

## EWALUACJA W ADMINISTRACJI PUBLICZNEJ FUNKCJE, STANDARDY I WARUNKI STOSOWANIA

*Ewaluacja to systematyczna i obiektywna ocena programu lub polityki, ich założeń, procesu realizacji i rezultatów pod względem stosowności, skuteczności, trwałości, efektywności, a także użyteczności podjętych w ich ramach działań. Powinna dostarczać rzetelnych i przydatnych informacji o obiekcie badania wspierając w ten sposób proces decyzyjny oraz wspierając współdziałanie wszystkich partnerów zaangażowanych w realizację projektu.*

Należy się spodziewać, że akcesja Polski do Unii Europejskiej wzmocni znaczenie studiów ewaluacyjnych w szeroko pojętej sferze funkcjonowania polskiej administracji wszystkich szczebli, wśród organizacji III-sektora oraz wśród szerokiego grona potencjalnych instytucjonalnych i jednostkowych beneficjentów wsparcia z unijnych funduszy. Choć ewaluacja jest obecna w polskiej administracji już od kilku lat, wiedza na temat jej użytkowego potencjału, zakresu stosowania oraz efektów, jakie może przynieść, wymaga cały czas poszerzania i popularyzacji.

## SPIS TREŚCI

WNIOSKI I REKOMENDACJE .....	1
<u>1. Ewaluacja – podstawowe zagadnienia teoretyczne .....</u>	<u>6</u>
1.1. DEFINICJE EWALUACJI.....	6
1.2. KRÓTKA HISTORIA BADAŃ EWALUACYJNYCH.....	8
1.3. PODSTAWOWE FUNKCJE EWALUACJI .....	9
1.4. GŁÓWNE KRYTERIA OCENY PROGRAMÓW.....	11
1.5. STANDARDY JAKOŚCI BADAŃ EWALUACYJNYCH .....	13
1.6. TYPY I RODZAJE EWALUACJI .....	15
1.7. PROJEKTOWANIE BADANIA EWALUACYJNEGO .....	18
1.8. NAJPOPULARNIEJSZE METODY BADAWCZE EWALUACJI .....	20
<u>2. Techniki wzmacniające oddziaływanie badań ewaluacyjnych .....</u>	<u>25</u>
2.1. „DIALOGICZNOŚĆ” I „DEMOKRATYCZNOŚĆ” JAKO ZALECANE CECHY BADAŃ EWALUACYJNYCH.....	25
2.2. „PROCESS USE”, EWALUACJA JAKO OŚWIECENIE ORAZ EMPOWERMENT EVALUATION. MECHANIZMY WZMACNIAJĄCE WYKORZYSTANIE BADAŃ EWALUACYJNYCH .....	28
<u>3. Formy wykorzystania badań ewaluacyjnych w wybranych instytucjach i krajach Unii Europejskiej .....</u>	<u>29</u>
3.1. BADANIA EWALUACYJNE W PORZĄDKU PRAWNYM UNII EUROPEJSKIEJ ORAZ W PRAKTYCE DZIAŁANIA KOMISJI EUROPEJSKIEJ .....	30
3.2. EWALUACJA POLITYK PAŃSTWA WE FRANCJI NA PRZYKŁADZIE BADANIA STANU REALIZACJI POLITYKI W ZAKRESIE BEZPIECZEŃSTWA DROGOWEGO .....	38
3.3. EWALUACJA JAKO STANDARD ZARZĄDZANIA W ADMINISTRACJI NA PRZYKŁADZIE WIELKIEJ BRYTANII .....	41
3.4. WNIOSKI DLA POLSKI .....	45
<u>4. Wybrane aspekty obecności badań ewaluacyjnych w polskiej administracji .....</u>	<u>46</u>
4.1. OBECNOŚĆ EWALUACJI W PRAKTYCE PROGRAMOWANIA DZIAŁAŃ W ADMINISTRACJI CENTRALNEJ W POLSCE. ANALIZA NARZĘDZI I TECHNIK EWALUACJI W STRATEGIACH I PROGRAMACH RESORTOWYCH OPRACOWANYCH W LATACH 2000 – 2004 .....	46
4.2. WARUNKI OBECNOŚCI EWALUACJI W PROGRAMOWANIU EUROPEJSKICH FUNDUSZY STRUKTURALNYCH .....	50

## WNIOSKI I REKOMENDACJE

1. Badania ewaluacyjne są metodą wieloaspektowej (uwzględniającej wskaźniki o charakterze społecznym i ekonomicznym), demokratycznej oceny jakości i efektywności programów, projektów, a nawet polityk publicznych. Ich zaletami są przede wszystkim: uwzględnianie perspektyw podstawowych grup zaangażowanych w przedmiot badania, budowanie zaufania pomiędzy badającymi i badanymi, a także inicjowanie zmian postaw jednostek i grup w kierunku większej otwartości na zmiany, innowacje i podnoszenie jakości własnych działań.
2. *Ewaluacja* jest terminem o szerokim znaczeniu i różnorodnych często definicjach. Oprócz jej formy „klasycznej” (badania szeroko rozumianego wpływu danego programu na otoczenie społeczno-ekonomiczne, budującego mechanizmy zaufania, współpracy i współdziałania pomiędzy realizatorami programu a innymi zainteresowanymi programem grupami) nazwą „ewaluacja” określa się w praktyce działania administracji zadania o charakterze ściśle kontrolnym, często dotyczące wyłącznie sfery środków finansowych programu, które nie mają faktycznie charakteru ewaluacji. Jednocześnie występuje ona coraz szerzej w formie nienazwanej, jako standard dobrego zarządzania organizacją w ogóle.
3. Mimo ewidentnego chaosu definicyjnego znaczenie badań o charakterze ewaluacyjnym w sferze zarządzania w sektorze publicznym wzrasta. Świadczy o tym wielość dokumentów prawnych i kierunkowych dotyczących ewaluacji powstających w administracjach obszaru Unii Europejskiej, realizowanych badań ewaluacyjnych, a także proces instytucjonalizacji działań o charakterze ewaluacyjnym w tych administracjach.
4. Obecna rosnąca popularność badań ewaluacyjnych wynika ze współczesnych tendencji w zakresie zarządzania organizacjami (m.in. *Total Quality Management*, *New Public Management*), a w kontekście Unii Europejskiej ze wzrastającą rolą funduszy strukturalnych. Biorąc pod uwagę dynamiczny rozwój teorii zarządzania, a także rozwój polityki strukturalnej Unii Europejskiej (w przypadku Polski – także świeżo uzyskane członkostwo w Unii), należy przypuszczać, że zainteresowanie ewaluacją będzie nadal

systematycznie rosło i będzie odgrywać rolę swoistego „znaku jakości” zarządzania, także w administracji.

5. Aczkolwiek obecność badań ewaluacyjnych w polskiej administracji publicznej jest obecnie związana przede wszystkim z obsługą funduszy strukturalnych Unii Europejskiej, mogą być one potencjalnie wykorzystywane do analizy szerokiej gamy programów, projektów, a nawet polityk, nie związanych z polityką strukturalną Unii. Należy zatem przypuszczać, że obecny zasięg przedmiotowy ewaluacji w polskiej administracji będzie się stopniowo poszerzał, a tym samym wzrośnie jej ogólne znaczenie w zarządzaniu tym sektorem.
6. Mimo stałej obecności tradycji kojarzenia ewaluacji wyłącznie z analizami finansowymi lub kontrolą, znacznie zyskuje model ewaluacji o profilu społecznym, odwołujący się do potrzeby dialogu z grupami zainteresowanymi programem, swego rodzaju „negocjowania perspektyw”, w jakich program jest postrzegany. Ewaluacja w tym ujęciu może nawet odgrywać rolę promocji badanego programu, budować wokół niego dobry klimat społeczny.
7. Wartość zrealizowanej ewaluacji mierzy się przede wszystkim jej pozytywnym wpływem na badany program. Oprócz trudno mierzalnych efektów o charakterze „miękkim”, takich jak zintegrowanie grup zainteresowanych programem, polepszenie przepływu komunikacji pomiędzy grupami odpowiedzialnymi za jego realizację, skutkiem ewaluacji powinny być również m.in. istotne decyzje o charakterze finansowym czy organizacyjnym. Decyzje te wymagają jednak zazwyczaj zgody politycznej, która jest trudna do osiągnięcia bez „kultury ewaluacyjnej” – poszanowania dla wyników badania, postrzegania go jako środka naprawy, a nie narzędzia kontroli podejmowanych działań.
8. W praktyce polskiej administracji programy i strategie traktowane są zazwyczaj jako kierunkowe, nie tworzące potrzeby monitorowania ich realizacji w czasie. Wydaje się, że w ten sposób utracona zostaje nie tylko możliwość wzmocnienia programu poprzez dostarczenie mu analiz społeczno-ekonomicznych, ale także niewykorzystany zostaje potencjał promocyjny poddawania danego programu ewaluacji.
9. Niezbędnym warunkiem efektywnej ewaluacji jest niepodważalne zaufanie do jej wyników. Dlatego traktowanie jej na sposób fasadowy może nie tylko ograniczać

korzyści z jej przeprowadzenia, ale także może podważać zaufanie do możliwości przeprowadzenia badania tego typu w sposób godny zaufania.

Niezbędna jest zatem promocja badań ewaluacyjnych w polskiej administracji jako działań zwiększających efektywność, przejrzystość i etyczność zarządzania w tym sektorze. Obecny sposób postrzegania ewaluacji przede wszystkim jako narzędzie kontroli nie ujmuje charakterystycznej dla niej funkcji kulturotwórczej.

- A. Biorąc pod uwagę, że systematyczna realizacja badań ewaluacyjnych staje się standardem funkcjonowania administracji w Unii Europejskiej, wydaje się, że konieczne jest stworzenie programu szkoleń z zakresu metod i technik ewaluacji, z uwzględnieniem specyfiki administracyjnej, skierowanych do szerokiego grona pracowników administracji publicznej. Podstawowa wiedza z tego zakresu będzie przydatna nie tylko w jednostkach, które elementy badań ewaluacyjnych wykonują samodzielnie, ale także w takich, które badania tego rodzaju zlecają lub nadzorują.
- B. Budowanie swoistej „kultury ewaluacyjnej” polskiej administracji powinno odbywać się dwutorowo. Obok ewaluacji dużych programów takich jak Narodowy Plan Rozwoju i Programy Operacyjne ewaluacji poddawać należy także projekty mniejsze (np. cykle szkoleń prowadzone przez poszczególne ministerstwa, projekty informacyjne prowadzone przez resorty, pilotaże inicjatyw samorządowych itd.). Istotne jest, aby badania ewaluacyjne prowadzone były na wszystkich szczeblach administracji i by prowadzone były na sposób demokratyczny, tak by mogły zyskać zaufanie w środowisku pracowników administracji, zapewnić sobie ich przychylność do dzielenia się informacjami na temat badanych programów, a w przyszłości - akceptację dla decyzji, jakie będą podejmowane w oparciu o wyniki ewaluacji.
- C. Budowaniu w administracji publicznej zaufania wobec badań ewaluacyjnych sprzyjać będzie wymiana doświadczeń z bliźniaczymi instytucjami administracji publicznej z innych krajów Unii Europejskiej. Chodzi o przekazywanie praktycznych, popartych przykładami, informacji na temat efektów prowadzenia badań ewaluacyjnych, możliwych trudności związanych z ich realizacją, wykorzystywanej metodologii, zakresu odbiorców raportu z badań itd.



