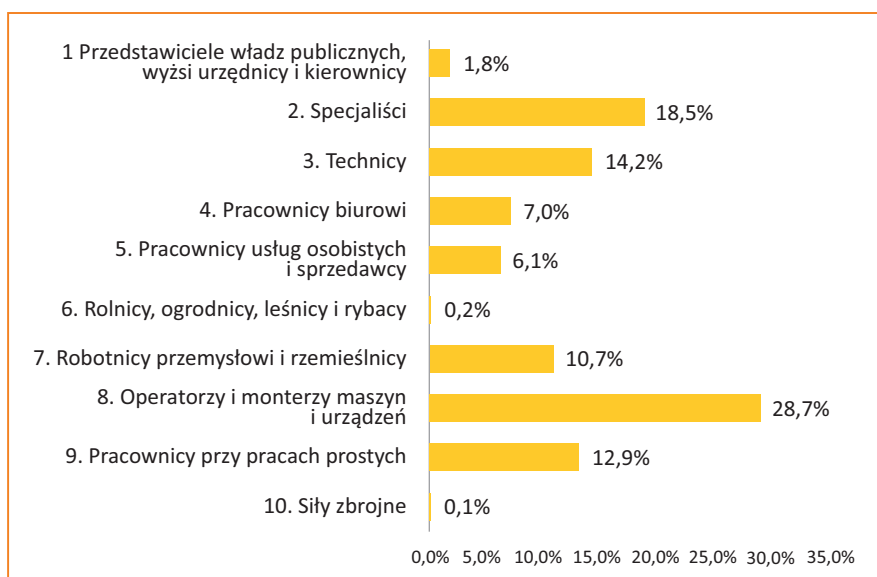
Wyniki analiz zrealizowanych w ramach niniejszego badania zgodnie z Klasyfikacją Zawodów i Specjalności MPiPS wykazały, że najwięcej osób zostało zatrudnionych na stanowiskach operatorów i monterów maszyn i urządzeń. Zaprezentowany rozkład miejsc pracy jest spójny z danymi dotyczącymi zmiany liczby bezrobotnych w regionie, jakie notowały urzędy pracy.

**Rycina 8. Rozkład nowoutworzonych miejsc pracy według wykształcenia zatrudnionych (n=1305)**



Źródło: Opracowanie własne Fundacji IDEA Rozwoju na podstawie wyników badania CAWI.

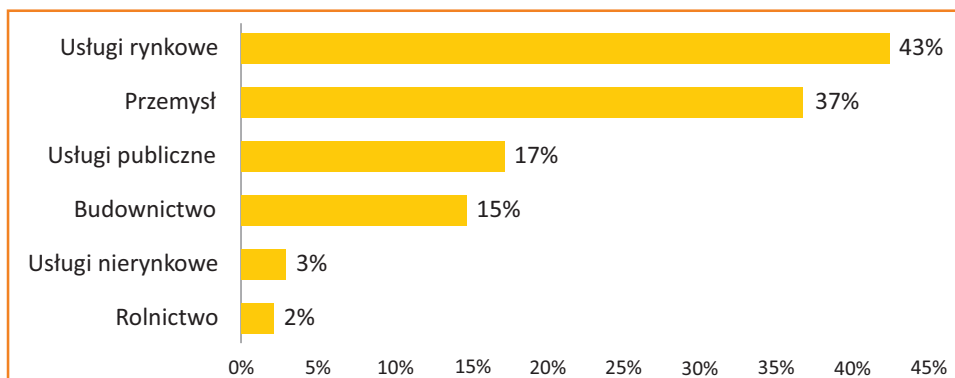
**Rycina 9. Zatrudnienie według stanowisk (zgodnie z Klasyfikacją Zawodów i Specjalności MPiPS) (n=1304)**



Źródło: Opracowanie własne Fundacji IDEA Rozwoju na podstawie wyników badania CAWI.

