

Tomasz Kupiec

Akademia Leona Koźmińskiego w Warszawie, ul. Jagiellońska 57/59, 03-301 Warszawa;
e-mail: t.kupiec@evaluation.pl

UŻYTECZNOŚĆ EWALUACJI JAKO NARZĘDZIA ZARZĄDZANIA REGIONALNYMI PROGRAMAMI OPERACYJNYMI¹

Streszczenie: Artykuł prezentuje wyniki badania, którego celem była ocena użyteczności ewaluacji realizowanych na potrzeby regionalnych programów operacyjnych 2007–2013 (RPO). Badanymi elementami, składającymi się na użyteczność ewaluacji, były: jakość, adekwatność i wiarygodność wniosków wynikających z ewaluacji. Badanie zrealizowano w formie analizy dokumentów na reprezentatywnej próbie 71 raportów ewaluacyjnych ukończonych w latach 2008–2012. Wyniki badania pokazują, że wiele raportów nie dostarcza istotnych rekomendacji oraz nie udziela odpowiedzi na kluczowe pytania badawcze. Najczęściej w trakcie ewaluacji nie udaje się odpowiedzieć na pytanie o wpływ RPO na rozwój społeczno-gospodarczy województw. Zaobserwowana niska użyteczność dokonywanych ocen stanowi podstawę do sformułowania wniosków o negatywnych trendach w rozwoju systemu ewaluacji RPO, obejmujących nastawienie systemu na dostarczanie prostych informacji i „produkcję raportów”, a nie zaspokajanie faktycznego popytu na wiedzę.

Słowa kluczowe: użyteczność ewaluacji, system ewaluacji, regionalny program operacyjny, polityka spójności.

THE USABILITY OF EVALUATION AS A TOOL FOR MANAGING REGIONAL OPERATIONAL PROGRAMMES

Abstract: The paper presents an analysis of usability of regional operational programmes 2007–2013 (ROP) evaluations. The elements of usability under study are: quality, relevance, and credibility of the evaluation assumptions. The study consists in desk research of a representative sample (n = 71) of the evaluation reports completed between 2008 and 2012. The results show that many reports do not contain important recommendations and do not answer key research questions. In most cases, the evaluations fail to supply an answer on the impact of ROP on the socio-economic development of regions. The limited usefulness of evaluation reports leads to conclusions about negative trends in the development of the ROP evaluation system, which is focused on providing simple information, reports production, and does not respond to actual demand for knowledge.

Keywords: evaluation usability, evaluation system, regional operational programme, cohesion policy.

¹ Prezentowana analiza została wykonana w ramach projektu realizowanego przez Akademię Leona Koźmińskiego przy wsparciu finansowym z inicjatywy Santander Universidades.

W ogólnym ujęciu ewaluacja definiowana jest najczęściej jako zestaw działań stosowanych do określenia jakości, ustalenia wartości czy znaczenia danego obiektu, programu (Scriven 2007, s. 1; Farrell et al. 2002, s. 8; Rossi et al. 2004, s. 2). Jednym z charakterystycznych elementów, wskazywanych powszechnie jako dodatkowo wyróżniające ewaluację (w tym ewaluację interwencji publicznych), jest jej użyteczność. Ewaluacja powinna wspierać podejmowanie decyzji odnośnie do obecnych i przyszłych działań publicznych poprzez dostarczanie wiedzy oraz danych niezbędnych do ich podjęcia (Patton 1997; Royse et al. 2001); powinna pomagać decydującym w dokonywaniu mądrych wyborów dotyczących programowania interwencji (Weiss, za: Alkin 1990, s. 83) i tym samym przyczyniać się do poprawy warunków życia społeczeństwa (Rossi et al. 2004). Poprawa dobrobytu społecznego jest nazywana podstawowym i ostatecznym celem ewaluacji (Henry, Mark 2003; Mark, Henry 2004). Jak ujęli to Stufflebeam i Shinkfield (1985, za: Clarke 1999, s. 2), najważniejszym celem ewaluacji jest nie dowieść, a poprawić (ang. *not to prove but to improve*).

Racjonalnie działająca instytucja powinna, w teorii, dokonywać ewaluacji jedynie wtedy, gdy służą one podejmowaniu decyzji, a korzyści wynikające z lepszych (wspartych ewaluacją) decyzji przewyższają koszty realizacji badań (Dahler-Larsen 1998). Rzeczywistość jest jednak zgoła inna i praktyce ewaluacji interwencji publicznych od początku towarzyszyło, potwierdzone licznymi badaniami empirycznymi, rozczarowanie znikomym wpływem dokonywanych ocen na podejmowane decyzje i kształt interwencji (Olejniczak 2008; Alkin, Coyle 1988; Henry 2000). W odpowiedzi stworzono bogatą literaturę na temat wykorzystania ewaluacji, wskazującą m.in. liczne propozycje czynników i całościowych modeli wykorzystania ewaluacji².

W kontekście przedmiotu niniejszego artykułu wartym przytoczenia jest opracowanie Leeuwa i Furubo (2008), którzy postawili kilka hipotez wyjaśniających obserwowaną pozorną nieracjonalność, tzn. rozwój w ramach administracji publicznej systemów ewaluacji (m.in. ich kadr, zasobów) i realizację znacznej liczby badań, mimo że ich wpływ na kształt polityk jest niewielki. Zdaniem autorów:

1. Systemy ewaluacji produkują w przeważającej mierze rutynowe, proste informacje wykorzystywane jedynie w bieżących działaniach wdrożeniowych.
2. Przesłanką rozwoju systemu ewaluacji jest często nie chęć prowadzenia polityki opartej na wiedzy czy dowodach, ale jedynie stworzenie wrażenia (w oczach społeczeństwa), że taką politykę się prowadzi.
3. System ewaluacji zorientowany jest przede wszystkim na własny rozwój. Osoby zaangażowane w system ewaluacji realizują swoje cele, a nie te, dla których system został powołany. Dążą do tego, by uzasadnić swoje istnienie, by system się rozwijał, by rósł jego prestiż, znaczenie, zasoby, niezależnie od faktycznych potrzeb.

² Jako przykłady rozbudowanych katalogów czynników wykorzystania ewaluacji wskazać można m.in. opracowania Alkina (1985), Cousins i Leithwood (1986) czy Saunders (2012). Całościowe modele wykorzystania ewaluacji zaproponowali m.in.: Johnson (1998), Cousins (2003), Mark i Henry (2004).

Polska wydaje się krajem, w którym ryzyko wystąpienia negatywnych zjawisk w procesie rozwoju systemu ewaluacji jest relatywnie wysokie. Nie należy bowiem do państw o długiej krajowej tradycji ewaluacji³, rozwiniętej samodzielnie w odpowiedzi na zidentyfikowane potrzeby i przekonanie, że ewaluacja może je zaspokoić. Ewaluacja zawitała do Polski w zasadzie dopiero po przystąpieniu do UE, jako wymóg formalny, towarzyszący realizacji polityki spójności⁴.