Ostatnie dziesięciolecie można nazwać czasem „boomu” badań ewaluacyjnych. Przyczyn zjawiska można upatrywać w obserwowanej w wielu krajach tendencji do decentralizacji zarządzania środkami publicznymi oraz w rosnącym nacisku na poprawę jakości zarządzania, także w sektorze publicznym. Obok powstających dynamicznie w całej Europie instytucji zajmujących się teoretyczną refleksją nad zjawiskiem ewaluacji, duży udział w rozwoju i promocji badań ewaluacyjnych ma Bank Światowy, Organizacja Współpracy Gospodarczej i Rozwoju (OECD) oraz Program Narodów Zjednoczonych ds. Rozwoju (UNDP). Szczególnie Bank Światowy przyczynił się w ostatnich latach do popularyzacji badań ewaluacyjnych wśród kadry zarządzającej różnego rodzaju programami społecznymi oraz wśród szerokiego beneficjentów tych programów.<sup>1</sup>

Rozwój zapotrzebowania na badania ewaluacyjne jest także związany z intensywnym rozwojem funduszy strukturalnych Unii Europejskiej. Obowiązek ewaluacji programów przez nie współfinansowanych stał się w 1999 roku częścią *acquis communautaire*<sup>2</sup>, a różne formy badań ewaluacyjnych stały się istotnym elementem cyklu zarządzania funduszami europejskimi. O skali znaczenia ewaluacji w tym kontekście świadczy przykład Irlandii, gdzie środki przeznaczone na ewaluację programów realizowanych przy udziale funduszy strukturalnych UE wynoszą średnio nawet od 0,6 do 1,4% ogólnej sumy środków zaangażowanych w te programy.<sup>3</sup>

Należy się spodziewać, że akcesja Polski do Unii Europejskiej wzmocni znaczenie studiów ewaluacyjnych w szeroko pojętej sferze funkcjonowania polskiej administracji wszystkich szczebli, wśród organizacji III-sektora oraz wśród szerokiego grona potencjalnych instytucjonalnych i jednostkowych beneficjentów wsparcia z unijnych funduszy. Choć zjawisko ewaluacji jest obecne w polskiej administracji już od kilku lat, wiedza na temat użytkowego potencjału tego rodzaju badań, zakresu ich stosowania oraz efektów, jakie mogą przynieść, wymaga jednak cały czas poszerzania i popularyzacji. W dłuższej perspektywie

---

<sup>1</sup> Obszerne informacje na temat badań ewaluacyjnych prowadzonych lub zamawianych przez Bank Światowy można znaleźć m.in. na następujących stronach: <http://www.worldbank.org/oed/> (stan z dnia 30.06.2004 r.) oraz <http://www.worldbank.org/wbi/evaluation/> (stan z dnia 30.06.2004 r.); interesujące informacje nt. ewaluacji prowadzonych w ramach Programu Narodów Zjednoczonych ds. Rozwoju można znaleźć na stronie <http://www.undp.org/eo/knowledge-management> (stan z dnia 01.07.2004 r.)

<sup>2</sup> rozporządzenie Rady nr 1260 z dnia 21 czerwca 1999 r.

<sup>3</sup> por. Evaluating Socio-Economic Programmes. Evaluation Design and Management. MEANS Collection Volume 1, European Commission 1999, s. 44, 79; średni koszt ewaluacji programu realizowanego z funduszy strukturalnych Unii Europejskiej wynosi obecnie ok. 0,02 jego wartości

istotne jest także określenie zakresu, w jakim badania ewaluacyjne powinny być obecne w sferze polskiej administracji publicznej, która oprócz utrzymywania odpowiednich standardów zarządzania w szerokiej sferze stosunków publicznych jest także odpowiedzialna za ich kreowanie.

## **1. Ewaluacja – podstawowe zagadnienia teoretyczne**

W ostatnim dziesięcioleciu dynamicznie zwiększyła się ilość literatury poruszającej nie tylko praktyczne, ale także ściśle teoretyczne aspekty badań ewaluacyjnych. Pośród sporów o definicję zjawiska w ciągu ostatnich kilkunastu lat powstała metodologia tych badań<sup>4</sup>, zestaw kryteriów oceny w odniesieniu do badanych programów, a także standardy etyczne pracy ewaluatorów.

W niniejszym rozdziale z szerokiego spektrum zagadnień, wokół których toczy się dyskurs na temat natury i roli badań ewaluacyjnych we współczesnych organizacjach, pomijając na razie specyficzny kontekst sfery administracji publicznej, poruszone zostaną m.in. kwestia definicji ewaluacji oraz funkcji, które może ona pełnić wobec obiektów swoich badań i jednostki zlecającej dane badanie. Ponadto, na tle schematu realizacji badania ewaluacyjnego zaprezentowane zostaną charakterystyczne metody badawcze, wpływające na unikalność i rozległe możliwości funkcjonalne tej praktyki.

### **1.1. Definicje ewaluacji**

Ewaluacja w administracji jest najczęściej rozumiana jako badanie zorientowane na ocenę polityk i programów podejmowanych przez władze publiczne pod kątem osiągniętych wyników, które może być wykorzystywane przez decydentów na różnych poziomach zarządzania do weryfikowania efektywności działania sektora publicznego.

Gdyby jednak chcieć podnieść podstawowe wątki toczącego się systematycznie dyskursu na temat teorii ewaluacji, należałoby przytoczyć definicję Michaela Quinna Pattona, wieloletniego prezesa Amerykańskiego Towarzystwa Ewaluacyjnego, który określił ewaluację jako *proces systematycznego gromadzenia informacji na temat działania, właściwości i rezultatów programów, personelu oraz produktów, które są wykorzystywane*

---

<sup>4</sup> Do najpopularniejszych podręczników metodologii badań ewaluacyjnych należą: John M. Owen., Patricia J. Rogers, Program Evaluation. Forms and Approaches, SAGE 1999; Michael Quinn Patton, Qualitative Research & Evaluation Methods. Third Edition, SAGE Publications, 2001 oraz Robert Stake, The Art of Case Study Research, SAGE Publications 1995

przez specjalistów do redukowania niepewności wokół programu i poprawienia jego efektywności, a także służą podejmowaniu decyzji z uwagi na to, co owe programy, personel lub produkty robią oraz czego dotyczą.”<sup>5</sup> Ta definicja nie tylko traktuje badanie ewaluacyjne systematyzuje wyróżniające cechy ewaluacji (*gromadzenie informacji, szeroki wachlarz problemów będących przedmiotem ewaluacji, cel i zastosowanie badań*), nie tylko pokazuje, że jest to technika uniwersalna, nie przypisana do żadnego określonego sektora gospodarki czy typu organizacji, ale także podkreśla, że jest ona efektywna tylko wtedy, gdy jej wyniki są przez decydentów w konstruktywny sposób spożytkowane. Oznacza to, że podstawowym kryterium jakości badań ewaluacyjnych jest ich *przydatność*.

Istotna jest też definicja ewaluacji Petera H. Rossiego, Howarda E. Freemana i Marka W. Lipselya. Zwraca ona uwagę na jej wymiar *naprawczy* – tj. na fakt, że jej celem długoterminowym jest poprawa sytuacji społecznej. W tej interpretacji ewaluację określa się jako *wykorzystanie metod badań społecznych do systematycznego określania efektywności programów społecznych pod względem tego w jaki sposób adaptują się do swojego otoczenia organizacyjnego i politycznego oraz w jaki sposób odpowiadają na potrzeby informacyjne, których zaspokojenie oznacza w założeniu poprawę sytuacji społecznej*.<sup>6</sup>

Definicja ewaluacji pojawiająca się w publikacjach Komisji Europejskiej jest znacznie mniej złożona i wężiej zakreśla ramy zjawiska. Ewaluacja jest tam *określeniem wartości danego działania publicznego, a więc polityki, programu lub projektu w odniesieniu do wcześniej zdefiniowanych kryteriów i w oparciu o specjalnie zebrane i zanalizowane informacje*.<sup>7</sup> Wydaje się jednak, że to przede wszystkim właśnie ono przenika najszerzej do praktyki administracyjnej.

Ze względu na wielość odmian badań ewaluacyjnych wiele jest terminów, których, czasem w sposób nieuprawniony, używa się zamiennie z terminem *ewaluacja*. Najczęściej jest ona błędnie utożsamiana z *kontrolą* lub *monitorowaniem*. Warto wyjaśnić zatem, że ewaluacja w przeciwieństwie do *monitoringu* i *kontroli* nie skupia się na aspekcie finansowym i rzeczowym efektów programu, ani wyłącznie na monitorowaniu jego postępów, ale przede wszystkim na efektach długoterminowych. Kryteriami oceny

---

<sup>5</sup> Colin Robson, Projektowanie ewaluacji w: Ewaluacja w edukacji, red. L. Korporowicz, Oficyna Naukowa, Warszawa 1997, s. 152

<sup>6</sup> Peter H. Rossi, Howard E. Freeman, Mark W. Lipsey, Evaluation. A Systematic Approach. Seventh Edition. SAGE 2004; por. Tomasz Kierzkowski, Ocena (Ewaluacja) Programów i Projektów o charakterze społeczno-gospodarczym w kontekście przystąpienia Polski do Unii Europejskiej, 2002 r., <http://www.parp.gov.pl/doc/raport11.pdf> (stan z dnia 02.05.2004 r.)

założonymi przez ewaluatorów nie jest legalność, bądź zgodność realizacji projektu z jego dokumentami programowymi, jak to jest w przypadku kontroli, lecz jego skuteczność, wykonalność, poprawność oraz stosowność. W szerszym sensie wyróżnikiem ewaluacji jest ponadto to, że pragnie ona opisać *społeczne* efekty badanego programu – tj. jest wyczulona na społeczny kontekst jego realizacji.

Warto przy tym zauważyć, że w polskich opracowaniach znacznie częściej niż terminu *ewaluacja* używa się, zgodnie z jego etymologią, słowa *ocena*. Taką praktykę, aczkolwiek dość powszechną, można nazwać nieprawidłową z dwóch powodów. Po pierwsze, zgodnie z trendami międzynarodowymi już od kilkudziesięciu lat ewaluację należy traktować jako odrębny rodzaj badań społecznych, który dynamicznie rozwija swoją teorię i praktykę. Po drugie, wielu uznanych teoretyków ewaluacji (m.in. E. Guba i Y. Lincoln, R. Stake, E. House) uważa, że badania ewaluacyjne nie powinny, paradoksalnie, mieć w ścisłym sensie charakteru oceniającego. Sytuacja wydawania oceny implikuje bowiem nierównowagę pomiędzy oceniającym a ocenianym, a wspomniani badacze uważają, że interwencja ewaluatorów powinna być, poprzez techniki angażowania w proces badawczy wszystkich zainteresowanych rodzajem dialogu, negocjacji i wzajemnego doskonalenia się. W tej interpretacji komponent oceniający badania traci zatem swe podstawowe znaczenie.

## **1.2. Krótka historia badań ewaluacyjnych**

Choć korzeni badań ewaluacyjnych upatruje się tradycyjnie we wczesnych badaniach nad programami szkolnymi prowadzonych w Stanach Zjednoczonych pod koniec XIX wieku<sup>8</sup>, to główną przyczyną ich rozwoju wydaje się być gwałtowny rozwój nauk społecznych w XX stuleciu oraz rozwój refleksji nad problematyką zarządzania w handlu i przemyśle.

Do lat sześćdziesiątych ewaluacja była utożsamiana z pomiarem osiągnięć szkolnych. Naturalnym sprawdzianem pracy szkoły stały się w tym czasie testy, które potrafiły mierzyć przystosowanie uczniów do precyzyjnie określonych standardów dydaktycznych.

O konieczności doskonalenia praktyk ewaluacyjnych zadecydowało upokorzenie amerykańskiego establishmentu, jakim było wystrzelenie przez Rosjan Sputnika w przestrzeń kosmiczną w roku 1957. Rząd Stanów Zjednoczonych za cel strategiczny przyjął

---

<sup>7</sup> Evaluating Socio-Economic Programmes. Evaluation Design and Management, MEANS Collection, Volume 1, European Commission 1999, s. 17

<sup>8</sup> Obszerny zarys historii rozwoju badań ewaluacyjnych znaleźć można w: Egon G. Guba, Yvonne S. Lincoln, Fourth Generation Evaluation, SAGE Publications, Newbury Park - London - New Dehli 1989

podniesienie poziomu edukacji w szkołach publicznych. Jednocześnie gwałtowny wzrost ekonomiczny po II Wojnie Światowej sprawił, że pomiędzy rokiem 1950 a 1979 podwoiła się w USA wartość produktu narodowego brutto oraz wydatki na programy opieki społecznej.<sup>9</sup> Realizowana w latach 60. polityka interwencjonizmu owocowała olbrzymimi inwestycjami. Badania ewaluacyjne miały zaś być sposobem na zrationalizowanie wydatków rządu. Chodziło o oszacowanie wyników programów oraz o znalezienie przyczyn konkretnych sukcesów i porażek. Często badano też skuteczność pilotażowych, pozarządowych alternatyw wobec programów publicznych.