Uwzględniając sektor ekonomiczny, w którym działał beneficjent, większość miejsc pracy utworzonych zostało w usługach rynkowych, drugą grupę stanowili zatrudnieni w przemyśle, tendencja ta pokrywa się ze statystyką publiczną określającą, że w tych właśnie sektorach pracuje najwięcej mieszkańców regionu. Wnioskowano zatem, że interwencja w ramach Programu odpowiadała potrzebom województwa również w tym zakresie.

**Rycina 10. Zatrudnienie według sektora ekonomicznego działania beneficjenta**

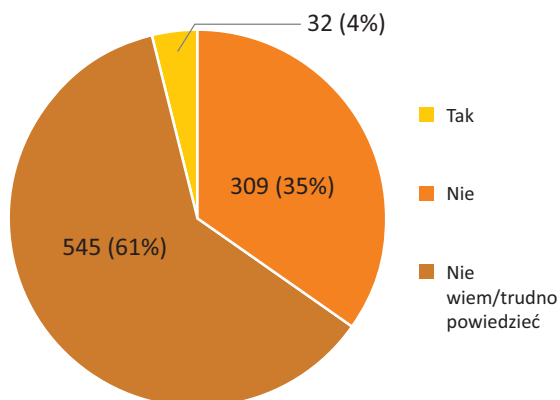


Źródło: Opracowanie własne Fundacji IDEA Rozwoju na podstawie wyników badania CAWI.

Jak już wcześniej wspomniano oprócz bezpośrednich efektów wsparcia w ramach RPO WK-P w omawianym badaniu określono udział projektów, dzięki którym utworzone zostały pośrednie miejsca pracy.

**Rycina 11. Nowe miejsca pracy jako efekty pośredniej realizacji projektów RPO WK-P**

Działanie	Liczba utworzonych miejsc pracy poza beneficjentami
01.01	8
02.02	20
03.02	14
04.03	1
05.01	467
05.02	36
05.06	32
06.01	2
06.02	24
07.01	44
<b>Ogółem</b>	<b>648</b>

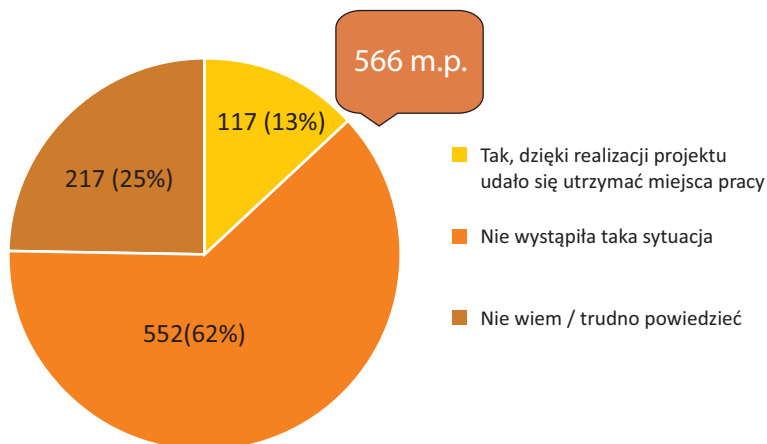


Źródło: Opracowanie własne Fundacji IDEA Rozwoju na podstawie wyników badania CAWI.

Na podstawie zrealizowanych badań terenowych stwierdzono, że w ramach interwencji finansowanych z RPO WK-P udało się utworzyć 648 miejsc pracy, które były pośrednim efektem wsparcia, tzn. zwiększenie zatrudnienia nie stanowiło głównego celu projektu (por. ryc.11).

Oprócz określenia liczby nowoutworzonych miejsc pracy (w sposób bezpośredni oraz pośredni) celem badania było również przedstawienie udziału projektów, które przełożyły się na utrzymanie miejsc pracy.

**Rycina 12. Udział projektów, które przełożyły się na utrzymanie miejsc pracy**



Źródło: Opracowanie własne Fundacji IDEA Rozwoju.

Na podstawie zebranych w ramach badania danych stwierdzono, że w przypadku 117 projektów fakt otrzymania wsparcia z RPO WK-P wpłynął na utrzymanie zatrudnienia, tzn. bez środków otrzymanych w ramach projektu beneficjent musiałby zlikwidować miejsca pracy. Łącznie beneficjenci zadeklarowali utrzymanie dzięki wsparciu 566 miejsc pracy (por. ryc. 12).