Zaszczepiona za pośrednictwem wymogów związanych z wdrażaniem funduszy strukturalnych praktyka ewaluacji – mimo braku tradycji – zaczęła się gwałtownie rozwijać⁵. Liczba zleczanych ewaluacji z pięciu w 2004 r. wzrosła do ok. 50 rocznie w latach 2006–2007 i ponad 120 rocznie po roku 2007 (National Evaluation Unit; Skórska 2011, s. 201). Roczny budżet na badania, wynoszący ok. 400 tys. zł w 2004 r. i 2,2 mln zł w roku 2006, w ostatnich latach osiągnął wartość ponad 25 mln zł (Bienias et al. 2009, s. 157).

Przedstawiciele administracji podkreślają często, np. w licznych publikacjach Ministerstwa Rozwoju Regionalnego⁶, rosnący potencjał systemu ewaluacji i zwiększającą się liczbę realizowanych badań. Ów wzrost, utożsamiany z rozwojem kultury ewaluacji, jest również zauważany i pozytywnie odbierany poza granicami kraju. Komisja Europejska (KE) postanowiła w ramach ewaluacji ex-post perspektywy 2000–2006 polityki spójności przygotować osobne studium przypadku opisujące dobrą praktykę budowania systemu ewaluacji w Polsce (Olejniczak 2009), zaś Bachtler (2011) stwierdza, że ewaluacja polityki regionalnej rozwija się w Polsce szybciej niż w pozostałych państwach Europy Środkowo-Wschodniej.

Odwołując się do koncepcji modeli logicznych teorii programu⁷, można jednak odnieść wrażenie, że większość dyskusji koncentruje się na wkładach i produktach systemu ewaluacji. Nie towarzyszy im natomiast wystarczająco głęboka refleksja odnośnie do jakości i użyteczności ewaluacji oraz jej faktycznego wykorzystania. Prezentowane poniżej badanie jest próbą takiej refleksji.

Opisany powyżej ilościowy wzrost aktywności ewaluacyjnej w Polsce jest w dużej mierze zasługą samorządów województw, które od 2007 r. zarządzają samodzielnie programami operacyjnymi i odpowiadają m.in. za ich ewaluację. Realizowane w ramach polityki spójności perspektywy 2007–2013 regionalne programy operacyjne stanowią kluczowy (i zdecydowanie największy w sensie alokacji finansowej) instrument prowadzenia polityki rozwoju na poziomie województwa, wpływający w zauważalny sposób na jakość życia jego mieszkańców. W obliczu niebagatelnej roli RPO kluczowego znaczenia nabiera również ewalu-

³ Takich jak USA, a w Europie np. Wielka Brytania, Niemcy, Szwecja (Bachtler 2001, Toulemonde 2000).

⁴ Jak pisał w 2005 r. Olejniczak, „świadomość roli ewaluacji budzi się bardzo powoli w naszej administracji krajowej, czy raczej jest budzona wymogami funduszy”.

⁵ Choć jedynie w obszarze polityki spójności.

⁶ M.in. przytoczonych już w poprzednich akapitach. Ministerstwo Rozwoju Regionalnego stanowi centralną instytucję systemu ewaluacji polityki spójności.

⁷ Modele te wyjaśniają mechanizmy zmiany wywołanej przez podjętą interwencję, poczynając od poniesionych nakładów, poprzez realizowane działania, ich bezpośrednie produkty, aż po efekty zwane często oddziaływaniami.

acja tych programów i jej wykorzystanie. Użyteczność ewaluacji jako narzędzia wspierającego zarządzanie strategiczne i operacyjne rozwojem województwa wpływa bowiem na skuteczność realizacji celów RPO i ostatecznie – jak już wskazano powyżej – na poprawę dobrobytu społecznego.

Prezentowane badanie można potraktować jako element rozliczenia dotychczasowych działań w zakresie ewaluacji na poziomie województw oraz, przede wszystkim, jako wstęp do dyskusji o poprawie użyteczności ewaluacji w nadchodzącej perspektywie 2014–2020. W okresie tym samorządy województw będą dysponowały jeszcze większą pulą środków a trafność (wspieranych ewaluacją) decyzji w zakresie ich wykorzystania w jeszcze większym stopniu odcisnie się na poziomie rozwoju regionów.

Cel, zakres i metodologia badania

Niniejszy artykuł prezentuje wyniki badania, którego celem było określenie użyteczności produktów dostarczanych przez system ewaluacji regionalnych programów operacyjnych⁸, czyli raportów ewaluacyjnych.

Użyteczność badań stanowi jedną z dwóch podstawowych kategorii czynników wskazywanych w modelach wykorzystania ewaluacji⁹. Składają się na nią m.in.: jakość, adekwatność i wiarygodność wniosków (Saunders 2012; Cousins, Leithwood 1986; Owen 2007). W ramach niniejszego badania elementy te, na potrzeby ich oceny, zdefiniowano w następujący sposób:

- adekwatność ustaleń: Czy przedstawione w raporcie rekomendacje spełniają potrzeby informacyjne zamawiającego? Czy są spójne z celami badania?
- wiarygodność ustaleń: Czy wnioski, w oparciu o które sformułowano rekomendacje, są wiarygodne?
- jakość ustaleń: Czy rekomendacje są czytelne (a nie niezrozumiałe bądź nierealne), czy dają się wdrożyć, są istotne, prowadzą do znaczących zmian w zakresie wdrażania lub kształtu programu operacyjnego?

Dodatkowo w badaniu odniesiono się do kwestii kompletności raportu (czyli do tego, czy ewaluacja odpowiada na wszystkie kluczowe pytania badawcze), która bezpośrednio wpływa na powyższe elementy.

Zasadnicze znaczenie kwestii użyteczności badań wynika z faktu, iż można ją traktować jako test instytucji bądź jej systemu ewaluacji. Brak użyteczności jest wskazówką, że system ten nie działa, że instytucja nie ma potencjału do realizacji wartościowych badań bądź nie jest nastawiona na faktyczne wykorzystanie wyników. Dlatego też ewentualna niska ocena użyteczności raportów ewaluacyjnych stanowić będzie argument wspierający część hipotez Furubo i Leeuwa w odniesieniu do systemu ewaluacji RPO. Przede wszystkim potwierdzi, iż system ten skupia się na samym procesie produkcji informacji, a nie na dostarczaniu wiedzy mogącej wspierać proces podejmowania decyzji. Tym samym niska użyteczność

⁸ Stanowi on ważne ogniwo systemu ewaluacji polityki spójności w Polsce.

⁹ Obok kontekstu organizacyjnego, czyli potencjału instytucji, jej struktury, zwyczajów, praktyk wykorzystania badań.

raportów powinna stanowić przesłankę podjęcia pogłębionych badań funkcjonowania systemu ewaluacji RPO i szerzej polityki spójności.