Mimo że w Stanach Zjednoczonych ewaluacja zinstytucjonalizowała się ostatecznie już około czterdzieści lat temu, do praktyki zarządzania w krajach europejskich przenikała powoli. Najszybciej zaczęły wykorzystywać ją Wielka Brytania (głównie w sferze edukacji) i kraje skandynawskie. Obecnie działa w Europie 11 krajowych towarzystw ewaluacyjnych oraz umiejscowione w Brukseli Europejskie Towarzystwo Ewaluacyjne.<sup>10</sup> Komisja Europejska regularnie zamawia badania ewaluacyjne u ekspertów zewnętrznych. Jednocześnie szereg funkcjonujących w jej ramach Dyrekcji Generalnych utrzymuje własne komórki zajmujące się sprawami ewaluacji (szczególnie aktywna działa w ramach *DG Budget*<sup>11</sup>), a dodatkowo funkcjonuje w ramach KE sieć ewaluatorów wewnętrznych, m.in. reprezentujących wszystkie Dyrekcje Generalne Komisji.<sup>12</sup>

W Polsce od 2000 roku działa Polskie Towarzystwo Ewaluacyjne, szereg ośrodków akademickich (Uniwersytet Warszawski, Uniwersytet Gdański) prowadzą studia podyplomowe z zakresu technik i metod ewaluacji, a na rynku funkcjonują prywatne firmy specjalizujące się w realizacji tego typu badań oraz coraz szersza grupa badaczy niezależnych. Ich zleceniodawcami są jednostki administracji publicznej, agencje państwowe, zakłady publiczne, placówki edukacyjne i organizacje pozarządowe.

### **1.3. Podstawowe funkcje ewaluacji**

W teorii badań ewaluacyjnych wymienia się zazwyczaj cztery podstawowe ich funkcje:

- funkcję *formatywną*,

---

<sup>9</sup> W. R. Shadish, Jr., Cook T. D., Leviton L., Foundations of Program Evaluation. Theories of Practice, SAGE Publications 1991, s. 22

<sup>10</sup> <http://www.europeanevaluation.org> (stan z dnia 30.06.2004 r.)

<sup>11</sup> [http://www.europa.eu.int/comm/budget/evaluation/index\\_en.htm](http://www.europa.eu.int/comm/budget/evaluation/index_en.htm) oraz

- funkcję *konkluzywną*,
- funkcję *psychologiczną* (nazywaną też funkcją *społeczną* bądź *polityczną*) oraz
- funkcję *administracyjną*.<sup>13</sup>

Funkcję *formatywną* (doskonalącą) ewaluacja pełni wtedy, kiedy badanie ewaluacyjne towarzyszy realizacji danego programu, lokalizuje te jego obszary, które wymagają poprawy, uzasadnia potrzebę dokonania poprawek, a także podpowiada ich kierunek. Kiedy ewaluacja służy celom sprawozdawczości bądź stanowi podstawę dla podjęcia wiążących decyzji o programie (np. o jego zakończeniu, kontynuacji, nadaniu certyfikatu itd.) pełni tzw. funkcję *konkluzywną*. Funkcję *psychologiczną* (*społeczną/polityczną*) badanie spełnia, o ile ono samo lub jego wyniki są wykorzystywane do kształtowania świadomości dotyczącej poszczególnych działań, motywowania pożądaných zachowań lub promowania określonych stosunków społecznych. Funkcję *administracyjną* – kiedy służy sprawowaniu władzy i komponent oceniający ewaluacji bywa wykorzystywany jako narzędzie nacisku.

W praktyce działania administracji publicznej funkcje ewaluacji to przede wszystkim:

1. *badanie efektywności realizowanych programów* (w tym m.in. określanie czynników wpływających na poziom realizacji postawionych przez program celów oraz formułowanie rekomendacji na rzecz lepszego wykorzystania zasobów)
2. *tworzenie pozytywnego klimatu społecznego wokół badanego projektu* (m.in. poprawa komunikacji pomiędzy grupami zaangażowanymi w badany projekt i w samo badanie, wymiana informacji pozwalająca na uzyskanie wieloaspektowego wizerunku projektu)
3. *wspomaganie zarządzania strategicznego* (dostarczanie analiz i rekomendacji będących podstawą dla budowania strategii oraz podejmowania decyzji)

Obszar zastosowania badań ewaluacyjnych jest w rzeczywistości bardzo szeroki. W USA i krajach zachodnioeuropejskich zlecenie takich badań wchodzi do kultury organizacyjnej administracji publicznej, jednostek edukacyjnych i organizacji pozarządowych. Poszczególne kraje UE różnią się, jeśli chodzi o popularność ewaluacji i wpływ na procesy podejmowania decyzji. Zakres wykorzystania ewaluacji można rozpatrywać na trzech poziomach: pierwszym – tzw. *regulacyjnym* (– badania ewaluacyjne wykorzystywane są wyłącznie jako technika prowadzenia obligatoryjnej kontroli

<sup>12</sup> [http://www.europa.eu.int/comm/budget/evaluation/Internal\\_network/internalnetwork\\_en.htm](http://www.europa.eu.int/comm/budget/evaluation/Internal_network/internalnetwork_en.htm) (stan z dnia 30.06.2004 r.)

<sup>13</sup> por. David Nevo, *Konceptualizacja ewaluacji edukacyjnej. Analityczny przegląd literatury w: Ewaluacja w edukacji*, red. L. Korporowicz, Oficyna Naukowa, Warszawa 1997

w instytucjach i programach), drugim - tzw. *menedżerskim* (który zakłada, że ewaluacja jest wykorzystywana jako technika wspomagająca kadre zarządzającą w projektowaniu i realizacji działań) lub najwyższym - tzw. poziomie decyzji politycznych (kiedy opinia publiczna jest systematycznie informowana o wynikach ewaluacji wydatkowania publicznych funduszy i staje się uznanym narzędziem demokracji). Krajami o najbardziej zakorzenionej kulturze ewaluacji, gdzie ewaluacja funkcjonuje na poziomie decyzji politycznych, są w Europie: Szwecja, Holandia, Dania, Irlandia i Wielka Brytania. Krajami, w których ewaluacja jest postrzegana przede wszystkim jako forma obowiązkowej kontroli, a tym samym ma tylko doraźny wpływ na jakość zarządzania funduszami publicznymi są: Włochy (regiony południowe), Grecja, terytoria zamorskie Francji, Hiszpania, Portugalia oraz wschodnie landy Niemiec.<sup>14</sup>

#### 1.4. Główne kryteria oceny programów

Zakres przedmiotowy badań ewaluacyjnych, w tym - problemy, które muszą zostać rozwiązane dzięki realizacji badania, określają tzw. *pytania ewaluacyjne*. Formułowane są zazwyczaj przede wszystkim w oparciu o analizę dostępności niezbędnych danych oraz wielkość budżetu badania. Grupuje się je zazwyczaj wedle pięciu klasycznych kryteriów ewaluacyjnych: *skuteczności*, *efektywności*, *odpowiedniości*, *użyteczności* oraz *długotrwałości* (*perspektywiczności*).

Kryterium *skuteczności* (*effectiveness*) – pozwala ocenić, czy cele programu zdefiniowane na etapie planowania zostały osiągnięte lub są osiągane według ustalonego harmonogramu. Oceniając program pod względem skuteczności porównuje się jego rzeczywiste efekty na poziomie produktu, rezultatu i oddziaływania z założeniami zawartymi w dokumentach programowych. Program może być oceniany pod względem skuteczności w połowie okresu programowania (ewaluacja *mid-term*) oraz w ramach ewaluacji *ex post*.

Oceniając program pod względem jego *efektywności* (wydajności) (*efficiency*) porównuje się zasoby zaangażowane przy jego realizacji z jego rzeczywistymi osiągnięciami. Podobnie, jak w przypadku kryterium skuteczności, efektywność programu można oceniać na półmetku realizacji programu i po jego zakończeniu.

---

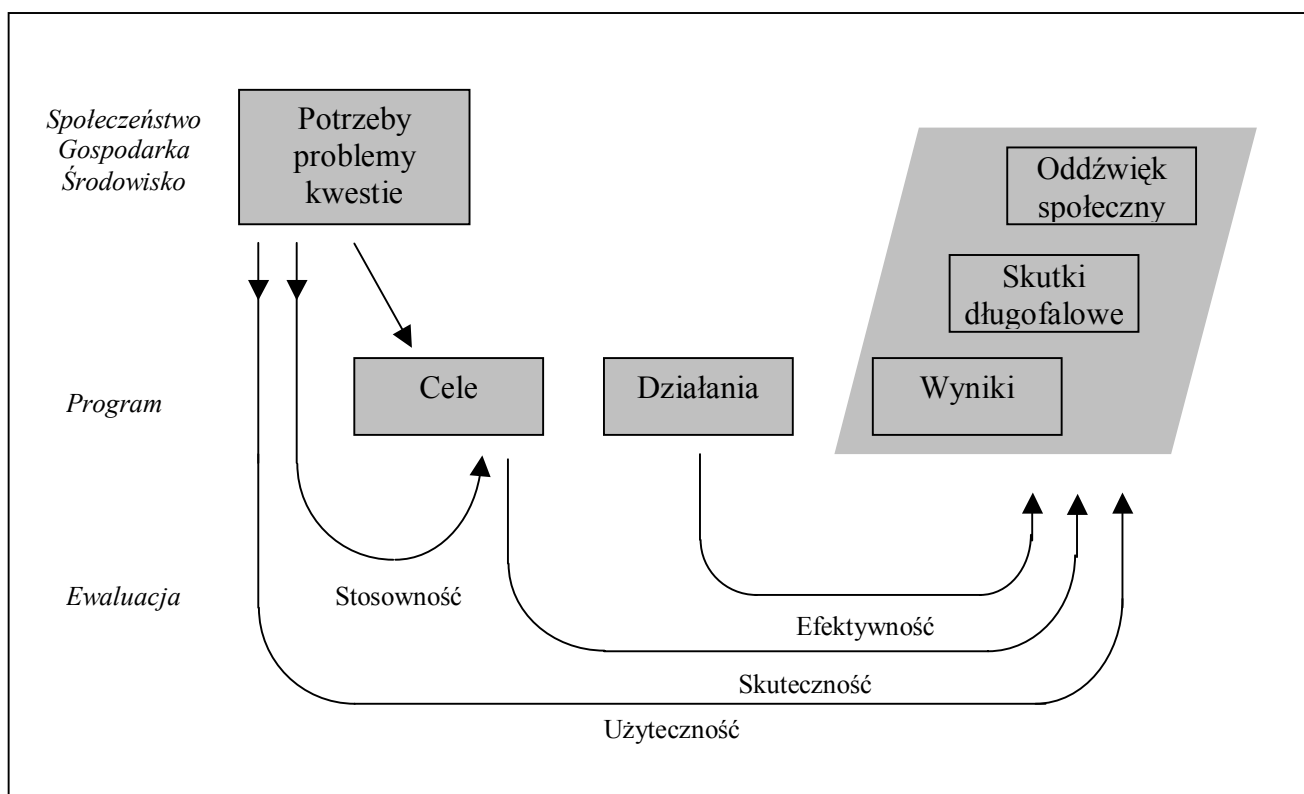
<sup>14</sup> Evaluating Socio-Economic Programmes. Evaluation Design and Management, MEANS Collection, Volume 1, European Commission 1999, s. 44-45

Za pomocą kryterium *odpowiedności* (*relevance*) określa się zgodność celów programu będącego przedmiotem oceny z potrzebami i problemami występującymi w danym sektorze/regionie itd. Kryterium to jest eksploatowane szczególnie w ewaluacjach *ex-ante*, wtedy, gdy zespół ewaluacyjny dokonuje analizy przygotowywanego programu w świetle aktualnej charakterystyki jego otoczenia społeczno-gospodarczego.

