

Użyteczność ewaluacji jako narzędzia zarządzania regionalnymi programami operacyjnymi

Tomasz Kupiec

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 ewaluacji nie udaje się odpowiedzieć na pytanie o wpływ RPO na rozwój społeczno-gospodarczy województw. Zaobserwowana niska użyteczność ewaluacji 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 programs

The article presents an analysis of usability of regional operational programs 2007-2013 (ROP) evaluations. Among studied elements of usability were: quality, relevance and credibility of the evaluation assumptions. The study comprised desk research of evaluation reports representative sample (n=71) completed between 2008 and 2012. Study results show that many reports do not contain important recommendations and do not answer key research questions. In most cases evaluations fail to answer the question about the impact of ROP on socio-economic development of regions. Limited usefulness of evaluation reports leads to conclusions about negative trends in the development of ROP evaluation system. System is focused on providing simple information, reports production, and does not respond to actual demand for knowledge.

Key words: evaluation usability, evaluation system, regional operational program, cohesion policy

Wprowadzenie¹

W ogólnym ujęciu, ewaluacja w literaturze definiowana jest najczęściej jako zestaw działań stosowanych do określenia jakości, ustalenia wartości, znaczenia danego obiektu, programu (Scriven 2007, s.1; Farrell et al. 2002, s.8; Rossi et al. 2004, s. 2). Jednym z charakterystycznych elementów, wskazywanym powszechnie jako dodatkowo wyróżniający ewaluację (w tym ewaluację interwencji publicznych), jest jej użyteczność. Ewaluacja powinna wspierać podejmowanie decyzji odnośnie obecnych i przyszłych działań publicznych, poprzez dostarczanie wiedzy, danych, niezbędnych do ich podjęcia (Patton 1997; Royse et al. 2001), powinna pomagać decydentom w dokonywaniu mądrych wyborów odnośnie 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*).

W teorii, racjonalnie działająca instytucja powinna realizować ewaluacje 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 realizowanych ewaluacji na podejmowane decyzje i kształt interwencji (Olejniczak 2008, Alkin, Coyle 1988, Henry 2000). W odpowiedzi wypracowano 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 pozorną nieracjonalność obserwowaną w rzeczywistości, 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 / dowodach, ale jedynie stworzenie wrażenia (w oczach społeczeństwa), że taką politykę się prowadzi.
3. System ewaluacji zorientowany jest przede wszystkim na swój własny rozwój. Osoby zaangażowane w system ewaluacji realizują swoje własne 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.

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

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

Mimo braku wcześniejszych tradycji, zaszczerpiąca za pośrednictwem wymogów związanych z wdrażaniem funduszy strukturalnych, praktyka ewaluacji zaczęła się gwałtownie rozwijać⁵. Liczba zlecanych ewaluacji z 5 w 2004 roku 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. PLN w 2004 r. i 2,2 mln PLN w roku 2006, w ostatnich latach osiągnął wartość ponad 25 mln PLN (Bienias i in. 2009, s.157).

Rosnący potencjał systemu ewaluacji i liczba realizowanych badań podkreślana jest często przez przedstawicieli administracji, w tym omawiana w licznych publikacjach Ministerstwa Rozwoju Regionalnego⁶. Ów wzrost, utożsamiany z rozwojem kultury ewaluacji, jest również zauważany i pozytywnie odbierany poza granicami kraju. KE postanowiła w ramach ewaluacji ex-post perspektywy 2000-2006 polityki spójności przygotować specjalny 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 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 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 ewaluacja tych programów i jej wykorzystanie. Użyteczność ewaluacji jako narzędzia wspierającego zarządzanie strategiczne i

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

⁴ Jak pisał w 2005 roku 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 Które wyjaśniają mechanizmy zmiany wywołanej przez podjętą interwencję, zaczynając od poniesionych nakładów, poprzez realizowane działania, ich bezpośrednie produkty, aż po efekty zwane często oddziaływaniami.

operacyjne rozwojem województwa wpływa bowiem na skuteczność realizacji celów RPO i ostatecznie – jak już wskazano powyżej – na poprawę dobrobytu społecznego.