Zakres prezentowanego badania obejmuje raporty ewaluacyjne zamawiane przez jednostki ewaluacyjne w ramach instytucji zarządzających (IZ RPO) 2007–2013. Do końca 2012 r. zrealizowano 236 ewaluacji wszystkich RPO. Analizą objęto próbę $n=71$ (30%) badań. Dobór próby miał charakter losowo-warstwowy, a celem było zachowanie różnorodności względem: momentu realizacji badania, tematyki, instytucji zamawiającej, wykonawcy. Podstawową metodą stosowaną w badaniu była analiza dokumentów, obejmująca dla każdego badania szczegółowy opis przedmiotu zamówienia (SOPZ), raport metodologiczny i przede wszystkim raport końcowy.

Wyniki badania

Uwagi wprowadzające

Badania ewaluacyjne zlecane są przez usytuowane w strukturze IZ RPO komórki zwane jednostkami ewaluacyjnymi, zaś ich odbiorcami są najczęściej inne komórki IZ, zaangażowane w proces zarządzania i wdrażania RPO. Badania realizowane są przeważnie (95%) przez zewnętrznych wykonawców wybieranych w trybie zamówienia publicznego. Wstępny zakres pytań i metod badawczych określany jest przez IZ RPO i stanowi zwykle wypadkową potrzeb różnych komórek merytorycznych danej instytucji oraz pomysłów jednostki ewaluacyjnej. Następnie katalog pytań i metod jest najczęściej rozszerzany propozycjami zawartymi w ofertach wykonawców i odzwierciedlającymi ich spojrzenie i pomysł na badanie.

Każde badanie ewaluacyjne powinno kończyć się sformułowaniem rekomendacji zmian sposobu wdrażania programu czy jego zakresu wsparcia. Rekomendacje formułowane są w raporcie wg określonego wytycznymi MRR wzoru – tabeli, w której konieczne jest wskazanie m.in. wniosków stanowiących podstawę rekomendacji, realnego sposobu jej wdrożenia, instytucji odpowiedzialnych za wdrożenie i terminu, w którym powinno ono nastąpić. Tabela rekomendacji jest kluczowym elementem raportu ewaluacyjnego, stanowiącym wyciąg najważniejszych ustaleń badania i przepis na ich wykorzystanie w procesie zarządzania RPO.

Adekwatność i wiarygodność ustaleń ewaluacji

Jedynie w 20% analizowanych ewaluacji RPO wszystkie bądź większość przedstawionych przez wykonawcę rekomendacji, mających służyć poprawie programu operacyjnego, można uznać za zgodne z celami ewaluacji i potrzebami informacyjnymi zamawiającego. W ok. 50% ewaluacji tylko część rekomendacji jest zgodna z oczekiwaniami odbiorców, zaś w 30% badań rekomendacje są w całości niezgodne z celami badawczymi. Typowy przykład braku zgodności pomiędzy celami badania a proponowanymi rekomendacjami stanowi sytuacja,

w której celem ewaluacji jest określenie efektywności poszczególnych typów interwencji i zaproponowanie pożądanych kierunków wsparcia na przyszłość, zaś rekomendacje koncentrują się na prostych technicznych kwestiach, związanych z procesem wdrażania programu operacyjnego, a nie na zakresie wsparcia.

Udział rekomendacji wiarygodnych, tzn. wynikających z wniosków, które zostały sformułowane w ramach danego badania, wynosi średnio 77%. W co szóstym badaniu udział ten wynosi jednak poniżej 50%. W większości wypadków brak wiarygodności rekomendacji wynika z faktu, że jest ona w oczywisty sposób niespójna z wnioskiem wskazanym jako jej uzasadnienie w tabeli rekomendacji. Przykłady takich sytuacji zawarto w tabeli 1.

Tab. 1. Przykłady rekomendacji, które nie są powiązane z wnioskami z badania

Wniosek	Rekomendacja	Komentarz
Departament Zarządzania RPO czerpie pomysły oraz wymienia doświadczenia w zakresie działań informacyjno-promocyjnych podczas różnego rodzaju spotkań.	Zaleca się branie udziału w spotkaniach służących wymianie doświadczeń oraz zintensyfikowanie działań zmierzających do kreowania innowacyjnych rozwiązań informacyjno-promocyjnych.	Nie odnosząc się w żaden sposób do skuteczności wymiany doświadczeń (pozytywnych efektów), zaleca się dalsze jej prowadzenie.
Wysoki poziom zainteresowania projektami z zakresu oczyszczania ścieków.	Kontynuowanie bieżącej polityki z zakresu oczyszczania ścieków.	Celem badania było wskazanie efektywności wsparcia. To właśnie wykazana efektywność powinna uzasadniać kontynuację wsparcia, a nie sam fakt zainteresowania ze strony potencjalnych beneficjentów.
Niski poziom zainteresowania aplikowaniem o środki z zakresu gospodarki odpadami komunalnymi.	Zapewnienie większej ilości środków na realizację projektów w zakresie gospodarki odpadami komunalnymi.	zupełnie niezrozumiała propozycja zwiększenia alokacji, w sytuacji gdy obecne środki nie są wykorzystywane.

Źródło: opracowanie własne na podstawie analizy dokumentów.

Równie często spotkać można sytuację, w której rekomendacja jest spójna z wnioskiem wskazanym w tabeli, jednakże sam wniosek sformułowany jest w oderwaniu od ustaleń poczynionych w procesie badawczym i w raporcie nie można znaleźć dla niego żadnego potwierdzenia czy uzasadnienia.

Jakość ustaleń ewaluacji

Średni udział rekomendacji, które można ocenić jako czytelne i możliwe do wdrożenia wynosi 68%. Jednak w co piątej ewaluacji ponad połowa rekomendacji jest niezrozumiała, zbyt ogólna bądź niewykonalna, czego przykłady zaprezentowano w tabeli 2.

Tab. 2. Przykłady rekomendacji niezrozumiałych bądź niewykonalnych.