Kryterium *użyteczności* jest do pewnego stopnia lustrzanym odbiciem kryterium *odpowiedności*. Różni się od niego o tyle, że jest wykorzystywane do oceny programów już zakończonych lub znajdujących się w trakcie realizacji. Potrzeby otoczenia społeczno-gospodarczego konfrontuje się z osiągnięciami programu – przedmiotu badania.

Ostatnie kryterium – kryterium *długotrwałości* (*perspektywiczności*) służy do określenia, czy efekty badanego programu będą czy też nie będą trwałe w perspektywie średnio- i długookresowej. Opisywane kryteria obrazuje Rysunek 1.<sup>15</sup>

**Rysunek 1**



<sup>15</sup>op. cit., s. 71

Przyjmuje się, że poddanie programu ewaluacji uwzględniającej tych pięć przedstawionych powyżej wątków czyni ją kompletną, opisującą wszystkie podstawowe aspekty jego funkcjonowania i jego oddziaływania na otoczenie.

### 1.5. Standardy jakości badań ewaluacyjnych

Kluczowymi standardami jakości badań ewaluacyjnych jest ich *rzetelność* i *obiektywizm*. Są one zresztą kluczowe nie tylko dla jakości oferowanych przez ewaluatorów usług, ale też dla budowania/utrzymania zaufania szerokiego grona decydentów do wyników tych analiz, szczególnie w sytuacji, gdy mają one implikacje o charakterze ekonomicznym bądź politycznym.

W literaturze przedmiotu przyjmuje się, że standardy, jakie powinny spełniać badania ewaluacyjne, w tym standardy etyczne, mają dwa źródła: teorię sprawiedliwości Johna Rawlsa oraz doświadczenia praktyczne ewaluatorów opisane i podsumowane po raz pierwszy przez Joint Committee on Standards for Educational Evaluation<sup>16</sup>.

J. Rawls szukał uniwersalnej definicji sprawiedliwości w oświeceniowej koncepcji umowy społecznej. Stworzył konstrukcję tzw. sytuacji pierwotnej, w której racjonalnie myślący ludzie uzgadniają sprawiedliwe prawa znajdując się za "zasłoną niewiedzy", nie znając swego miejsca w społeczeństwie, swojego materialnego statusu, ani swych jednostkowych przymiotów i uzdolnień.<sup>17</sup> Rawls zakładał, że tylko w takich warunkach może być wypracowana uniwersalna koncepcja sprawiedliwości. Sprawiedliwość pojmuje, w uproszczeniu, jako bezstronność. Teoria ta prowadzi do stworzenia koncepcji idealnego badacza, którego walorami etycznymi są bezstronność, życzliwość i empatia.<sup>18</sup>

Z praktycznego punktu widzenia za kanon standardów jakości ewaluacji uznać można natomiast zestaw opracowany przez Joint Committee on Standards for Educational Evaluation w 1981 r. i do dziś uważany za aktualny. Zespół ten, składający się z reprezentantów kilkunastu organizacji zawodowych związanych z ewaluacją edukacyjną, wyróżnił 30 standardów zorganizowanych wokół czterech podstawowych kryteriów jakości badań ewaluacyjnych: *przydatności, wykonalności, przyzwoitości oraz rzetelności*.<sup>19</sup>

---

<sup>16</sup> D. Nevo, *Konceptualizacja ewaluacji edukacyjnej. Analityczny przegląd literatury* w: *Ewaluacja w edukacji*, red. L. Korporowicz, Oficyna Naukowa, Warszawa 1997, s. 61

<sup>17</sup> J. Rawls, *Teoria sprawiedliwości*, Wydawnictwo Naukowe PWN, Warszawa 1994, s. 24

<sup>18</sup> *op. cit.*, s. 257

<sup>19</sup> D. Nevo, *Konceptualizacja ewaluacji edukacyjnej. Analityczny przegląd literatury* w: *Ewaluacja w edukacji*, red. L. Korporowicz, Oficyna Naukowa, Warszawa 1997, s. 61; standardy ewaluacji zostały również

*Przydatność (utility)* ewaluacji oznacza, że badania dostarczają informacji istotnych dla programu, są przeprowadzane w czasie pozwalającym na zrobienie użytku z jej wyników i mają wpływ na obiekt ewaluacji. Oznacza to, że ewaluacja ma przynieść praktyczne informacje właściwym odbiorcom. By badanie było przydatne powinien zostać spełniony szereg warunków szczegółowych: m.in. wiarygodność ewaluatora, określenie zakresu badania i selekcja informacji, dokonanie interpretacji kryteriów oceny programu, przejrzystość raportu oraz określenie zasad udostępniania.

Cecha *wykonalności (feasibility)*, drugiego ze wspomnianych standardów jakości ewaluacji, wiąże się ze stosunkiem wartości efektów ewaluacji do poniesionych w związku z nią nakładów. Wykonalność badania związana jest m.in. z dostępnością danych niezbędnych do przeprowadzenia badania, wykonalnością założonych w projekcie ewaluacji procedur i z uwarunkowaniami politycznymi.

Ewaluacja jest *przyzwoita (propriety)*, kiedy zapewnia poszanowanie praw osób zaangażowanych w program i zarazem biorących udział w badaniu. Warunkami spełnienia kryterium przyzwoitości są m.in.: formalny charakter zobowiązań podjętych przez ewaluatorów, panowanie nad potencjalnymi konfliktami interesów grup zainteresowanych wynikami badania, jasne określenie warunków dostępności wyników badania, wyważony sposób raportowania obejmujący mocne i słabe strony przedmiotu analizy, ale także - przedstawianie szczegółowych sprawozdań z wydatków dokonywanych na poczet prowadzenia badania.

Dotrzymanie standardu *rzetelności (accuracy)* oznacza dbałość o prawdziwość wyników ewaluacji. Szczegółowe zagadnienia poruszane w badaniu faktycznie powinny odpowiadać kluczowym kwestiom wyróżnionym w fazie planowania a otrzymane informacje powinny być adekwatne do zgłoszonych przez zamawiających potrzeb. Realizacja kryterium rzetelności zakłada również dokładną identyfikację obiektu badania, dokonanie analizy kontekstu realizacji programu, jasne sformułowanie celów badania i procedur, które mają być tam zastosowane, zapewnienie wiarygodnych źródeł informacji, dokonywanie pomiaru w sposób metodologicznie uprawniony oraz tak by był on godny zaufania i niezawodny, systematyczne kontrolowanie procesu zbierania danych, formułowanie jedynie całkowicie usprawiedliwionych - uprawnionych wniosków oraz obiektywne raportowanie.<sup>20</sup>

---

opracowane przez szereg organizacji profesjonalnych ewaluatorów, m.in. American Evaluation Association: <http://www.eval.org/EvaluationDocuments/progeval.html> (stan z dnia 02.07.2004 r.)

<sup>20</sup> John M. Owen, Patricia J. Rogers, Program Evaluation. Forms and Approaches, SAGE Publications (London - Thousand Oaks - New Delhi) 1999, s. 154 - 158

Wysokie oczekiwania profesjonalne jak i etyczne wobec osób prowadzących badania ewaluacyjne doprowadziły w naturalny sposób do powstania zawodu ewaluatora. Jednak twierdzi się niekiedy, że społeczny kontekst badania istotnie ogranicza jego działania, w tym także te, które mają służyć dotrzymaniu powyżej wspomnianych standardów, a niekiedy nawet całkowicie determinuje efekty pracy badacza. Ewaluatorzy są przecież w sposób formalny lub praktyczny zależni od swych klientów. Zjawisko prezentowania “wdzięczniejszych” niż faktyczne rezultatów ewaluacji zyskało w literaturze przedmiotu nazwę “pozytywnego uprzedzenia” (*positive bias*).<sup>21</sup>

## 1.6. Typy i rodzaje ewaluacji<sup>22</sup>

Terminem *ewaluacja* określa się wiele różnych rodzajów działań tak o charakterze kontrolnym jak i eksploracyjnym. Dla uporządkowania wielości tych zjawisk poniżej przedstawiono kilka najważniejszych typologii ewaluacji, które pozwolą na określanie charakteru i funkcji konkretnych przykładów tych badań stosowanych.

Najczęściej stosowany podział badań ewaluacyjnych zasadza się na ich rozróżnieniu ze względu na pełnioną przez nie rolę w poszczególnych etapach badanego projektu. W tym kontekście wyróżniamy: ewaluację *ex ante*, ewaluację *mid term* oraz ewaluację *ex post*.

*Ewaluacja ex ante* ma miejsce na początku cyklu projektu, przed rozpoczęciem jego realizacji. Jej zadaniem jest zapewnienie, aby założenia programu były wewnętrznie spójne i odpowiadały faktycznemu społecznemu zapotrzebowaniu. Skupia się na analizie mocnych i słabych stron programu, diagnozuje jakość założonej strategii osiągania celów, samych celów, a także wykonalności programu. Ewaluacja *ex ante* stanowi także podstawę dla działań monitorujących oraz przyszłych ewaluacji poprzez określenie mierzalnych wskaźników osiągania celów założonych w programie. W ten sposób w dłuższym horyzoncie czasowym wspomaga proces podejmowania decyzji odnośnie programu. W praktyce realizowana jest w różnych formach – od uproszczonej analizy ekonomicznej do wielowątkowego badania społeczno-ekonomicznego o charakterze eksploracyjnym.

*Ewaluację mid-term* podejmuje się w trakcie (zazwyczaj w połowie) realizacji programu. Badanie analizuje jego pierwsze efekty, jakość zarządzania jego finansami, jakość sposobów implementacji programu oraz prowadzonego monitoringu. Bada też, czy od czasu ewaluacji

---

<sup>21</sup> M. Q. Patton, *Utilization-Focused Evaluation*. The New Century Text, SAGE Publications, Thousand Oaks CA 1997, s. 178

<sup>22</sup> *Evaluating Socio-Economic Programmes*. Evaluation Design and Management, MEANS Collection, Volume 1, European Commission 1999, s. 51

*ex ante* nie zaszły zmiany w otoczeniu ekonomiczno-społecznym programu. Ewaluacja *mid-term* bazuje na informacjach zbieranych w ramach systemu monitoringu opierając się na wskaźnikach wyznaczonych w trakcie badania *ex ante*. Odnosi się także do wyników badań podobnych programów realizowanych w przeszłości, o ile takie istnieją. Ma naturę formatywną – tzn. w zależności od jej wyników mogą być dokonywane poprawki w zakresie sposobów realizacji programu będącego przedmiotem badania. (Wnioski z ewaluacji *mid-term* programów realizowanych przy pomocy funduszy strukturalnych Unii Europejskiej są także jednym z ważnych czynników pozwalających Komisji Europejskiej na podjęcie decyzji w sprawie alokowania tzw. rezerwy wykonania. Zgodnie z rozporządzeniem Rady nr 1260/99, 4% środków z funduszy strukturalnych przyznanych danemu krajowi członkowskiemu na początku okresu programowania nie jest faktycznie programowane, ale zostaje zachowane jako tzw. *rezerwa wykonania*. Służy ona do finansowania tych programów operacyjnych lub ich priorytetów, które w oparciu o zestaw wskaźników, uzgodniony wcześniej pomiędzy Komisją a danym krajem członkowskim lub regionem, wykazują najlepsze wyniki jeśli chodzi o stopień osiągnięcia zakładanych celów, jakość zarządzania oraz postępy we wdrażaniu.)

Ewaluacja *ex post* podsumowuje oraz ocenia program jako całość, szczególnie pod względem jego efektów, po jego zrealizowaniu. Analizuje go przede wszystkim w kategoriach efektywności, skuteczności oraz jakości wykorzystania zasobów przeznaczonych na realizację. Ma też zazwyczaj ambicję formułować konkluzje wykraczające swym zasięgiem poza badany program, takie, które można odnosić do innych programów, innych regionów itd.