Prezentowane badanie służyć może zarówno jako element rozliczenia dotychczasowych działań w zakresie ewaluacji na poziomie województw, jak i 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.

1. Cel, zakres i metodologia badania

Niniejszy artykuł prezentuje wyniki badania, którego celem było określenie użyteczności produktów dostarczanych przez system ewaluacji regionalnych programów operacyjnych (RPO)⁸, czyli raportów ewaluacyjnych.

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

- adekwatności ustaleń - czy przedstawione w raporcie rekomendacje spełniają potrzeby informacyjne zamawiającego, czy są spójne z celami badania,
- wiarygodności ustaleń – czy wnioski, w oparciu o które sformułowano rekomendacje są wiarygodne,
- jakości ustaleń – czy rekomendacje są czytelne, możliwe do wdrożenia (a nie niezrozumiałe, bądź nierealne), istotne, prowadzące do znaczących zmian w zakresie wdrażania, czy kształtu programu operacyjnego.

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

Kluczowe 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 dostarczaniu wiedzy mogącej wspierać proces podejmowania decyzji. Tym samym, niska użyteczność raportów

⁸ Który stanowi ważne ogniwo systemu ewaluacji polityki spójności w Polsce.

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

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.

2. Wyniki badania

2.1. 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 to nastąpić. Tabela rekomendacji jest kluczowym elementem raportu ewaluacyjnego, stanowiącym wyciąg najważniejszych ustaleń badania i przepis na ich wykorzystanie w procesie zarządzania RPO.

2.2. Adekwatność i wiarygodność ustaleń ewaluacji

Jedynie w 20% analizowanych ewaluacji RPO wszystkie, bądź większość przedstawionych przez wykonawcę rekomendacji, mających służyć poprawie programu operacyjnego, okazało się 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. Typowym przykładem braku zgodności pomiędzy celami badania a proponowanymi rekomendacjami jest 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 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 przypadków 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 poniższej tabeli.

Tabela 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.	<i>Nie odnosząc się w żaden sposób do skuteczności wymiany doświadczeń (pozytywnych efektów) zaleca się dalsze jej prowadzenie.</i>
Wysoki poziom zainteresowania projektami z zakresu oczyszczania ścieków.	Kontynuowanie bieżącej polityki z zakresu oczyszczania ścieków.	<i>Celem badania było wskazanie efektywności wsparcia. To właśnie wykazana efektywności powinna uzasadniać kontynuację wsparcia, a nie sam fakt zainteresowania ze strony potencjalnych beneficjentów.</i>
Niski poziom zainteresowania aplikowaniem o środki z zakresu gospodarki odpadami komunalnymi.	Zapewnienie większej ilości środków na realizację projektów w zakresie gospodarki odpadami komunalnymi.	<i>Zupełnie niezrozumiała propozycja zwiększenia alokacji, w sytuacji gdy już obecne środki nie są wykorzystywane.</i>

Źródło: opracowanie własne na podstawie analizy dokumentów.

Równie często spotkać można sytuację gdy 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.

2.3. 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, co zaprezentowano w poniższej tabeli.

Tabela 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 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, bowiem skutkowałoby to 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.</i> <i>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 projekty 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 poza tym, iż są powiązane z wnioskami z badania oraz czytelne i możliwe do wdrożenia, 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 istotne wdrożeniowo, a w 20% badań nie ma ani jednej takiej rekomendacji. Wśród ewaluacji efektów oraz efektów i procesów, aż 89% zawiera 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 nieodpowiadając 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 i rozwiązań. Użytkownicy badania - osoby, które zadały pytania dostają bowiem jedynie odpowiedzi, które już wcześniej znały 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, jedynie 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 wdrażania¹⁰. Dotyczy to w szczególności rekomendacji, które są mocno osadzone w kontekście obecnego programu operacyjnego, ale na ich wdrożenie jest już za późno, dlatego z konieczności, by w ogóle zwrócić uwagę na daną kwestię, zaleca się ich wykorzystanie po 2013 roku. Średni udział rekomendacji z terminem wdrożenia jeszcze w tej perspektywie (2007-2013) wynosi 43%. W ponad połowie badań rekomendacji takich jest mniej niż 30% (w tym w co czwartej ewaluacji procesów¹¹), a w 41% nie ma ani jednej (w tym w 14% ewaluacji procesów).