Z punktu widzenia celu badania, a także polityki prozatrudnieniowej w nowej perspektywie 2014-2020 bardzo istotnym elementem określonym w ramach badania były również plany zatrudnieniowe beneficjentów (por. tab. 3.).

**Tabela 3. Efektywność, skuteczność i trwałości tworzenia miejsc pracy w ramach RPO WK-P**

		sektor własności		ogółem (N=511)
		publiczny (N=171)	prywatny (N=341)	
<b>Czy do końca bieżącego roku planują Państwo zwiększać zatrudnienie?</b>	Zdecydowanie tak	1,80%	7,30%	5,50%
	Raczej tak	6,50%	26,40%	19,80%
	Raczej nie	54,70%	27,30%	36,40%
	Zdecydowanie nie	11,20%	16,40%	14,70%
	Trudno powiedzieć	25,90%	22,60%	23,70%
<b>Ogółem</b>		100,00%	100,00%	100,00%

Źródło: Opracowanie własne Fundacji IDEA Rozwoju.

Z uzyskanych danych wynikało, że zaobserwowany w trakcie realizacji badania wzrost zatrudnienia, będzie dalej występował. Dotyczy to przede wszystkim podmiotów sektora prywatnego, w przypadku którego ponad jedna trzecia badanych podmiotów zadeklarowała, że ma plany zwiększenia zatrudnienia do końca 2015 roku.

Ostatnim obszarem, którego dotyczyło badanie „Wpływ projektów realizowanych w ramach RPO WK-P na lata 2013-2020 na zatrudnienie w województwie kujawsko-pomorskim” była kwestia adekwatności miejsc pracy powstałych w wyniku RPO WK-P do potrzeb regionalnych.

Analizując utworzone nowe miejsca pracy wobec istniejących potrzeb regionalnych pod kątem struktury bezrobotnych zauważono się dużą korelację obu zjawisk. Tworzone miejsca pracy dotyczyły w większości osób z wykształceniem zawodowym, co koreluje ze strukturą bezrobocia regionu, gdzie przeważały osoby z wykształceniem zawodowym.

Zaobserwowano natomiast występowanie negatywnej korelacji pomiędzy poziomem bezrobocia a intensywnością wsparcia pochodząca z RPO WK-P 2007-2013 (ryc. 13). Inną ważną obserwacją było występowanie negatywnej zależności pomiędzy poziomem przedsiębiorczości i stopą bezrobocia, a wpływem Programu na zatrudnienie. W powiatach zlokalizowanych na obrzeżach województwa, najsilniej dotkniętych bezrobociem i najniższą relatywnie liczbą podmiotów gospodarczych, utworzono najmniej nowych miejsc pracy.

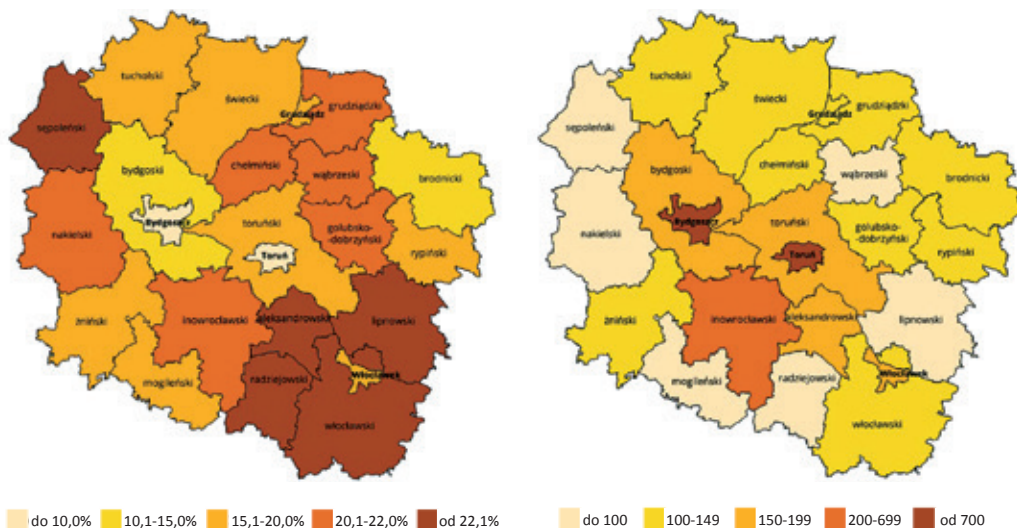
Najwięcej miejsc pracy dzięki projektom z regionalnego programu utworzono w Toruniu i Bydgoszczy oraz powiecie inowrocławskim, a także w mieście Włocławek i powiatach: aleksandrowskim, toruńskim, bydgoskim.