O rodzaju ewaluacji decyduje także jej przedmiot. W przypadku ewaluacji programu rodzaj jej określa np. zasięg oddziaływania tego programu. W odniesieniu do programów i projektów finansowanych z funduszy strukturalnych Unii Europejskiej wyróżnia się między innymi: *ocenę wstępną* w odniesieniu np. do selekcji projektów biorących udział w konkursach o dotacje (tzw. *prior appraisal*), *ewaluację dużych projektów* oraz *ewaluację polityk publicznych*. Oznacza to między innymi, że suma ewaluacji poszczególnych projektów w ramach danego programu nie może być traktowana jako ewaluacja tego programu - ewaluacji programu nie można sprowadzać do sumy ewaluacji poszczególnych jego części.

Ocena wstępna zakłada zazwyczaj sprawdzenie, czy jej przedmiot (zazwyczaj aplikacja o dotację na rzecz programu) spełnia kryteria formalne zawarte w ogłoszeniu o konkursie, i jednocześnie wydanie oceny na podstawie zestawu kryteriów, do których należą m.in.: związek z celami programu, przewidywana efektywność, szacunek dla środowiska

naturalnego, innowacyjność, współpraca z sektorem prywatnym i z partnerami lokalnymi oraz komplementarność z innymi projektami.

*Ewaluacja dużych projektów* w świetle rozporządzenia Rady nr 2082/93<sup>23</sup> odnosi się do takich projektów będących realizacją polityki regionalnej Unii Europejskiej, których koszt przekracza 25 mln euro w zakresie projektów infrastrukturalnych i 15 mln euro w zakresie projektów inwestycyjnych. W przypadku ewaluacji tych projektów, obok przedstawionej powyżej oceny wstępnej, analizy kosztów i korzyści podejmowane są szerokie analizy skutków społeczno-ekonomicznych, m.in. przewidywanej stopy zwrotu, przewidywanego wpływu na rozwój lub rekonwersję regionu, perspektyw danego sektora produkcji, wpływu na rynek pracy itd. Istotne jest jednak, że każda decyzja Komisji Europejskiej wobec zgłaszanych projektów – czy to pozytywna czy negatywna – jest rezultatem dialogu z jednostką aplikującą o dotację.

*Ewaluacja polityk publicznych* opiera się na systematycznej ewaluacji realizowanych programów będących narzędziami danej polityki. Służą wspieraniu szeroko pojętego procesu podejmowania decyzji. Szczególnym rodzajem ewaluacji na poziomie polityk jest tak zwana ewaluacja *przekrojowa* (*transversal evaluation*) lub – inaczej - *tematyczna* (*thematic evaluation*). Analizy te określają udział ewaluowanych programów w realizacji priorytetów wspólnotowych. (Na poziomie polityk Unii Europejskiej badania przekrojowe dotyczą zazwyczaj kwestii takich jak: wpływ na rynek pracy, konkurencyjność sektora MSP, równości szans kobiet i mężczyzn, ochrony środowiska itd.) Badanie zakłada analizę kilku programów pod względem realizacji danej zasady/kwestii. Ma zanalizować procesy rządzące jej implementacją w różnych kontekstach.<sup>24</sup>

Przy realizacji badań ewaluacyjnych zaleca się zabieg triangulacji danych - uwzględnienie wielu źródeł danych dla lepszego „oświetlenia” przedmiotu badania. Zakłada on operowanie szerokim zestawem metod badawczych i wykorzystywanie szeregu tzw. *podejść*, czy też *modeli* teoretycznych (*approaches*). Do najbardziej wpływowych podejść w ewaluacji należą: *analiza systemowa*, *podejście zorientowane na podjęcie decyzji*, *studium przypadku* oraz *autoewaluacja*.

*Analiza systemowa* (*system analysis*) – jest określeniem skuteczności, wydajności lub produktywności ewaluowanego programu poprzez ustalenie zależności pomiędzy „wkładem” programu do przedmiotu jego interwencji a wynikami uzyskanymi na końcu realizacji

---

<sup>23</sup> op. cit., s. 57

<sup>24</sup> op. cit., s. 59, por. również Evaluating Socio-Economic Programmes. Evaluation Design and Management, MEANS Collection, Volume 5 (*Transversal evaluation of impacts on the environment, employment and other intervention priorities*), European Commission 1999

programu. Ma odpowiedzieć na pytanie, czy spowodowana przez program zmiana jest zgodna z projektowaną i czy następuje w pożądanym kierunku.

*Podejście zorientowane na podjęcie decyzji* ma ułatwiać decydującym podejmowanie decyzji. Ewaluacja ma przynieść odpowiedzi na sygnalizowane przez nich pytania, oszacować skutki alternatywnych sposobów działania.

W przypadku podejścia zwanego *studium przypadku* uwaga jest skoncentrowana na procesie realizacji programu i na tym jak odbierają go ludzie. Charakterystyczne dla tego podejścia jest nacisk na pozyskiwanie i analizę danych jakościowych. Ma zainteresować odbiorców przedmiotem swojej analizy kładąc nacisk na jego wyjątkowość i złożoność jego uwarunkowań. Powinien uwzględniać pluralizm wartości podzielanych przez grupy zaangażowane w realizację programu.

*Autoewaluacja*, zwana inaczej samoewaluacją, to badanie prowadzone nie przez ekspertów zewnętrznych bądź wewnętrznych, ale przez samych uczestników badanego programu, którego celem jest podniesienie ich własnej wiedzy o programie, i określenie możliwości podniesienia jego jakości.

Zaprezentowane powyżej podejścia, typy i rodzaje badań ewaluacyjnych często nie występują w czystej formie. Postrzegać je należy przede wszystkim jako typy idealne tworzące osie, na których można każdorazowo lokować nowo projektowane badanie dla lepszego zdefiniowania jego celów oraz funkcji, jakie może spełnić na rzecz zamawiającego.

## **1.7. Projektowanie badania ewaluacyjnego**

Zakłada się, że nawet w przypadku obligatoryjnych ewaluacji nie należy lekceważyć znaczenia decyzji o jej przeprowadzeniu. Oznacza ona obowiązek sporządzenia odpowiednio szczegółowego planu badania, w tym szczególnie przedmiotu badania oraz pytań badawczych, które będą organizować późniejszą pracę badaczy.

Konieczne jest określenie obszaru badania pod względem instytucjonalnym (lokalnym, krajowym lub międzynarodowym), sektorowym (np. społecznym, przemysłowym, wiejskim itd.), czasowym i geograficznym. Z praktyki wynika, że na późniejszych etapach badania zakres zainteresowania nim wzrasta, są mu przypisywane nowe cele badawcze, dlatego im dokładniejsze zakreślenie obszaru badania, tym większe prawdopodobieństwo sukcesu projektu badawczego i satysfakcji jego adresatów.<sup>25</sup> Równie strategiczne znaczenie

---

<sup>25</sup> Evaluating Socio-Economic Programmes. Evaluation Design and Management, MEANS Collection, Volume 1, European Commission 1999, s.66

ma precyzyjne określenie celów studium i sporządzenie listy potencjalnych czytelników raportu ewaluacyjnego. Przydatne jest bowiem określenie oczekiwań co do zakresu tematycznego raportu, oraz charakteru wniosków (które mogą mieć walory strategiczne lub być tylko nastawione na pomoc w rozwiązywaniu doraźnych problemów).

Fundamentalne znaczenie przy projektowaniu ewaluacji ma także formułowanie tzw. *pytań ewaluacyjnych*. Mogą mieć one charakter opisowy („co się stało?”), przyczynowe („jak i w jakim stopniu to, co się stało jest związane z badanym programem?”) oraz odwoływać się do oceny zidentyfikowanego oddziaływania programu („czy rezultaty i szersze oddziaływanie programu można ocenić jako satysfakcjonujące?”).

Podstawą relacji pomiędzy zleceniodawcą a zespołem ewaluatorów jest tzw. specyfikacja zamówienia (*terms of reference*). Jakkolwiek może się wydawać, że jej sporządzenie nie jest konieczne w przypadku ewaluacji wewnętrznej (realizowanej przez instytucję odpowiedzialną za projekt poddawany badaniu), to jednak ma ona bardzo istotną wartość porządkującą, określa bowiem szczegółowo oczekiwania zamawiających wobec badania pod względem merytorycznym i technicznym, co pozwala zapanować nad złożoną materią zamówienia. Specyfikacja zamówienia obejmuje zazwyczaj:

- podstawę prawną do przeprowadzenia ewaluacji,
- zakres i cel ewaluacji
- zestaw adresatów ewaluacji oraz planowany sposób wykorzystania danych i raportu z badania
- pytania dotyczące programu, na które ewaluatorzy będą musieli udzielić odpowiedzi po przeprowadzeniu badania oraz kryteria ewaluacji
- zakres informacji o programie, które już są w posiadaniu zamawiającego, i które będą dostępne ewaluatorom
- metody badawcze, których wykorzystanie w badaniu zamawiający uważa za niezbędne,
- założenia programu/projektu, który ma być poddany badaniu,
- przybliżony terminarz wykonania zlecenia,
- przybliżony budżet przeznaczony na wykonanie badania oraz
- niezbędne kwalifikacje ewaluatorów.<sup>26</sup>

Ważne jest, aby określony został tam inicjator badania (autor jego założeń), a także przesłanki stojące za jego uruchomieniem (o ile w danym przypadku ewaluacja nie jest obligatoryjna).

Wskazane jest również jasne określenie zadań, jakie ma spełnić badanie (różnić się będzie to nakierowane na zmodyfikowanie programu od takiego, którego celem jest usprawnienie systemu monitoringu lub relokacja funduszy).

Metody ewaluacji są zazwyczaj negocjowane ze zleceniodawcą na podstawie propozycji zespołu badaczy, preferuje się przy tym pozostawienie im marginesu elastyczności. Decydujący dla wyboru danej metody badawczej bywa niekiedy założony budżet przedsięwzięcia – od niego zależy zarówno ilość analizowanych kwestii jak i stopień zaawansowania tych analiz. Jego wielkość wynika z tzw. podejścia *top-down* – kiedy środki przeznaczone na badanie odzwierciedlają budżet badanego programu lub z podejścia *bottom-up*, które zakłada wstępną estymację kosztów uzyskania odpowiedzi na kolejne wyselekcjonowane pytania badawcze. Ostatecznie zatwierdzony budżet badania stanowi zazwyczaj połączenie obu podejść, przy czym decydującymi czynnikami są najczęściej: liczba pytań badawczych, założony poziom wiarygodności badania, jakość zastanego systemu monitoringu oraz, często ukryte i stąd niedoszacowane, koszty administracyjne.

### 1.8. Najpopularniejsze metody badawcze ewaluacji<sup>27</sup>

Proces realizacji ewaluacji zakłada wykorzystanie czterech podstawowych grup metod: metod *strukturyzacji*, metod *obserwacji*, metod *analizy* oraz metod *wydawania oceny*. To właśnie między innymi ta różnorodność technik analizy i informacji, którymi dzięki nim dysponują ewaluatorzy umożliwia sporządzenie analizy obiektywnej, uwzględniającej punkt widzenia możliwie wielu różnych grup zainteresowanych obiektem badania oraz szeroki zestaw wskaźników społecznych i ekonomicznych.

#### Metody strukturyzacji badania

Charakterystycznymi metodami strukturyzacji są w ewaluacji: *analiza SWOT*, *matryca logiczna* oraz *mapa oddziaływania*.

Najczęściej wykorzystywana w tej fazie jest *analiza SWOT* (akronim pochodzący od słów *strengths* – ang. silne strony, *weaknesses* – ang. słabe strony, *opportunities* – ang. szanse i *threats* – ang. zagrożenia). Silne i słabe strony projektu są tu konfrontowane z czynnikami zewnętrznymi, pozostającymi poza kontrolą władz odpowiedzialnych za jego realizację.