Rekomendacja	Komentarz
Opracowanie procedury weryfikacji	<i>Rekomendacja sformułowana w odpowiedzi na problem „brak procedury weryfikacji komplementarności w praktyce”. Nie wskazano żadnych szczegółów odnośnie do założeń samej procedury.</i>
Poszukiwanie odpowiednich kanałów komunikacji z beneficjentami RPO WO 2007–2013	<i>Rekomendacja sformułowana w odpowiedzi na problem niedoinformowania beneficjentów w zakresie szkoleń. Brakuje wskazania, jakie kanały komunikacji będą odpowiednie dla poszczególnych grup odbiorców.</i>
Zmiana jednostki wskaźnika	<i>Rekomendacja skierowana jest do IZ, a dotyczy wskaźnika obowiązkowego, zawartego w Krajowym Systemie Informatycznym. IZ nie może dokonać jego modyfikacji, skutkowałoby to bowiem niespójnością z systemem krajowym.</i>
Cykl szkoleń motywacyjnych i informacyjnych dla JST o ich znaczeniu w uzyskaniu efektu synergii dzięki komplementarności lokalnych projektów	<i>Rekomendacja sformułowana w odpowiedzi na problem bardzo ograniczonej roli organów JST w koordynowaniu, monitorowaniu, inicjowaniu działań służących realizacji celów strategicznych w gminie. Wskazany problem mocno wykracza poza tematykę RPO i nie wydaje się zasadne, by to IZ RPO miała szkolić i motywować JST w zakresie realizacji ich zadań.</i>
Wnioski o dofinansowanie powinny być oceniane jakościowo przez wybitnych specjalistów i ekspertów w danej dziedzinie	<i>Wdrożenie tej rekomendacji niewątpliwie podniosłoby jakość oceny projektów. Jest to jednakże jedynie nierealne życzenie. Jak pokazuje praktyka, nie tylko wybitni, ale po prostu specjaliści nie są zainteresowani angażowaniem się w ocenę projektów.</i>

Źródło: opracowanie własne na podstawie analizy dokumentów.

W toku analizy za rekomendacje istotne uznawano te, które nie tylko są powiązane z wnioskami z badania, czytelne i możliwe do wdrożenia, ale posiadają dodatkowo potencjał do wywołania znaczących pozytywnych zmian, rozwiązania zidentyfikowanego problemu (w odróżnieniu od rekomendacji oczywistych, nieodkrywczych bądź dotyczących drobnych, technicznych zmian).

Średni udział tak rozumianych rekomendacji istotnych to 37%, przy czym większość z nich (30%) to rekomendacje operacyjne, dotyczące usprawniania procesu wdrażania, rozwiązywania bieżących problemów związanych z wydatkowaniem środków. Jedynie 7% to istotne rekomendacje dotyczące zmian strategicznych – kierunków wsparcia, zakresu działań i priorytetów, wspieranych typów projektów i przeznaczonych na nie alokacji finansowych. Po uwzględnieniu rozróżnienia przedmiotów ewaluacji (na te dotyczące procesu wdrażania RPO, efektów RPO oraz łączących badanie procesów i efektów) okazuje się, że istotne rekomendacje wdrożeniowe stanowią połowę rekomendacji w ewaluacjach procesów (i 19% w ewaluacjach procesów i efektów), zaś istotne rekomendacje strategiczne stanowią średnio 17% w badaniach efektów (i 5% w ewaluacjach procesów i efektów). Wśród ewaluacji, których przedmiotem są procesy bądź procesy i efekty, aż 62% stanowią takie, w których mniej niż połowę rekomendacji można uznać za istotną wdrożeniowo, a w 20% badań nie ma ani jednej takiej rekomendacji. Wśród ewaluacji efektów oraz efektów i procesów aż 89% zawie-

ra mniej niż połowę istotnych strategicznie rekomendacji, a w 59% ewaluacji nie ma ani jednej takiej rekomendacji.

Przykłady rekomendacji, które nie posiadają potencjału do wywołania zauważalnej zmiany, znaleźć można m.in. w ewaluacjach wpływu RPO na strategię lizbońską. Mimo iż badanie wykazało, że wpływ ten zależy głównie od zakresu wsparcia RPO, w rekomendacjach odnoszono się do kształtu wniosku o dofinansowanie i tego, jak w nim ukazać spójność projektu ze strategią. Zabawnym przykładem innego typu rekomendacji, które w toku analiz uznano za nieistotne, jest konkluzja przedstawiona w ramach ewaluacji systemu wdrażania i zarządzania RPO, iż system będzie działał lepiej, jeśli pracownikom IZ zapewni się więcej miejsc parkingowych a pokoje będą bardziej przestronne. Jako nieistotne oceniano również rekomendacje słuszne w swojej treści, ale zupełnie nieodkrywcze, oczywiste, powtarzające powszechnie akceptowane już poglądy, np.:

- Podczas prac nad nowym programem operacyjnym należy wykorzystać wyniki dotychczasowych badań ewaluacyjnych.
- Przy wyborze projektów należy ze szczególną starannością ocenić ich potencjalny wpływ na osiągnięcie celów określonych w RPO.
- Wybrane projekty powinny spełniać najbardziej żywotne potrzeby społeczno-gospodarcze regionu.

Kilkukrotnie zdarzało się, że wykonawca, nie odpowiadając na zadane pytania badawcze, w rekomendacjach wskazywał, iż powinny być one przedmiotem innego badania ewaluacyjnego. Inną grupą rekomendacji nieistotnych są propozycje będące wiernymi cytatami wypowiedzi przedstawicieli IZ RPO. Ewaluacje, w których wykonawca mimo prowadzenia własnej analizy dokumentów, realizacji badania ankietowego, formułując rekomendacje, jedynie (czy przede wszystkim) przytacza opinie pracowników IZ RPO, należy ocenić jako nieużyteczne, niedostarczające odbiorcom żadnej nowej wiedzy ani rozwiązań. Użytkownicy badania – osoby, które zadały pytania – dostają bowiem jedynie odpowiedzi, które już wcześniej знаły i same ich sobie udzieliły. Jeszcze jednym przykładem rekomendacji o niewielkiej wartości może być wskazywanie w odpowiedzi na pytania o efektywny sposób alokacji pozostałych środków tylko tych projektów, które IZ już i tak wstępnie wcześniej zakwalifikowała do dofinansowania.

Za potencjalnie mniej istotne uznano również rekomendacje, których termin wdrożenia określono na przyszły okres wdrożeniowy¹⁰. Dotyczy to w szczególności rekomendacji mocno osadzonych w kontekście obecnego programu operacyjnego, na których wdrożenie jest jednak za późno, dlatego z konieczności, by w ogóle zwrócić uwagę na daną kwestię, zaleca się ich wykorzystanie po 2013 r. Średni udział rekomendacji z terminem wdrożenia jeszcze w perspektywie 2007–2013 wynosi 43%. W ponad połowie badań rekomendacji takich jest

¹⁰ Jeżeli rekomendacja nie zostaje wdrożona w krótkim czasie po zakończeniu ewaluacji, kiedy wnioski z danego badania stanowią jeszcze przedmiot zainteresowania i dyskusji, to szanse jej wykorzystania drastycznie maleją.

mniej niż 30% (w tym w co czwartej ewaluacji procesów¹¹), a w 41% nie ma ani jednej (w tym w 14% ewaluacji procesów).

Podobnie postrzegać można rekomendacje skierowane do instytucji niewchodzących w skład systemu wdrażania badanego RPO¹². Ich udział jest jednak niewielki, średnio wynosi 6% i tylko w nielicznych przypadkach (4%) przekracza 30%.