### Rycina 13. Zależność między intensywnością wsparcia a stopą bezrobocia w województwie

Stopa bezrobocia rejestrowanego wg powiatów województwa kujawsko-pomorskiego w 2014 roku

Liczba nowoutworzonych miejsc pracy wg miejsca realizacji projektu (stan na 31.12.2014 r.)



Źródło: Opracowanie własne Fundacji IDEA Rozwoju.

Z analizy adekwatności wsparcia udzielonego w ramach RPO WK-P wypływa wniosek wskazujący na konieczność pozyskiwania nowych inwestorów i skorelowanie kierunków kształcenia z rozwojem nowych technologii. Ogólnie stwierdzono także, że system edukacji zawodowej w województwie kujawsko-pomorskim (podobnie jak w całej Polsce), w niewystarczającym stopniu reaguje na dynamicznie zmieniające się potrzeby i oczekiwania pracodawców.

Przeprowadzone badanie potwierdziło, że realizacja projektów, dzięki wsparciu RPO WK-P 2007-2013 pozwoliła osiągnąć zakładane cele w obszarze zatrudnienia. Liczba miejsc pracy przekroczyła wyznaczoną wartość 3 200 etatów, a struktura nowych miejsc pracy okazała się adekwatna wobec kluczowych dla regionu wyzwań. Efekty Programu były jednak zbyt ograniczone, aby trwale rozwiązać problemy kujawsko-pomorskiego rynku pracy, który nadal cechuje się wysokim poziomem i niekorzystną strukturą bezrobocia oraz rosnącymi dysproporcjami pomiędzy największymi ośrodkami miejskimi a obszarami wiejskimi oraz między wschodnią i zachodnią częścią województwa.

Ważną rekomendacją dla przyszłej perspektywy jest położenie większej niż dotąd dbałości o synergię efektów interwencji o różnym charakterze (zwłaszcza projektów „miękkich” i instrumentów finansowych). Szczególnej uwagi wymaga dostosowanie do potrzeb rynku pracy systemu edukacji, szkolnictwa wyższego oraz kształcenia ustawicznego. Spełnienie tych warunków w opinii ewaluatorów pozwoli lepiej wykorzystać potencjał regionu oraz

stawić czoła zagrażającym jego rozwojowi trendom (np. niekorzystnym zjawiskom demograficznym).

Prowadząc wyżej opisane badanie i współpracując przy tym z wykonawcą, a także z komórkami dostarczającymi dane oraz zainteresowanymi wynikami badań, w jednostce ewaluacyjnej nasunęły się następujące refleksje:

- istnieje słabość baz danych – informacje gromadzone w KSI SIMIK (2007-13) są niewystarczające do kompleksowych badań ewaluacyjnych,
- spotykana jest różna postawa beneficjentów wobec oczekiwań ewaluatorów (co przekłada się na zwrotność informacji),
- konieczny jest duży wysiłek obu stron (zamawiającego i wykonawcy), aby wyniki badania były pełne i użyteczne,
- stała współpraca z wykonawcą powinna być wpisana jako dobra praktyka na każdym etapie realizacji badania,
- warto być otwartym na sugestie ekspertów w dziedzinie ewaluacji, gdyż może to spowodować uzyskanie dodatkowych, ciekawych wyników badania.