---

<sup>26</sup>op. cit., s. 76

<sup>27</sup> Dokładnego omówienia metod badawczych ewaluacji, szczególnie w odniesieniu do programów realizowanych z funduszy strukturalnych Unii Europejskiej szukać należy w: Evaluating socio-economic programmes. MEANS Collection, Vol. 3 *Principal evaluation techniques and tools*, European Commission 1999

Zaletą SWOT jest przejrzysty sposób prezentacji wyników - prezentowane są w postaci macierzy 2 x 2. Zgodnie z rozporządzeniem Rady 1260/99 analiza SWOT powinna być wykorzystywana przy ewaluacji *ex ante* planów rozwoju, dokumentów programowych, programów operacyjnych i ich uzupełnień. Choć jak wspomniano, jej przeprowadzenie zalecane jest szczególnie w badaniach *ex ante*, wykorzystywana jest także w analizach *mid-term* i *ex post*.

Metoda zwana **matrycą logiczną** (*logical framework*) opiera się na identyfikacji oczekiwanych efektów projektu na poziomie produktu, rezultatu i szerszego oddziaływania, a następnie na przyporządkowaniu każdemu z tych poziomów wskaźników osiągnięć oraz założeń warunkujących te wskaźniki. Matryca powinna odzwierciedlać związek przyczynowo-skutkowy pomiędzy różnymi poziomami celów, wskazywać sposoby weryfikowania, czy zostały one osiągnięte, oraz określić, jakie założenia poza obszarem kontrolowanym przez projekt mogą wpływać na jego udane zakończenie.<sup>28</sup> Dlatego często matryca logiczna staje się *de facto* narzędziem zarządzania poszczególnymi fazami cyklu projektu i instrumentem do tworzenia m.in. budżetu, harmonogramu wdrażania, planu monitorowania itd. Metoda jest skonstruowana w taki sposób, że oczekiwane efekty projektu na danym poziomie wraz z przyjętymi założeniami pozwalają na osiągnięcie efektów na poziomie wyższym. Tak uzyskuje się logiczne powiązanie pomiędzy poszczególnymi poziomami celów. Zgodnie z nazwą metody, do prezentacji efektów projektu wykorzystuje się matrycę. Matryca logiczna wykorzystywana jest głównie w ewaluacji *ex ante*.

**Mapa oddziaływania** (*concept mapping of impact*) jest wykorzystywana do określania oczekiwanych długoterminowych efektów programu. Użyteczna jest szczególnie, gdy zakładane efekty programu nie zostały w dostateczny sposób opisane w dokumentach programowych i istnieje konieczność odtworzenia tych założeń. Stosowanie tej metody opiera się na szerokim udziale partnerów - osób i instytucji zaangażowanych w realizację programu. W trakcie spotkań uczestnicy wraz z ewaluatorami identyfikują efekty programu, porządkują je, grupują ze względu na ich wzajemne związki, a powstałym grupom nadają wagi ze względu na ich znaczenie w osiągnięciu celów programu. Ewaluatorzy analizują otrzymane wyniki i przedstawiają je graficznie w formie mapy. Dzięki zastosowaniu mapy oddziaływania przedstawić można oczekiwane efekty programu, określić ich wzajemne relacje i ocenić wagę poszczególnych efektów danego programu. Metoda ta wykorzystywana jest zarówno w ewaluacji *ex ante* jak i w analizie *ex post*.

## Metody obserwacji

Najpopularniejszymi metodami obserwacyjnymi są w ewaluacji: *wywiady indywidualne, zogniskowane wywiady grupowe, badania ankietowe, studia przypadku oraz obserwacja (sensu stricto)*. Słusznie kojarzyć się one mogą z klasyczną metodologią badań społecznych, jako że ewaluacja przejęła od nich całą teorię i tradycję badań terenowych.

**Wywiady indywidualne**, metoda możliwa do zastosowania na każdym etapie ewaluacji, polega na przeprowadzaniu ustrukturalizowanych rozmów dotyczących badanego programu z osobami zaangażowanymi w projektowanie dokumentów programowych, odpowiedzialnymi za wdrażanie programu, z bezpośrednimi i pośrednimi jego beneficjentami, oraz osobami, które mimo spełnienia odpowiednich warunków nie były zaangażowane w realizację programu. Wywiady wykorzystywane są na wszystkich etapach badania ewaluacyjnego.

Metoda **zogniskowanych wywiadów grupowych**, przydatna we wszystkich rodzajach i etapach ewaluacji, zasadza się na dyskusji grupy, na którą składa się kilka lub kilkanaście osób na różne sposoby zaangażowanych w program, podczas której prowokowane są przez moderatora do przedstawiania swych opinii na temat programu, w tym – do wyrażania sądów wartościujących. Metoda ma na celu uruchomienie kreatywności uczestników dyskusji, która pozwoliłaby na pogłębione zbadanie szczegółowych kwestii poprzez określenie, w jaki sposób odbierają je dyskutanci. Podobnie jak wywiady indywidualne, wywiady grupowe wykorzystuje się we wszystkich etapach studium ewaluacyjnego.

**Badania ankietowe** służą najczęściej do prowadzenia oceny okresowej lub końcowej. Obejmują za to zazwyczaj znacznie większą grupę adresatów niż wywiady indywidualne lub grupowe. Narzędziem badawczym jest kwestionariusz ankiety zawierający jasno sformułowane pytania do ankietowanych odnośnie badanego programu. Trudnością w efektywnym stosowaniu metody jest częsty brak zainteresowania adresatów ankiety jej wypełnieniem. Ważne jest też ustalenie odpowiedniej wielkości badanej próby respondentów – to znaczy takiej, by wyniki były reprezentatywne dla danej populacji. Badania ankietowe wykorzystuje się we wszystkich etapach ewaluacji.

**Studium przypadku (case study)** faktycznie wykorzystuje wiele klasycznych metod badań społecznych jednocześnie (wywiady, zogniskowane wywiady grupowe, obserwacja uczestnicząca itd.). Najczęściej wykorzystuje się je przy ewaluacji określonego, szczegółowego zagadnienia. Do badania wybierane są projekty reprezentatywne dla całego

---

<sup>28</sup> por. broszura pt. Zarządzanie cyklem projektu, Urząd Komitetu Integracji Europejskiej, 2003

programu lub jego komponentu, a następnie są przedmiotem pogłębionej analizy – identyfikacji i interpretacji poddane są mechanizmy oraz procesy, które próbuje się następnie ekstrapolować na cały program. Metodę studium przypadku używa się przede wszystkim w ewaluacjach typu *ex post*.

Odrębną metodą badawczą jest też **obserwacja**. Członkowie zespołu ewaluacyjnego zbierają wtedy dane o realizacji programu poprzez bezpośredni (oficjalny lub anonimowy) udział w działaniach podejmowanych w jego ramach. Metoda ma ograniczony zasięg użyteczności, skądinąd jest jednak klasyczną metodą ewaluacji szczegółowej. Wykorzystywana jest szczególnie w ewaluacji *ex post*.

Dodać należy, że uzupełnieniem uzyskanych przy pomocy opisanych powyżej metod danych są często informacje pochodzące z systemu monitoringu.

### **Metody analizy**

Metody analizy służą oddzieleniu efektów brutto programu (ogół wszystkich efektów związanych z danym programem) od jego efektów netto (tych, które powstały wyłącznie w wyniku wdrożenia programu/projektu). Najczęściej stosowanymi metodami analizy są: analiza *grup porównawczych*, *analiza shift – share*, *model makroekonomiczny* oraz *analiza czynnikowa*.

Metoda **analizy grup porównawczych** zasadza się na porównywaniu dwóch grup podmiotów, z których jedna była beneficjentem danego programu/projektu, a druga pozostawała poza jego wpływem. Jej zastosowanie sprawdza się przede wszystkim w ewaluacji szczegółowej. Problemem tu może być wyselekcjonowanie takich grup, których wszystkie cechy (poza kwestią uczestnictwa w programie) są takie same. Metoda wykorzystywana jest w ewaluacji *ex post*.

**Analiza shift – share**, wykorzystywana do ewaluacji polityki regionalnej, projektuje trendy obserwowalne na poziomie kraju na gospodarkę regionów operując na poziomie makroekonomicznym. Punktem wyjścia metody *shift – share* nie jest analiza konsekwencji programu, ale analiza jego wymiarów makroekonomicznych, tak aby określić, co stałoby się, gdyby badanego programu nie zrealizowano w danym miejscu i czasie. Tzw. stan *policy-on* – to znaczy stan faktyczny, w regionie, będący efektem zrealizowania programu - porównywany jest z tzw. stanem *policy-off*, czyli sytuacją, która miałaby miejsce w regionie,

o ile przebiegałaby zgodnie z ogólnokrajowymi trendami. Podobnie jak analiza grup porównawczych, metoda *shift - share* wykorzystywana jest w badaniu *ex post*.<sup>29</sup>

Celem budowania **modelu makroekonomicznego** jest przedstawienie zasad i warunków funkcjonowania całości gospodarki oraz będących jej częścią rynków (rynek produktów, rynek pracy, rynki finansowe itd.). Model pokazuje funkcjonowanie gospodarki w stanie równowagi oraz dokonuje symulacji ewolucji stanu gospodarki w wyniku bodźców zewnętrznych (np. wahania cen benzyny, wahania kursów walutowych itd.) oraz scenariusze powrotu do stanu pierwotnego.<sup>30</sup> W badaniach ewaluacyjnych umożliwia symulację według scenariusza interwencji za pomocą realizacji programu lub w razie kontynuacji *status quo*. Metodę stosuje się najczęściej w badaniach typu *ex-ante*, poprzez porównanie dwóch scenariuszy właśnie.

**Analiza czynnikowa** jest techniką analizy statystycznej wykorzystywaną do analizy dużych zbiorów danych ilościowych odnośnie grup jednostek, których cechy oznaczone są jako zmienne. Narzędzie to identyfikuje związki i korelacje pomiędzy jednostkami i charakteryzującymi je zmiennymi oraz umożliwia przedstawienie ich w przejrzystej formie graficznej. Wykorzystywane jest zarówno w ewaluacji *ex ante*, jak i w badaniu *ex post*.

## Metody oceny

Popularnymi metodami tej fazy są *panel ekspertów*, *benchmarking* oraz *analiza kosztów i korzyści*.

**Panel ekspertów** wywodzi się z modelu uniwersyteckich rad naukowych i polega na wykorzystaniu wiedzy szeregu niezależnych specjalistów w dziedzinie objętej badaniem. Metoda ta jest szczególnie przydatna w przypadku złożonych i wadliwie zbudowanych pytań badawczych - zapewnia największą możliwą wiarygodność odpowiedzi. Charakterystyczna jest dla badania *ex ante* i *ex post*.

**Benchmarking** oznacza ocenianie efektów programu poprzez ich porównanie z efektami podobnych programów, które zostały uznane za udane.<sup>31</sup> Technika oparta jest na wymianie informacji dotyczących wyników i procesu ich osiągnięcia. Jej celem jest, w procesie porównywania, zidentyfikowanie silnych i słabych stron programu lub organizacji oraz wyszukiwanie nowych rozwiązań, które mogą podnieść jakość osiągniętych efektów.

---

<sup>29</sup> Evaluating socio-economic programmes. MEANS Collection, Vol. 3, European Commission 1999, s. 89-91

<sup>30</sup> op. cit., s. 103-110

<sup>31</sup> więcej informacji na temat techniki *benchmarking* w odniesieniu do badania poziomu usług sektora publicznego znaleźć można m.in. na stronie <http://www.benchmarking.gov.uk/default1.asp>

Celem *analizy kosztów i korzyści* (*cost-benefit analysis*) jest określenie, czy realizacja programu jest pożądana z punktu widzenia danej społeczności. Analizuje pozytywne i negatywne efekty programu, przypisując im ich wartość finansową, w perspektywie interesów poszczególnych grup społecznych, na przestrzeni całego czasu jego trwania. Aby ustalić ostateczną wartość programu, skupia się nie tylko na analizie zmian, które faktycznie spowodował, ale także na analizie scenariuszy alternatywnych wobec realizacji badanego programu. Metoda charakterystyczna jest dla ewaluacji *ex ante*.