Kompletność raportów ewaluacyjnych

Duży udział ewaluacji, które nie prowadzą do istotnych rekomendacji, wynika m.in. z tego, że w wypadku większości analizowanych badań (72%) odpowiedzi udało się udzielić nie na wszystkie kluczowe pytania postawione przez zamawiających. Pod tym względem trudniejsze w realizacji okazują się ewaluacje efektów – w 85% tego typu badań brakuje odpowiedzi na niektóre pytania, jednak także w dużej części ewaluacji procesów (63%) brak rozstrzygnięcia niektórych zagadnień badawczych jest problemem.

Brak odpowiedzi na pytania badawcze czasami w oczywisty sposób wynika z raportu końcowego. Najczęściej jednak wykonawca pozoruje odpowiedź, podejmując w raporcie rozważania nie na temat, bądź formułuje wnioski tak rażąco niewiarygodne (ze względu na błędne źródło, metodę pozyskania danych, brak interpretacji), że trzeba uznać, iż odpowiedź na postawione pytania ewaluacyjne nie została udzielona. Zjawisko to obrazują poniższe przykłady:

- Zamiast analizy wpływu wsparcia RPO w dziedzinie transportu na gospodarkę województwa w raporcie zawarto jedynie ogólne wywody na temat znaczenia transportu dla rynku pracy i turystyki (bez jakiegokolwiek odniesienia do projektów dofinansowanych w RPO).
- Zamiast analizy wpływu RPO na realizację celów strategii w raporcie wskazano jedynie, czy projekty są spójne, możliwe do powiązania z celami strategii. W odniesieniu do celu „Rozwój szkolnictwa zawodowego zgodnie z potrzebami gospodarki” stwierdzono na przykład, że kilka projektów wiąże się z promocją i poprawą kształcenia zawodowego. Natomiast nie ma informacji o jakości tych projektów, trafności, adekwatności skali wsparcia w RPO i odpowiedzi na pytanie, czy te projekty zaowocowały zauważalną zmianą w poziomie realizacji celów strategii.
- W odpowiedzi na pytanie o komplementarność projektów priorytetu „środowiskowego” RPO z projektami realizowanymi w ramach innych programów wykonawca podał jedynie po trzy, cztery przykłady projektów komplementar-

¹¹ Można przyjąć, że to właśnie z ewaluacji procesów wynikają wnioski i rekomendacje raczej adekwatne dla badanego programu operacyjnego, choć oczywiście są od tej reguły wyjątki. W wypadku ewaluacji efektów również formułowany jest szereg rekomendacji wdrożeniowych, o czym mowa powyżej.

¹² Wyzwaniem dla jednostek ewaluacyjnych jest już zarządzanie procesem wdrażania rekomendacji we własnej instytucji. Trudno wyobrazić sobie, by były one w stanie skutecznie zainteresować daną ewaluacją RPO instytucje spoza systemu realizacji NSRO, ze szczebla krajowego czy unijnego.

nych dla każdego z programów. Nie odniesiono się natomiast w żaden sposób do skali zjawiska w całej populacji¹³. Nie wiadomo, jaki jest udział projektów komplementarnych w każdym z programów, nie wiadomo, czy skala zjawiska jest zadowalająca, odpowiednia dla tego obszaru wsparcia.

- Zamawiający oczekiwał analizy wpływu wdrożenia rekomendacji z poprzedniej ewaluacji na sprawność działania systemu zarządzania i wdrażania RPO, w tym m.in. wskazania ewentualnych negatywnych skutków rekomendacji oraz wprowadzonych do systemu zmian niewynikających z rekomendacji. Wykonawca wskazał rekomendacje wdrożone i niewdrożone, czasami odnosząc się do ich zasadności. To odniesienie się ma jednak charakter opinii nieopartej w żaden sposób analizą faktycznych zmian efektywności systemu. Zupełnie pominięto kwestię negatywnych skutków czy zmian w systemie niewynikających z rekomendacji.
- Zamawiający oczekiwał identyfikacji funkcji metropolitalnych miast stołecznych regionu, analizy wpływu RPO na ich rozwój oraz wskazania funkcji wymagających dalszego wsparcia i narzędzi ich wspierania. Wykonawca stworzył szeroki katalog potencjalnych funkcji metropolitalnych, każdą z nich omówił, wykorzystując dane ze statystyk publicznych, ale nie dokonał żadnej oceny, nie wskazał funkcji wiodących dla badanych miast, funkcji, które mogą stanowić przewagę konkurencyjną i powinny być rozwijane.

W raportach pojawiają się jednak również stwierdzenia, z których wprost wynika (choć zapewne nie było to intencją autorów), że w toku badania nie udało się uzyskać odpowiedzi na nurtujące zamawiającego pytania:

- „Odpowiedź na pytanie o kierunki inwestycji ze środków UE w ramach RPO w służbę zdrowia w kolejnej perspektywie finansowania ma swoje ograniczenia, gdyż brakuje diagnozy obecnego stanu wyposażenia placówek oraz pełnej diagnozy potrzeb, co pozwoliłoby zrationalizować i w większym stopniu ukierunkować wydatki ze środków UE, a także z krajowych środków publicznych” – celem ewaluacji było właśnie określenie kierunków interwencji, niezrozumiałe jest więc, dlaczego potrzebna diagnoza nie została w ramach badania przeprowadzona i jaki był sens ewaluacji nierozstrzygującej podstawowego problemu badawczego.
- „Wszelkie działania prowadzone w województwie w obszarze edukacji wymagają szczegółowej diagnozy i prognozy stanu edukacji wraz z określeniem dostosowania jej do zmieniających się potrzeb rynku pracy” – podobnie jak powyżej wspomniana diagnoza powinna być elementem ewaluacji.

Najczęściej spotykanym problemem – pytaniem pozostawionym bez odpowiedzi w raporcie – jest kwestia wpływu realizacji projektów wspartych w RPO na rozwój społeczno-gospodarczy województwa, szacowanie oddziaływań, szerszego wpływu, realizacja potrzeb społecznych. Określenie wpływu projektów na rozwój społeczno-gospodarczy regionu utrudnia wiele czynników. Po pierwsze, w wielu wypadkach w momencie przeprowadzania ewaluacji brakuje wystarczającej liczby na tyle dawno zakończonych projektów, żeby ich efektów

¹³ W priorytecie podpisano umowy na realizację ponad 90 projektów.

można już było oczekiwać i poszukiwać. Po drugie, skala wsparcia RPO jest raczej nieznaczną w porównaniu z pozostałymi środkami unijnymi, wszystkimi środkami publicznymi czy inwestycjami sektora prywatnego – co sprawia, że wyodrębnienie ze zmian dokonujących się w otoczeniu tej części, za którą odpowiada RPO, jest bardzo trudne. Po trzecie, dostępny budżet i metody zwykle uniemożliwiają uchwycenie kwestii oddziaływania RPO. W praktyce zatem wykonawcy uciekają się do różnych półśrodków, które jednak można nazwać jedynie imitacją odpowiedzi. Cytowane są zapisy wniosków o dofinansowanie, w których wnioskodawcy zapewniają o pozytywnym wpływie swoich projektów. Rzadziej podawane są dane statystyczne obrazujące zmiany sytuacji społeczno-gospodarczej w regionie, powiatach czy gminach, a ewentualne zmiany na lepsze są utożsamiane z wpływem RPO, jeśli na danym obszarze był realizowany projekt. Najczęściej w raportach końcowych zamiast oczekiwanej odpowiedzi na pytanie, „jaki jest wpływ projektów RPO na rozwój społeczno-gospodarczy regionu”, prezentuje się odpowiedź na pytanie, „jaka jest opinia beneficjentów o wpływie realizacji ich własnych projektów na rozwój społeczno-gospodarczy”. Opinie te najczęściej nie są w żaden sposób weryfikowane, za to prawie zawsze są pozytywne.