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, kiedy raport trafia już na przysłowiową półkę.

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

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%.

2.4. Kompletność raportów ewaluacyjnych

Duży udział ewaluacji, które nie prowadzą do istotnych rekomendacji wynika, m.in. z tego, iż w przypadku 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 są projekty spójne, możliwe do powiązania z celami strategii, np. w odniesieniu do celu *Rozwój szkolnictwa zawodowego zgodnie z potrzebami gospodarki* wskazano, że jest kilka projektów związanych z promocją i poprawą kształcenia zawodowego. Natomiast nie ma informacji o jakości tych projektów, trafności, adekwatności skali wsparcia w RPO i czy te projekty wywołały zauważalną zmianę w poziomie realizacji celów strategii.**
- **W odpowiedzi na pytanie o komplementarność projektów priorytetu „środowiskowego” RPO z projektami w ramach innych programów wykonawca podał jedynie po 3-4 przykłady projektów komplementarnych dla każdego z programów. Nie odniesiono się natomiast w żaden sposób do skali zjawiska w całej populacji¹³, nie wiadomo jaki jest udział projektów komplementarnych dla każdego 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 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.**

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

¹³ W priorytecie podpisano umowy na realizację ponad 90 projektów.

→ Zamawiający oczekiwał identyfikacji funkcji metropolitalnych miast stołecznych regionu, analizy wpływu RPO na ich rozwój i wskazania funkcji wymagających dalszego wsparcia i narzędzi ich wspierania. Wykonawca stworzył szeroki katalog potencjalnych obszarów, funkcji metropolitalnych, każdy 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 wspartych w RPO projektów 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 przypadkach w momencie realizacji ewaluacji brakuje wystarczającej liczby, odpowiednio dawno zakończonych projektów, których efektów można by już 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 / 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” prezentowana jest 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 jest próba odpowiedzi na 40 pytań badawczych w jednym raporcie, a zdarzają się i takie ewaluacje, gdzie pytań jest ponad 100. Nawet w przypadku, gdy pytań badawczych jest znacznie mniej, to niektóre z nich są na tyle rozległe i skomplikowane, że jedno pytanie mogłoby z powodzeniem stanowić samodzielny przedmiot badania o typowym dla ewaluacji RPO budżecie i czasie na realizację. Często prosta suma pytań badawczych nie oddaje faktycznego zakresu badania, bowiem SOPZ zawiera dodatkowe założenia, np. przy 44 pytaniach każda odpowiedź wymagana jest w przekroju powiatów, gmin i województwa, w innym przypadku odpowiedzi na 27 pytań badawczych mają być udzielane odrębnie dla każdej z 4 dziedzin: edukacji, kultury, turystyki, sportu.

Tabela 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 ¹⁴	39	62
Ogółem	29	36

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

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

¹⁴ Ewaluacje obejmującą swoim zakresem zarówno bieżące problemu procesu wdrażania, jak i analizę efektów rzeczowych, oddziaływania realizacji RPO na otoczenie.

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, ale następnie prowokują dalsze ich rozszerzanie na etapie wyboru oferty. Kryteria merytoryczne oceny ofert najczęściej premiują 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 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, ale znacznie odbiega od pierwotnych intencji zamawiającego. W konsekwencji wykonawca na etapie realizacji poświęca ponad połowę czasu i zasobów zajmując się zagadnieniami, które nie są zgodne z potrzebami 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 badania ewaluacyjnego została przedstawiona z odmiennych perspektyw zlecającego (Haber 2008) i wykonawcy (Ledzion 2008).