## **2. Techniki wzmacniające oddziaływanie badań ewaluacyjnych**

Istnieje bogata literatura przedmiotu poświęcona podejściom, technikom oraz metodom wzmacniania oddziaływania badań ewaluacyjnych na badaną instytucję, program, czy politykę. Wiele z nich postrzega ewaluację wyłącznie jako pomiar określonego stanu czy też etapu rozwoju przedmiotu badania, tj. jako narzędzie służące zwiększeniu efektywności i podwyższeniu jakości danego działania. Coraz powszechniej przypisuje się jednak ewaluacji także funkcje o charakterze głębszym, edukacyjnym. Uznaje się, że może dokonywać trwałych, pozytywnych zmian w badanych organizacjach, przekraczając swym oddziaływaniem ramy programu, który ją wywołał.

Podejścia wzmacniające negocjacyjne funkcje badań ewaluacyjnych, zaprezentowane poniżej, służyć mają, oprócz zapewnienia wynikom badania odpowiedniego oddźwięku ze strony zleceniodawców, prowokowaniu takich właśnie głębszych zmian.

### **2.1. „Dialogiczność” i „demokratyczność” jako zalecane cechy badań ewaluacyjnych**

Co najmniej od początku lat 70. wśród teoretyków ewaluacji w Stanach Zjednoczonych i w Wielkiej Brytanii dominuje teza, że jedynie badania prowadzone z odpowiednią wrażliwością na społeczny kontekst projektowania, realizacji badania oraz prezentacji jego wyników umożliwiają dokonanie dokładnego pomiaru, i że dopiero tak prowadzone badania zbliżają badacza do obiektywności.

Podstawowe znaczenie wydaje się mieć koncepcja ewaluacji dialogicznej (*responsive evaluation*). Dialog jest tu pojmowany na dwa sposoby: jako źródło danych, które są niezbędne do odpowiedniej oceny badanego programu, po drugie jako proces, w którym postrzeganie realizowanego programu przez samych jego realizatorów i uczestników zmienia się pod wpływem wzajemnych konfrontacji poszczególnych tych grup. W tym ujęciu dialog

jest zatem sposobem kreowania, animowania, a często także modyfikowania określonych praktyk.<sup>32</sup> Termin „ewaluacja dialogiczna” został po raz pierwszy użyty przez Roberta Stake’a w 1975 r. Koncepcja ewaluacji dialogicznej, określona poprzez teoretyczne propozycje Roberta Stake’a, Ernesta House’a i Barry’ego MacDonalda, za podstawę przyjmuje zasadę interaktywności i negocjacyjności procesu badawczego. Uznaje się z jednej strony, że negocjacyjny charakter badań ewaluacyjnych jest gwarantem ich wysokiej wartości etycznej. Zakłada się jednocześnie, że otwartość, „uspołecznienie” badania podnosi ostatecznie funkcjonalną, użytkową wartość raportu. W sytuacji, gdy wyniki studiów ewaluacyjnych często nie są akceptowane – i co za tym idzie – wykorzystywane, można uznać, iż ich istotny wkład w procesy decyzyjne może polegać choćby na pełnieniu przez nie funkcji edukacyjnej.<sup>33</sup> W tym kontekście więc ewaluatora należałoby postrzegać jako „nauczyciela, którego sukces należy szacować według tego, czego nauczą się inni”.<sup>34</sup> Choć ewaluacja poprzez decyzje, które podejmowane są na podstawie raportu badawczego, istotnie wpływa na swoich odbiorców, uznaje się, że samo badanie, poprowadzone w atmosferze dialogu, może nabrać wymiaru bardziej uniwersalnego - usprawnić komunikację wewnątrz organizacji, zintegrować zespół wokół zadań, i spowodować, że bardziej się on zidentyfikuje z wyznaczonymi mu celami.

Równoległe, i niejako komplementarnie, wobec podejścia dialogicznego, Barry Mac Donald stworzył w pierwszej połowie lat 70. koncepcję ewaluacji demokratycznej w opozycji do sposobu postrzegania badań oceniających jako instrumentu władzy. Wcześniej celem badania było zazwyczaj skontrolowanie realizatorów ocenianego programu, a raport ewaluacyjny funkcjonował nierzadko jako symbol przewagi kierownictwa nad podwładnymi. Wyniki takiego badania były zazwyczaj poufne. W tym kontekście koncepcja ewaluacji demokratycznej może być uznana za przełomową. „Demokratyczne studia ewaluacyjne to służba na rzecz społeczności, którą informuje się o właściwościach programu (...). Uwzględnia się tu pluralizm wartości podzielanych przez członków społeczności, a w toku formułowania podstawowych pytań badawczych bierze się pod uwagę możliwie szeroki zakres ich interesów. Podstawową wartością dla ewaluatora jest poinformowana opinia publiczna, a jego rola polega przede wszystkim na pośredniczeniu w wymianie informacji

---

<sup>32</sup> por. Guy A. M., Widdershoven, Dialogue in Evaluation: A Hermeneutic Perspective w: Evaluation, SAGE Publications (London - Thousand Oaks - New Delhi), Vol 7(2) 2001, s. 253

<sup>33</sup> por. Helen Simons, Specyfika i rozwój studiów ewaluacyjnych w: Ewaluacja w edukacji, red. L. Korporowicz, Oficyna Naukowa, Warszawa 1997, s. 37

między różnymi grupami.”<sup>35</sup> Stawiało to przed ewaluatorami całkowicie nowe zadania. “Ewaluator powinien dowiedzieć się, jak ludzie reagują na określony program zmiany społecznej i jak go definiują”.<sup>36</sup> Także wykorzystanie i prezentacja zebranych danych miały stać się przedmiotem regularnych konsultacji ewaluatora z badanymi. Miał on pełnić rolę strażnika “prawa do informacji” przy jednoczesnym zachowaniu “prawa do poufności”. To zaś oznaczało potrzebę stałego dialogu i negocjacji pomiędzy grupami osób zaangażowanych w program i w jego ocenę. Przyjęcie tej orientacji metodologicznej może skądinąd implikować istotne zmiany w stosunkach władzy w organizacjach i wkraczać w sferę kultury organizacyjnej badanych instytucji. Faktycznie proponuje ona bowiem podjęcie długofalowego wysiłku demokratyzacji sposobu realizacji programów i zachęca do wypracowania własnych ram dialogu wewnątrz organizacji.

Podsumowując, założenia ewaluacji demokratycznej i dialogicznej można zamknąć w pięciu uniwersalnych postulatach:

- ewaluacja powinna być badaniem, w którym może uczestniczyć szerokie grono uczestników programu/grupy w innym kontekście zainteresowane programem
- metodologia ewaluacji musi być na tyle prosta, by była zrozumiała dla osób/grup w nią zaangażowanych,
- ewaluacja powinna odpowiadać na kwestie kluczowe, a nie dowodzić prawdziwości hipotez,
- wyniki ewaluacji, na podstawie których można doprowadzić do poprawy programu, powinny być w możliwie największym stopniu przedstawiane jako własne wnioski “organizacji uczącej się” (*learning organization*), jako że to realizatorzy badanego programu/projektu, a nie zewnętrzni eksperci czy agencje badawcze, są odpowiedzialni za odpowiednie działanie przedmiotu badania.<sup>37</sup>

---

<sup>34</sup> David Nevo, *Konceptualizacja ewaluacji edukacyjnej. Analityczny przegląd literatury w: Ewaluacja w edukacji*, red. L. Korporowicz, Oficyna Naukowa, Warszawa 1997, s. 53

<sup>35</sup> Helen Simons, *Polityczne implikacje teorii ewaluacyjnych. Przybliżenie problemu w: Ewaluacja w edukacji*, red. L. Korporowicz, Oficyna Naukowa, Warszawa 1997, ss. 82

<sup>36</sup> op. cit. s. 83

<sup>37</sup> por. M. A. Dugan (red.), *Participatory and Empowerment Evaluation: Lessons Learned in Training and Technical Assistance*, SAGE Publications, Thousand Oaks, 1996

## 2.2. „*Process use*”, ewaluacja jako oświecenie oraz *empowerment evaluation*. Mechanizmy wzmacniające wykorzystanie badań ewaluacyjnych

Michael Quinn Patton, w latach 80. prezes Amerykańskiego Towarzystwa Ewaluacyjnego, oprócz wymiernej użyteczności ewaluacji poprzez wspomaganie procesu decyzyjnego w organizacjach, silnie akcentuje także znaczenie jej jako procesu dokonującego się w poszczególnych jednostkach. Według niego udział w badaniu ewaluacyjnym powoduje zwykle także istotne zmiany w sposobie myślenia i zachowania osób zaangażowanych w badanie.<sup>38</sup> Jak pisze Patton, aktywny udział w ewaluacji trenuje w jednostkach umiejętności tak uniwersalne jak: zdolność systematyzowania wiedzy, jasnego wyrażania przekonań, operacjonalizacji pomysłów i celów, a także umiejętność rozróżniania zainwestowanego wkładu (input) od otrzymanego wyniku (output) i odróżniania faktów od sądów i interpretacji. Tzw. *process use* to nauka myślenia „ewaluacyjnego” - nauka o nabywaniu wiedzy w ogólności, pozwalająca radzić sobie ze zmianami zachodzącymi systematycznie w otoczeniu i na trwale zmieniająca sposób patrzenia na rzeczywistość. Co więcej, ewaluacja jako proces w naturalny sposób zatwierdza niejako założone przez program priorytety i wspiera osiągnięte przez niego efekty (zakłada się bowiem zazwyczaj podświadomie, że „jeśli coś poddawane jest badaniu, to znaczy, że to coś jest coś warte”).<sup>39</sup>

Rozważaniom Carol Hirshon Weiss, amerykańskiego teoretyka ewaluacji społecznych nad możliwymi technikami wzmacniającymi wykorzystanie wyników badań ewaluacyjnych towarzyszy przekonanie, że jest ono ograniczane nie tylko względami organizacyjnymi czy politycznymi, ale także immanentnymi cechami samych badanych programów. Programy te pełnią zazwyczaj bardzo złożone funkcje, realizują nieraz wewnętrznie sprzeczne ze sobą cele, angażują różnorodne interesy zaangażowanych grup itd.). Weiss zakłada przy tym, że decyzje polityczne nie są podejmowane jednorazowo, lecz raczej narastają stopniowo jako ciąg małych wyborów i stopniowe zawężanie ilości możliwych do zaakceptowania alternatyw, aż do momentu, kiedy w danych warunkach tylko jedna opcja jawi się jako słuszna. Wobec tego ewaluacja może pełnić rolę katalizatora dokonywanych wyborów, powodując, że podejmowanie decyzji będzie prostsze, bo poparte danymi, racjonalne, „naukowo” ugruntowane, a więc niejako „oświecone“ (*enlightment*

---

<sup>38</sup> M. Q. Patton, *Discovering Process Use w: Evaluation*, SAGE Publications (London - Thousand Oaks - New Delhi), Vol 4(2) 1998, s. 225-226

<sup>39</sup> op. cit., s. 231

*theory of use*)<sup>40</sup>. Ewaluacja działa tu “bezinwazyjnie” – jej wpływ rozłożony jest w czasie, jej oddziaływanie jest jednak trwalsze, głębiej osadzone, uwewnętrznione. Choć nie skutkuje od razu sprawnym rozwiązywaniem konkretnych problemów, uzdalnia do wydawania samodzielnych osądów opartych na wynikach analiz badaczy.