Brakowi faktycznej odpowiedzi na pytania towarzyszą często w raportach zaskakujące stwierdzenia i wnioski, czasami komiczne, które rodzą wątpliwości co do kompetencji autorów, wiarygodności ustaleń i sensu realizacji danej ewaluacji. W analizowanych raportach natrafiono m.in. na następujące sformułowania:

- „Każdy projekt turystyczny wpływa na wydłużenie sezonu turystycznego”.
- „Wyniki badań empirycznych podkreślają bezsprzeczny fakt, że niemożliwe jest osiągnięcie sukcesu aplikacyjnego bez poprawnie napisanego wniosku spełniającego kryteria formalne i merytoryczne”.
- W badaniu działań informacyjno-promocyjnych: „Informacje do młodzieży powinny trafiać w postaci gadżetów”.
- „Realizowane projekty stanowią elementy komplementarnego systemu. Z uwagi na duże potrzeby regionu, każdy nawet mały projekt stanowi element tego systemu”.

Bezpośrednią przyczyną opisanego powyżej problemu (braku odpowiedzi na kluczowe pytania badawcze) wydają się być m.in. nierealne oczekiwania zamawiających przejawiające się w zdecydowanie zbyt szerokich zakresach realizowanych badań. Niczym wyjątkowym nie jest próba odpowiedzi na 40 pytań badawczych w jednym raporcie, a zdarzają się i takie ewaluacje, w których pytań jest ponad 100. Nawet gdy pytań badawczych jest znacznie mniej, niektóre z nich bywają na tyle rozległe i skomplikowane, że jedno mogłoby z powodzeniem stanowić samodzielny przedmiot badania o typowych dla ewaluacji RPO budżecie i czasie na realizację. Częstokroć prosta suma pytań badawczych nie oddaje faktycznego zakresu badania, SOPZ zawiera bowiem dodatkowe założenia, np. przy 44 pytaniach każda odpowiedź powinna zostać przedstawiona w przekroju powiatów, gmin i województwa, w innym przypadku odpowiedzi na 27 pytań badawczych mają być udzielane odrębnie dla każdej z czterech dziedzin: edukacji, kultury, turystyki, sportu.

Tab. 3. Średnia liczba pytań badawczych w ewaluacjach różnych typów

Typ ewaluacji	Liczba pytań I	Liczba pytań II*
Operacyjne	25	27
Strategiczne	34	41
Operacyjno-strategiczne**	39	62
Ogółem	29	36

* Większe wartości w kolumnie „Liczba pytań II” wynikają z uwzględnienia pytań szczegółowych bądź różnych przekrojów czy dziedzin wskazanych w treści pytań, które de facto stanowią osobne zagadnienia.

** Ewaluacje obejmujące swoim zakresem zarówno bieżące problemy procesu wdrażania, jak i analizę efektów rzeczowych oddziaływania realizacji RPO na otoczenie.

Źródło: opracowanie własne na podstawie analizy dokumentów.

Co ciekawe, choć większa liczba pytań byłaby uzasadniona w ewaluacjach operacyjnych – powiązanych ściśle z procesem wdrażania i monitorowania, mających na celu wsparcie realizacji programu operacyjnego poprzez dostarczanie wniosków, rekomendacji, rozwiązań bieżących problemów, to w rzeczywistości więcej zagadnień, problemów badawczych ujmowanych jest w koncepcjach ewaluacji strategicznych – obejmujących głównie analizę i ocenę trafności ogólnych kierunków interwencji wyznaczonych na etapie programowania, służących głównie wspieraniu procesu uczenia się z efektów interwencji.

Zamawiający nie tylko formułują zbyt szerokie zakresy badań w SOPZ, lecz prowokują także dalsze ich rozszerzanie na etapie wyboru oferty. Kryteria merytoryczne oceny ofert najczęściej premiuje dodatkowe pytania i obszary badawcze, a nawet cele badawcze. Wykonawca chcący otrzymać zlecenie zmuszony jest więc „prześcignąć” konkurentów w liczbie wymyślonych dodatkowych pytań, obszarów, celów badawczych, które uzupełniają te wskazane (i tak najczęściej już zbyt liczne) w SOPZ. Zjawisko to przyjmuje bardzo poważne rozmiary. Liczba celów badawczych potrafi wzrosnąć z 4 do 8 czy z 3 do 5, a liczba pytań badawczych z 11 do 35 czy z 28 do 105. W związku z tym ostateczna koncepcja badania, wynikająca z raportu metodologicznego i bazująca na zapisach SOPZ uzupełnionych o propozycje z oferty jest często nie tylko zbyt szeroka, lecz znacznie odbiega także od pierwotnych intencji zamawiającego. W konsekwencji wykonawca na etapie realizacji poświęca ponad połowę czasu i zasobów na zagadnienia, które nie wynikają z potrzeb odbiorców.

Tak rozległe i czasami niespójne wewnętrznie zakresy badań są w dużej mierze rezultatem specyficznych uwarunkowań, a konkretnie swoistego trójkąta interesów i interakcji, w którym powstaje koncepcja badania. Wierzchołki owego trójkąta stanowią: 1) jednostka ewaluacyjna i jej pomysł na ewaluację, 2) komórki zarządzające, wdrażające RPO i ich potrzeby informacyjne, 3) wykonawca badania i jego własne ambicje, interesy, spojrzenie na badaną tematykę. Problem odmiennych spojrzeń, podejść i stawianych sobie celów przez różnych uczestników procesu strukturalizacji badania przybliżyć może lektura dwóch rozdziałów pierwszego polskiego podręcznika ewaluacji, w których kwestia organizacji ba-

dania ewaluacyjnego została przedstawiona z odmiennych perspektyw: zlecającego (Haber 2008) i wykonawcy (Ledzion 2008).