2.5. Ewaluacja jako produkcja raportów

Opisane powyżej zjawiska (znaczna liczba realizowanych ewaluacji, ich szeroki zakres oraz niska użyteczność) i ich powszechność sugerują, 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, bądź ich braku. To natomiast 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 dosłownie przepisane całe grupy pytań badawczych. Są to więc raczej przykłady 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 kopiuje pytania które proponowali w innym województwie, bądź pytania, które stawiali zamawiający w innych województwach¹⁵. Stawianie tych samych pytań

¹⁵ To samo dotyczy metod, celów, obszarów badawczych.

proceeds subsequently often to the granting by contractors of the same answers and formulating such recommendations. No examples of creative use of knowledge from previous research, acceptance of proposals for evaluation of other RPO as a basis for asking further questions, or viewing the current research as a further step forward, using verification tools of earlier proposals, their deepening, or looking at it from another angle. Observing all units and their actions as one evaluation system, it is not possible to notice that the process of knowledge accumulation would proceed, that further research would complement previous ones and constitute a contribution to the expansion of knowledge about effects and mechanisms of change through intervention within RPO.

Another question, which may suggest that the need for specific knowledge is not always the basic motivation for the realization of evaluation, are cases of research, in which the contractor (usually at the stage of creating a methodological report) makes significant modifications to the concept of the client, which facilitate his task, but cause that the primary goals of the evaluation are not achieved. The fact that the client agrees to such changes allows us to suspect that the declared goals and information needs are not important for him. Below are examples of the phenomenon:

- **W ewaluacji wsparcia przedsiębiorców wykonawca w ramach badania kontrfaktycznego zamiast ankiety telefonicznej zrealizował internetową. Zanotowano niewielki zwrot ankiet, poza tym utracono możliwości kontroli kto odpowiada a kto nie na ankietę, zatem nie uzyskano dopasowania grup kontrolnej i eksperymentalnej i utracono możliwość zbadania efektów wsparcia RPO.**
- **Zamawiający zlecił badanie efektów wsparcia przedsiębiorców w perspektywie 04-06 wraz ze wskazaniem skutecznych typów wsparcia, wartych kontynuowania w perspektywie 07-13. Wykonawca przeprowadził jedynie analizę jakości procesu obsługi przedsiębiorców przez instytucje wdrażające i barier wsparcia w perspektywie 04-06.**
- **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.**

3. Konkluzje

The conducted research showed that the implemented so far RPO evaluations are characterized by a high level of usefulness. In more than 2/3 of the analyzed research, less than half of the recommendations can be considered important, capable of serving to improve the program. In 72% of the research, no answer was given to some of the key questions.

The observations made are credible, at least partly, of the hypotheses of Leeuwa and Furubo (2008) in relation to the RPO evaluation system. The majority of the formulated in the reports of the research recommendations concern simple implementation issues and do not have an impact on the shape and effectiveness of the programs. The attitude of the institutions commissioning the evaluation (formulating unrealistic expectations, accepting incomplete reports without answers to questions and important recommendations, accepting unhelpful modifications to the scope of the research) also suggests that it is not useful knowledge (and at least not the only one) that is their priority. The fact that, despite the above-mentioned shortcomings, the evaluation enjoys such popularity and that more than 230 RPO evaluations have been carried out, suggests that these actions may

mieć charakter rutynowy, fasadowy. 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, wiedząc, 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. Opisany problem ma charakter błędnego koła, bowiem niska użyteczność raportów ewaluacyjnych utwierdza decydentów w przekonaniu, że ewaluacja nie jest przydatna do wspierania decyzji i odwraca ich uwagę od tego narzędzia. Zaś brak zainteresowania decydentów ewaluacją sprawia, że kolejne badania nie są osadzone w 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 znaczące zasoby przeznaczone na ewaluację RPO są marnowane, a decyzje odnośnie zakresu i sposobu wdrażania programów operacyjnych nie są wspierane wiedzą, która mogłaby podnieść ich efektywność.