### Podsumowanie

Dotychczasowe, zdobyte przez JE doświadczenia w dziedzinie ewaluacji programu operacyjnego mogą być wykorzystane i stanowić pewne wyzwania w nowej perspektywie. Wśród nich warto wymienić:

- odpowiednie zaplanowanie badań ewaluacyjnych (dobre Plany Ewaluacji), przy respektowaniu wszystkich procedur i przepisów prawnych;
- wypracowanie dobrej koncepcji badania opartej na dostępnych danych;
- gromadzenie danych i tworzenie banków informacji dla celów ewaluacyjnych;
- aktualizowanie i ciągłe doskonalenie kryteriów wyboru Wykonawców (indywidualne podejście do każdego postępowania);
- kontynuowanie współpracy z GSE RPO WK-P oraz z IZ RPO i departamentami odpowiedzialnymi za wdrażanie programu;
- upowszechnianie wyników badań i wskazywanie na ich użyteczność, zwłaszcza wśród decydentów;
- podnoszenie wiedzy z zakresu ewaluacji – udział w szkoleniach oraz śledzenie nowości metodologicznych;
- wykorzystanie dobrych praktyk, przykładów na polu ewaluacyjnym z innych regionów i państw.

## Bibliografia

- Bienias S.**, *Przyszłość i rola ewaluacji w nowym okresie programowania*, [w:] *Ewaluacja programów operacyjnych, aspekty teoretyczne, metodologiczne i praktyczne*, praca zbiorowa, Urząd Marszałkowski Województwa Kujawsko-Pomorskiego, Toruń 2013
- Jąkałski J.** (red.), *Proces ewaluacji Polityki Spójności w Polsce 2004-2014*, opracowanie Krajowej Jednostki Ewaluacji, Ministerstwo Infrastruktury i Rozwoju, Warszawa 2014
- Matusiak M.**, *Wybrane problemy kształtowania polityki rozwoju na poziomie lokalnym*, [w:] *Monitorowanie i ewaluacja polityk publicznych – doświadczenia przeszłości wyzwaniem przyszłości*, praca zbiorowa, Urząd Marszałkowski Województwa Opolskiego, Opole 2015
- Olejniczak K.**, *System ewaluacji Polityki Spójności. Tezy do dyskusji*, [w:] *Monitorowanie i ewaluacja polityk publicznych – doświadczenia przeszłości wyzwaniem przyszłości*, praca zbiorowa, Urząd Marszałkowski Województwa Opolskiego, Opole 2015
- Plan ewaluacji dla RPO WK-P na lata 2007-2013*, Urząd Marszałkowski Województwa Kujawsko-Pomorskiego, Toruń 2008
- Plan Ewaluacji Regionalnego Programu Operacyjnego Województwa Kujawsko-Pomorskiego na lata 2014-2020*, Urząd Marszałkowski Województwa Kujawsko-Pomorskiego, Toruń 2015
- Rudnicka J.**, *Problemy ewaluacji na przykładzie badań prowadzonych przez Jednostkę Ewaluacyjną RPO w województwie kujawsko-pomorskim*, [w:] *Ewaluacja programów operacyjnych, aspekty teoretyczne, metodologiczne i praktyczne*, praca zbiorowa, Urząd Marszałkowski Województwa Kujawsko-Pomorskiego, Toruń 2013
- Wpływ projektów realizowanych w ramach RPO WK-P na lata 2007-2013 na zatrudnienie w województwie kujawsko-pomorskim*, raport z badania ewaluacyjnego przygotowany na zlecenie UM WK-P, Fundacja IDEA Rozwoju, Toruń 2015



### JOLANTA RUDNICKA

Absolwentka Uniwersytetu Mikołaja Kopernika w Toruniu – kierunek geografia. Pierwsze doświadczenie zawodowe zdobyła w zakresie planowania i zagospodarowania przestrzennego w Wojewódzkim Biurze Planowania Przestrzennego we Włocławku. Od 2004 r. jako pracownik samorządowy Urzędu Marszałkowskiego Województwa Kujawsko-Pomorskiego uczestniczyła m.in. w pracach nad aktualizacją strategii rozwoju województwa kujawsko-pomorskiego oraz programowaniem perspektywy 2007-2013 i 2014-2020. Od 2009 roku kieruje pracami Biura Ewaluacji a od 2015 również Wydziałem Analiz i Ewaluacji.