Ewaluacja jako produkcja raportów

Opisane powyżej zjawiska (znaczna liczba realizowanych ewaluacji, ich szeroki zakres oraz niska użyteczność) i ich powszechność świadczą o tym, że obecny system ewaluacji RPO koncentruje się na ilości, a nie jakości raportów. Mimo rosnącej liczby raportów ewaluacyjnych brakuje wiarygodnych odpowiedzi, dowodów skuteczności bądź nieskuteczności interwencji w ramach RPO i wyjaśnień mechanizmów powstawania efektów albo ich braku. To z kolei może powodować, że ewaluacja nie jest postrzegana jako użyteczne narzędzie zdobywania wiedzy i podstawa podejmowania decyzji. Ewaluacja zostaje sprowadzona do „produkcji raportów”, a nie tworzenia użytecznej wiedzy.

Przejawem owego nastawienia na produkcję może być zauważony fakt powielania tych samych badań w różnych regionach. Powielanie to obejmuje cele badawcze, ale i dosłowne przepisywanie całych grup pytań badawczych. Mamy więc do czynienia z przykładami procesu „produkcji raportów”, w ramach którego jednostka ewaluacyjna ułatwia sobie zadanie wymyślenia tematu, celów, pytań. Podobne zachowania mają miejsce po stronie wykonawców, którzy przygotowując dodatkowe pytania do oferty dla jednego województwa, wiernie kopiują pytania, które proponowali w innym województwie, bądź pytania, które stawiali zamawiający w innych województwach¹⁴. Stawianie tych samych pytań prowadzi następnie często do udzielania przez wykonawców tych samych odpowiedzi i formułowania takich samych rekomendacji. Nie napotkano natomiast przykładów twórczego wykorzystania wiedzy z poprzednich badań, przyjęcia wniosków z ewaluacji innych RPO jako bazy do zadawania kolejnych pytań, a swojego badania jako kolejnego kroku naprzód, narzędzia weryfikacji wcześniejszych wniosków, ich pogłębienia, spojrzenia z innej strony. Postrzegając więc wszystkie jednostki i ich działania jako jeden system ewaluacji, nie da się zauważyć, by następował proces akumulacji wiedzy, by kolejne badania uzupełniały poprzednie i stanowiły wkład w poszerzanie wiedzy o efektach i mechanizmach doprowadzania do zmian przez interwencję w ramach RPO.

Kolejnym przykładem mogącym świadczyć o tym, że potrzeba pozyskania konkretnej wiedzy nie zawsze jest podstawową motywacją do realizacji ewaluacji, są przypadki badań, w których wykonawca (zwykle na etapie tworzenia raportu metodologicznego) dokonuje istotnych modyfikacji koncepcji zamawiającego, ułatwiających mu zadanie, ale sprawiających, że pierwotne cele ewaluacji nie zostają osiągnięte. Fakt, że zamawiający godzi się na takie zmiany, pozwala podejrzewać, iż deklarowane cele i potrzeby informacyjne nie były jednak dla niego istotne. Poniżej przykłady opisywanego zjawiska.

- W ewaluacji wsparcia przedsiębiorców wykonawca w ramach badania kontryfaktycznego zamiast ankiety telefonicznej zrealizował internetową.

¹⁴ To samo dotyczy metod, celów, obszarów badawczych.

Zanotowano niewielki zwrot ankiet, poza tym utracono możliwości kontroli nad tym, kto odpowiada, a kto nie na ankietę. Nie uzyskano zatem dopasowania grup kontrolnej i eksperymentalnej i utracono możliwość zbadania efektów wsparcia RPO.

- Zamawiający zlecił badanie efektów wsparcia przedsiębiorców w perspektywie 2004–2006 wraz ze wskazaniem skutecznych typów wsparcia, wartych kontynuowania w perspektywie 2007–2013. Wykonawca przeprowadził jedynie analizę jakości procesu obsługi przedsiębiorców przez instytucje wdrażające i barier wsparcia w perspektywie 2004–2006.
- Jednym z elementów badania wymaganym przez zamawiającego była analiza komplementarności między realizowanymi projektami, którą zamawiający zamiennie nazywał spójnością. Wykonawca zdefiniował sobie spójność jako powiązania projektów z celami dokumentów strategicznych JST i ten aspekt poddał badaniu, pomijając kwestię komplementarności między projektami.

Konkluzje

Przeprowadzone badanie pokazało, że realizowane dotychczas ewaluacje RPO charakteryzują się w dużej części niską użytecznością. W ponad 2/3 analizowanych badań mniej niż połowę rekomendacji można uznać za istotną, mogącą służyć poprawie programu. W 72% badań nie udzielono odpowiedzi na niektóre z kluczowych pytań.

Dokonane obserwacje czynią wiarygodnymi przynajmniej część hipotez Leeuwa i Furubo (2008) w odniesieniu do systemu ewaluacji RPO. Zdecydowana większość formułowanych w raportach z badań rekomendacji dotyczy prostych wdrożeniowych kwestii i nie ma wpływu na kształt i trafność programów. Postawa instytucji zamawiających ewaluacje (formułowanie nierealnych oczekiwań, przyjmowanie niepełnych raportów bez odpowiedzi na pytania i istotnych rekomendacji, akceptowanie niekorzystnych modyfikacji zakresu badania) każe sądzić, iż to nie użyteczna wiedza jest ich priorytetem (a przynajmniej nie jedynym). Fakt, że mimo powyższych uchybień ewaluacja cieszy się taką popularnością i zrealizowano już ponad 230 badań RPO, może świadczyć o rutynowym, fasadowym charakterze tych działań. Ewaluacje są realizowane, bo taki jest formalny wymóg. Osoby bezpośrednio odpowiedzialne za ewaluację produkują rocznie kilka raportów, bo takie powierzono im obowiązki. Działania te nie są poparte refleksją nad przydatnością i wykorzystaniem tworzonej wiedzy.

Na rynku ewaluacji RPO ustaliła się trwała i niekorzystna praktyka, w ramach której zamawiający mają zwyczaj oficjalnie wymagać rzeczy nierealnych (przy określonych ograniczeniach budżetowych, czasowych), bo przyjęło się, że tak wygląda zakres badania. Wykonawcy mimo to konkurują o zlecenie i podejmują się jego realizacji, podnosząc jeszcze skalę trudności zadania, gdyż wiedzą, że w rzeczywistości nie oczekuje się od nich wiarygodnych, całościowych odpowiedzi. Zamawiający akceptuje raporty, w których nie ma wiarygodnych odpowiedzi na zadane pytania, wiedząc, że niczego więcej oczekiwać nie może. Mamy do czynienia z błędnym kołem, niska użyteczność raportów ewaluacyjnych utwier-

dza bowiem decydentów w przekonaniu, że ewaluacja nie jest przydatna przy podejmowaniu decyzji, i odwraca ich uwagę od tego narzędzia. Zaś brak zainteresowania decydentów ewaluacją sprawia, że kolejne badania nie są osadzone w ich potrzebach, nie są traktowane priorytetowo i w konsekwencji znowu kończą się raportami o niskiej jakości i użyteczności.