Bibliografia

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*, nr 14(3), 331-40.

Bachtler J. (2001). „Quod erat demonstrandum ? Ewaluacja polityki regionalnej”, *Studia Regionalne i Lokalne*, nr 7(4), 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, 90-113.

Bienias S. i in. (2009). „Evaluation in Poland: Brief overview of evaluation process of EU Cohesion Policy funds”, w: P. Żuber (koordynacja) *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*, nr 56(3), 331-364.
- Cousins J.B. (2003). „Utilization effects of participatory evaluation”, w: T. Kellaghan, D. L. Stufflebeam, L. A. Wingate (Eds.), *International handbook of educational evaluation*, (pp. 245-265). Boston: Kluwer.
- Dahler-Larsen P., (1998). “Beyond non-utilization of evaluations: An institutional perspective”, *Knowledge, Technology, and Policy*, nr 11(1-2), 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 stycznia 2013.
- Haber A. (2008). „Zarządzanie badaniem ewaluacyjnym – perspektywa zlecającego”, w: Olejniczak K., Kozak M., Ledzion B. (red.), *Teoria i praktyka ewaluacji interwencji publicznych. Podręcznik akademicki*. Warszawa, Wydawnictwa Akademickie i Profesjonalne, s. 212-230.
- Henry G.T. (2000). „Why Not Use ?”, *New Directions for Evaluation*, nr 88, 85-98.
- Henry G.T., Mark M.M. (2003). „Beyond Use: Understanding Evaluation's Influence on Attitudes and Actions”, *American Journal of Evaluation*, nr 24(3), s. 293-314.
- Johnson, R. B. (1998). “Toward a theoretical model of evaluation utilization”, *Evaluation and Program Planning*, nr 21(4), 93-110.
- Ledzion B. (2008). „Zarządzanie kontraktem ewaluacyjnym – perspektywa wykonawcy”, w: Olejniczak K., Kozak M., Ledzion B. (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*, nr 14(2), 157-169.
- Mark M.M., Henry G.T. (2004). „The Mechanisms and Outcomes of Evaluation Influence”, *Evaluation*, nr 10(1), 35-57.
- National Evaluation Unit, Skórska P. (2011). *Poland*, w: *Evaluation of Cohesion Policy: Overview of practices*, edited by K. Olejniczak, P. Strzęboszewski, S. Bienias, Warsaw: Ministry of Regional Development.
- Olejniczak K. (2005). „Teoria i praktyka ewaluacji w świetle doświadczeń europejskich”, *Studia Regionalne i Lokalne*, nr 4(22), 35-58.
- Olejniczak K. (2008). *Mechanizmy wykorzystania ewaluacji. Studium ewaluacji średniookresowych INTERREG III*, Warszawa: Wydawnictwo Naukowe Scholar.
- Olejniczak K. (2009). „Mini-Case Study: Building the evaluation system in Poland”, w ramach: 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* (3rd ed.), Thousand Oaks, CA: Sage.
- Rossi, P.H., Freeman H.E., M.W. Lipsey (2004). *Evaluation: A Systematic Approach*, 7th edition Thousand Oaks, London, Sage.
- Royse D., Thyer B., Padget D., Logan T. (2001). *Program evaluation: An introduction* (3rd ed.), Belmont, CA: Wadsworth.

- Saunders M. (2012). "The use and usability of evaluation outputs: A social practice approach", *Evaluation*, nr 18(4), 421–436.
- Scriven M. (2007). *Key Evaluation Checklist, Evaluation Checklists Project*. University of Michigan, http://dmeformpeace.org/sites/default/files/Scriven_Key%20Evaluation%20Checklist.pdf, dostęp: 22 stycznia 2013.
- Toulemonde J. (2000). "Evaluation Culture(s) in Europe: Differences and Convergence between National Practices", *Vierteljahrshefte zur Wirtschaftsforschung*, nr 69(4), 350-357.