Na zdjęciu: Otwarcie konferencji  
*Paweł Peńsko (kierownik Biura Promocji Projektów Multimedialnych)*



Na zdjęciu: Otwarcie konferencji (fot. M. Kuras)  
*Zbigniew Ostrowski (Wicemarszałek Województwa Kujawsko-Pomorskiego)*





Na zdjęciu: Uczestnicy konferencji (fot. M. Kuras)



Na zdjęciu: Uczestnicy konferencji





Na zdjęciu: Uczestnicy konferencji



Na zdjęciu: Uczestnicy konferencji



Na zdjęciu: Prelegenci konferencji

Od lewej: Stanisław Bienias (Fundacja IDEA Rozwoju)

Tomasz Kot (Krajowa Jednostka Ewaluacji, Ministerstwo Rozwoju)



Na zdjęciu: Prelegenci konferencji

Przemysław Herman (Naczelnik Wydziału Ewaluacji w Departamencie Europejskiego Funduszu Społecznego w Ministerstwie Rozwoju)



Na zdjęciu: Uczestnicy konferencji



Na zdjęciu: Uczestnicy konferencji





Na zdjęciu: Prelegenci konferencji  
*Agnieszka Haber (Polska Agencja rozwoju Przedsiębiorczości)*



Na zdjęciu: Prelegenci konferencji  
*Małgorzata Wiśniewska (Zastępca Dyrektora Departamentu Spraw Społecznych i Zdrowia UM Województwa Kujawsko-Pomorskiego)*



Początek wdrażania programów operacyjnych perspektywy finansowej 2014-2020 to czas niezwykle ważny dla procesu ewaluacji. Procesu, który rozwija się i zmienia wraz z nowymi wyzwaniami stawianymi przez politykę spójności. Do tej pory ewaluacja skoncentrowana była na kwestiach wdrożeniowych, w mniejszym zaś stopniu dotyczyła oceny efektów realizacji wsparcia. Obecnie jej rola w procesie decyzyjnym jeszcze się zwiększa. Szczególnie ważne staje się bowiem zapewnienie użyteczności ewaluacji. Również podniesienie wiarygodności badań ewaluacyjnych stanowi istotne wyzwanie do zrealizowania w bieżącym okresie programowania. Wymagać to będzie mobilizacji i wysiłku ze strony regionów, by ewaluacja była wykonywana nie tylko w odpowiednim czasie, ale żeby była też skutecznie wykorzystywana podczas realizacji polityki regionalnej przez interesariuszy zaangażowanych w osiąganie założonych celów. Żeby tak się stało, należy także zintensyfikować poszukiwania nowych metod prezentacji wyników badań, a sposób ich pokazywania dostosować do potrzeb i oczekiwań ich odbiorców. Dopiero dzięki temu ewaluacja stanie się skutecznym narzędziem polityki rozwoju.

**Zbigniew Ostrowski**  
**Wicemarszałek Województwa**  
**Kujawsko-Pomorskiego**

**Urząd Marszałkowski**  
**Województwa Kujawsko-Pomorskiego**  
**w Toruniu**  
Plac Teatralny 2  
87-100 Toruń

Departament Rozwoju Regionalnego  
Wydział Analiz i Ewaluacji  
Biuro Ewaluacji

tel. 56 62 18 653

[www.mojregion.eu](http://www.mojregion.eu)

ISBN: 978-83-65206-00-8



**Egzemplarz bezpłatny**

Publikacja jest finansowana ze środków Europejskiego Funduszu Społecznego w ramach Regionalnego Programu Operacyjnego Województwa Kujawsko-Pomorskiego na lata 2014-2020 oraz budżetu województwa kujawsko-pomorskiego