Pożądanym działaniem na przyszłość jest odejście od produkcji raportów na rzecz identyfikowania, kreowania i zaspokajania faktycznego popytu na wiedzę. Jasno komunikowanym celem i kryterium weryfikacji przydatności systemu ewaluacji nie może być liczba realizowanych badań, ale faktyczne wsparcie procesu decyzyjnego, konkretne wybory (i ich znaczenie), które zostały dokonane w oparciu o wiedzę pochodzącą z ewaluacji. Brak radykalnych działań w tym zakresie doprowadzi do utrzymania bądź pogłębienia obecnego, negatywnego stanu, w ramach którego marnuje się znaczące zasoby przeznaczane na ewaluację RPO, a decyzje odnośnie do zakresu i sposobu wdrażania programów operacyjnych nie są wspierane wiedzą, która mogłaby podnieść ich efektywność.

Literatura

- Alkin M.C., 1985, *A Guide for Evaluation Decision Makers*, Beverly Hills, CA: Sage.
- Alkin M.C., 1990, *Debates on Evaluation*, Newbury Park, CA: Sage.
- Alkin M.C., Coyle K., 1988, „Thoughts on evaluation utilization, misutilization and non-utilization”, *Studies in Educational Evaluation*, t. 14, nr 3, s. 331–340.
- Bachtler J., 2001, „*Quod erat demonstrandum* ? Ewaluacja polityki regionalnej”, *Studia Regionalne i Lokalne*, nr 4, s. 41–59.
- Bachtler J., 2011, „Evaluating the effects of regional policies – the European state of art”, w: K. Olejniczak, M. Kozak, S. Bienias (red.), *Evaluating the Effects of Regional Interventions. A Look Beyond Current Structural Funds' Practice*, Warsaw: Ministry of Regional Development, s. 90–113.
- Bienias S. et al., 2009, „Evaluation in Poland: Brief overview of evaluation process of EU Cohesion Policy funds”, w: P. Žuber (koord.), *Evaluation Systems in the Visegrad Member States*, Warszawa: Ministerstwo Rozwoju Regionalnego.
- Clarke A., 1999, *Evaluation Research: An Introduction to Principles, Methods and Practice*, London: Sage.
- Cousins J. B., Leithwood K.A., 1986, „Current empirical research on evaluation utilization”, *Review of Educational Research*, t. 56, nr 3, s. 331–364.
- Cousins J.B. (2003). „Utilization effects of participatory evaluation”, w: T. Kellaghan, D.L. Stufflebeam, L.A. Wingate (red.), *International Handbook of Educational Evaluation*, Boston: Kluwer, s. 245–265.
- Dahler-Larsen P., 1998, „Beyond non-utilization of evaluations: An institutional perspective”, *Knowledge, Technology, and Policy*, t. 11, nr 1–2, s. 64–90.
- Farell K. et al., 2002, *Evaluation Made Very Easy Accessible, and Logical*, Atlantic Centre of Excellence for Women's Health, <http://www.acewh.dal.ca/eng/reports/EVAL.pdf> [dostęp: 22.01.2013].
- Haber A., 2008, „Zarządzanie badaniem ewaluacyjnym – perspektywa zlecającego”, w: K. Olejniczak, M. Kozak, B. Ledzion (red.), *Teoria i praktyka ewaluacji inter-*

- wencji publicznych. *Podręcznik akademicki*, Warszawa: Wydawnictwa Akademickie i Profesjonalne, s. 212–230.
- Henry G.T., 2000, „Why not use?”, *New Directions for Evaluation*, nr 88, s. 85–98.
- Henry G.T., Mark M.M., 2003, „Beyond use: Understanding evaluation’s influence on attitudes and actions”, *American Journal of Evaluation*, t. 24, nr 3, s. 293–314.
- Johnson, R.B., 1998, „Toward a theoretical model of evaluation utilization”, *Evaluation and Program Planning*, t. 21, nr 1, s. 93–110.
- Ledzion B., 2008, „Zarządzanie kontraktem ewaluacyjnym – perspektywa wykonawcy”, w: K. Olejniczak, M. Kozak, B. Ledzion (red.), *Teoria i praktyka ewaluacji interwencji publicznych. Podręcznik akademicki*. Warszawa: Wydawnictwa Akademickie i Profesjonalne, s. 231–248.
- Leeuw F.L., Furubo J.E., 2008, „Evaluation systems: What are they and why study them?”, *Evaluation*, t. 14, nr 2, s. 157–169.
- Mark M.M., Henry G.T., 2004, „The mechanisms and outcomes of evaluation influence”, *Evaluation*, t. 10, nr 1, s. 35–57.
- National Evaluation Unit, Skórska P., 2011, „Poland”, w: K. Olejniczak, P. Strzęboszewski, S. Bienias (red.), *Evaluation of Cohesion Policy: Overview of Practices*, Warsaw: Ministry of Regional Development.
- Olejniczak K., 2005, „Teoria i praktyka ewaluacji w świetle doświadczeń europejskich”, *Studia Regionalne i Lokalne*, nr 4, s. 35–58.
- Olejniczak K., 2008, *Mechanizmy wykorzystania ewaluacji. Studium ewaluacji średnio-okresowych INTERREG III*, Warszawa: Wydawnictwo Naukowe Scholar.
- Olejniczak K., 2009, „Mini-case study: Building the evaluation system in Poland”, w: European Policies Research Centre, *Ex Post Evaluation of Cohesion Policy Programmes 2000-06 Co-financed by ERDF. Work Package 11: Management and Implementation System for Cohesion Policy*, Glasgow: European Commission.
- Owen J.M., 2007, *Program Evaluation: Forms and Approaches*, London: The Guilford Press.
- Patton M.Q., 1997, *Utilization-focused Evaluation: The New Century Text* (wyd. 3), Thousand Oaks, CA: Sage.
- Rossi P.H., Freeman H.E., Lipsey M.W. (2004). *Evaluation: A Systematic Approach* (wyd. 7), Thousand Oaks, London: Sage.
- Royse D., Thyer B., Padget D., Logan T., 2001, *Program Evaluation: An Introduction* (wyd. 3), Belmont, CA: Wadsworth.
- Saunders M., 2012, „The use and usability of evaluation outputs: A social practice approach”, *Evaluation*, t. 18, nr 4, s. 421–436.
- Scriven M., 2007, *Key Evaluation Checklist*, *Evaluation Checklists Project*, University of Michigan, http://dmeforpeace.org/sites/default/files/Scriven_Key%20Evaluation%20Checklist.pdf (dostęp: 22.01.2013).
- Toulemonde J., 2000, „Evaluation culture(s) in Europe: Differences and convergence between national practices”, *Vierteljahrshefte zur Wirtschaftsforschung*, t. 69, nr 3, s. 350–357.