Nieco kontrowersyjnym, ale także interesującym podejściem nakierowanym na zwiększenie siły oddziaływania badań ewaluacyjnych jest podejście zakładające, że najbardziej znaczącymi i pożądanymi zmianami, z punktu widzenia powodzenia danego programu, jakie inicjują badania ewaluacyjne, są takie, które dokonują się w ludzkim wnętrzu i odzwierciedlają zwiększoną zdolność do dokonywania zmian społecznych. Ta zdolność wyraża się raczej w autodeterminacji i samowystarczalności decyzyjnej niż w zmianach, które można by zmierzyć ilościowo. Sukces szacuje się tu miarą zdolności do identyfikowania własnych problemów oraz umiejętnością osiągnięcia consensusu co do możliwości i sposobów ich rozwiązywania. Ewaluacja ma zatem właściwość *wzmacniania* (stąd *empowerment evaluation*) jednostki, grupy, instytucji, i uzdalniania ich do przeprowadzania koniecznych zmian. Zakłada ścisłą współpracę realizatorów programu, jego uczestników oraz ewaluatorów wspólnie zaangażowanych w wypracowywanie zgody co do misji, wizji oraz oczekiwanych rezultatów, porządkowanie dotychczasowych działań w ramach programu oraz czynienie ich bazą dla działań planowanych w przyszłości. Wszystko to prowadzić ma do wykształcenia w grupach i organizacjach rodzaju “wewnętrznej kultury ewaluacji” i naturalnej adaptacji standardów, dzięki którym nabrałyby cech “organizacji uczącej się” (*learning organization*) – otwartej na zmiany, systematycznie dbającej o jak najwyższą jakość swych działań.<sup>41</sup>

### **3. Formy wykorzystania badań ewaluacyjnych w wybranych instytucjach i krajach Unii Europejskiej**

Na tle zarysowanych powyżej najważniejszych teoretycznych założeń i modeli badań ewaluacyjnych poniżej została przedstawiona garść praktycznych doświadczeń instytucjonalnych i narodowych z obszaru Unii Europejskiej. Obrazują one różnorodne zadania, jakie są przydawane ewaluacji w poszczególnych administracjach (np. administracja funduszami strukturalnymi - w Komisji Europejskiej, eksploracja badanie towarzyszące

---

<sup>40</sup>W. R. Shadish, Jr., T. D. Cook, L. C. Leviton, Foundations of Program Evaluation. Theories of Practice, SAGE Publications 1991, s. 196 - 197

<sup>41</sup> por. [www.stanford.edu/~davidf/empowermentevaluation.html](http://www.stanford.edu/~davidf/empowermentevaluation.html) (stan z 28.06.2004 r.)

realizacji wrażliwej społecznie dziedziny polityki rządu - we Francji), jaka rola jest im przypisywana w sferze przepisów prawa i w praktyce. Przedstawione przykłady próbują również pokazać, jakie efekty i jakie trudności wiążą się z ich stosowaniem.

### **3.1. Badania ewaluacyjne w porządku prawnym Unii Europejskiej oraz w praktyce działania Komisji Europejskiej**

#### **Zasady ewaluacji wsparcia strukturalnego w świetle rozporządzenia Rady nr 1260/99 i dokumentów Komisji Europejskiej**

Jedynym dokumentem zobowiązującym do prowadzenia systematycznych badań ewaluacyjnych projektów realizowanych w ramach funduszy strukturalnych Unii Europejskiej i jednocześnie będącym częścią porządku prawnego Unii Europejskiej, jest rozporządzenie Rady nr 1260/1999. Rozróżnia ono trzy formy ewaluacji projektów: *monitoring*, *kontrolę finansową* oraz *ewaluację (sensu stricto – przyp.)*

**Monitoring.** Za prowadzenie monitoringu rozporządzenie czyni odpowiedzialnym jednostkę zarządzającą programem oraz tzw. Komitety Monitorujące. Do kompetencji jednostki zarządzającej należy m.in.: stworzenie systemu zbierania informacji o przebiegu realizacji badanego programu, systematyczne przedkładanie raportów z postępów we wdrażaniu tego programu, organizacja jego ewaluacji *mid term* oraz zapewnienie prawidłowości operacji finansowych i zgodności podejmowanych działań z głównymi kierunkami polityki UE. Jednocześnie realizacja każdego dokumentu programującego fundusze strukturalne (Podstawy Wsparcia Wspólnoty, programy operacyjne itd.) podlega nadzorowi Komitetu Monitorującego, powoływanego za zgodą jednostki zarządzającej przez państwo – beneficjenta wsparcia i Komisję. Jego zadaniem jest m.in. udział w wypracowaniu kryteriów selekcji projektów w ramach poszczególnych działań danego programu, dokonywanie okresowych przeglądów postępów w osiąganiu celów wsparcia, badanie rezultatów wdrażania pomocy, a także proponowanie jednostce zarządzającej dokonywanie dodatkowych przeglądów i ewentualne wprowadzanie poprawek w celu osiągnięcia założonych przez program celów.

Wskaźniki monitoringu powinny być określone we wszystkich podstawowych dokumentach programujących. Powinny opisywać specyficzne cele działań i priorytetów, a także ich wzajemną zgodność i spójność, stan wykorzystania pomocy, rezultaty (*results*), szersze

oddziaływanie (*impacts*) oraz postępy w realizacji planu finansowego. W ramach procesu monitoringu jednostka zarządzająca wraz z Komitetem Monitorującym zobowiązane są do przedstawiania corocznych (lub końcowych – w przypadku projektów trwających nie dłużej niż 2 lata) raportów z procesu wdrażania, które powinny zawierać informacje o postępach tego procesu, w tym także w aspekcie finansowym oraz o działaniach podejmowanych na rzecz zapewnienia jego odpowiedniej jakości i efektywności.

**Kontrola finansowa.** Za kontrolę finansową realizacji projektów, ze strony beneficjenta, odpowiedzialne jest państwo członkowskie. Do jego obowiązków należy przede wszystkim weryfikacja systemu zarządzania i kontroli finansowej, dostarczanie Komisji pełnej informacji o tym systemie, zapewnienie jego zgodności ze stosownymi zasadami Wspólnoty, zapobieganie, wykrywanie i korygowanie nieprawidłowości, a także informowanie Komisji o ich ewentualnym wykryciu i podjętych w związku z tym krokach. W określonych przypadkach do państwa członkowskiego należy również wstrzymanie przyznanego wsparcia i jego redystrybucja.

**Ewaluacja.** Ewaluacja *sensu stricto* w ujęciu wspomnianego rozporządzenia Rady ma na celu zmierzenie efektywności wsparcia strukturalnego, oszacowanie jego oddziaływania w odniesieniu do zaprogramowanych celów oraz analizę uzyskanych efektów. Podstawowymi kryteriami efektywności jest oddziaływanie wsparcia na poziomie realizacji celów Wspólnoty zapisanych w Traktacie (art. 158) (m.in. wzmacnianie społecznej i ekonomicznej spójności UE) oraz oddziaływanie w sferze realizacji priorytetów zapisanych w planach rozwoju, Podstawach Wsparcia Wspólnoty itd. Za zbieranie danych niezbędnych do przeprowadzenia badania odpowiedzialne są państwa członkowskie oraz Komisja. W tym względzie badanie ewaluacyjne opiera się zresztą na danych uzyskanych w procesie monitoringu. Poza ewaluacją *ex ante*, *mid term* oraz *ex post* zarówno państwo członkowskie jak i Komisja może uruchamiać ewaluację o charakterze tematycznym, o ile zebrane w ten sposób doświadczenia miałyby wartość w kontekście realizacji innych projektów. Zakłada się również, że wyniki ewaluacji powinny być, na życzenie, upubliczniane.

Celem ewaluacji *ex ante* jest dostarczenie bazy dla przygotowania planów rozwoju, programów operacyjnych oraz uzupełnień programów. Są za nią odpowiedzialne władze opracowujące wspomniane dokumenty. Badanie obejmuje analizę silnych i słabych stron, potencjału danego państwa członkowskiego, regionu lub sektora, a także zgodności strategii i celów ze specyficznymi cechami regionu i z oczekiwanym oddziaływaniem badanego projektu. Ma za zadanie zweryfikować związki pomiędzy proponowanym systemem wdrażania a założeniami dotyczącymi procesu monitoringu oraz zgodność projektu

z politykami wspólnotowymi biorąc pod uwagę rezultaty ewaluacji z wcześniejszych okresów programowania, a także sytuację w zakresie konkurencyjności, innowacyjności, sytuacji sektora MSP, zatrudnienia i rynku, polityki w zakresie ochrony środowiska naturalnego oraz równości mężczyzn i kobiet.

Ewaluacja *mid term* określa w świetle wyników badania *ex ante*, wstępne rezultaty wsparcia, ich korelację z założonymi celami, a także poziom realizacji tych celów. Ocenia wykorzystanie środków finansowych oraz działanie systemu monitoringu i wdrażania. Za dokonanie ewaluacji *mid term* odpowiedzialna jest jednostka wdrażająca we współpracy z Komisją Europejską i państwem członkowskim. Powinna być prowadzona przez niezależnego biegłego i odnosić się do podstawowych dokumentów programowych.

Ewaluacja *ex-post*, w świetle omawianego rozporządzenia Rady, obejmuje analizę wykorzystania zasobów, skuteczności (*effectiveness*) i efektywności (*efficiency*) wsparcia oraz jego bezpośrednich skutków. Ma jednocześnie za zadanie opracować wnioski odnośnie wpływu zrealizowanych działań na rzecz poprawy spójności społecznej i ekonomicznej oraz zidentyfikować przyczyny sukcesów i porażek w ramach badanego programu. Ewaluacja *ex post* należy do kompetencji Komisji, badanie powinno być jednak dokonywane we współpracy z danym państwem członkowskim – beneficjentem funduszy oraz instytucją zarządzającą badanym programem. Podobnie jak w przypadku analizy *mid term* powinna być ona dokonywana przez niezależnych badaczy.

Jak widać formalne ramy ewaluacji programów w ramach europejskich funduszy strukturalnych są wyraźnie określone co do zasady, jednak ze względu na ich szeroki zakres oddziaływania – są jednocześnie dość ogólne. Wydają się za to odgrywać rolę bazy dla pewnego uniwersalnego standardu jakości zarządzania w ramach Komisji Europejskiej, dotyczącego nie tylko funduszy, ale także realizacji polityk wspólnotowych, działania instytucji wspólnotowych i ich satelitów.

Zresztą, oprócz rozporządzenia Rady 1260/99 wątek potrzeby rozwijania badań ewaluacyjnych podnosi także szereg innych dokumentów Komisji. Należą do nich Rozporządzenie Finansowe (art. 27 (4), 28 (1), 33 (2d) stanowiące, iż „*ewaluacja powinna być stosowana wobec wszystkich programów i działań zakładających znaczące wydatki a wyniki ewaluacji powinny być przekazywane władzom prawnym oraz finansowym*”<sup>42</sup> (wydaje się to być zresztą pierwszy akt prawny UE, podnoszący kwestię ewaluacji działań

---

<sup>42</sup> Financial Regulation, data publikacji 31.12.1977; OJ 356, pełna lista poprawek naniesionych od daty publikacji do znalezienia w bazie prawa europejskiego CELEX

w ramach Wspólnoty) oraz znaczna liczba komunikatów Komisji, przewodniczącego KE oraz poszczególnych Komisarzy. Są to:

- Evaluation of Community Expenditure „Concrete steps towards best practice across the Commission” (SEC 96/659 final)
- Spending more wisely: Implementation of the commission’s evaluation policy SEC (1999) 69/4;
- Focus on Results: Strengthening Evaluation of Commission Activities SEC (2000) 1051-26/07/2000
- Evaluation Standards and Good Practice C (2002) 5267-23/12/2002
- Putting Stratetic Evaluation into Practice within the Commission C (2001) 3661<sup>43</sup>

### **Formy wykorzystania badań ewaluacyjnych w praktyce działania Komisji Europejskiej**

Na temat wykorzystania ewaluacji w pracach samej Komisji Europejskiej zespół ewaluatorów z niezależnych ośrodków badawczych (The Tavistock Institute i Technopolis France) przeprowadził w roku 2002 odrębne badanie. W analizach wykorzystano dokumentację dotyczące ewaluacji przeprowadzonych w ramach lub na zlecenie Komisji, szereg studiów typu *case study*, wyniki badania kwestionariuszowego wśród pracowników Komisji wyższego szczebla oraz wywiady indywidualne.

Z badania wynika, że choć faktyczne miejsce badań ewaluacyjnych w strukturze zarządzania, jak i realny wpływ na sposób działania instytucji są, w zależności od Dyrekcji Generalnej w Komisji, silnie zróżnicowane, to jednak ewaluacja jest w Komisji systematycznie stosowana. Interesujące jest, że pojęcie „wykorzystania” ewaluacji jest przez badanych wyższych urzędników KE rozumiane szeroko. Uznaje się za nie m.in.

- podejmowanie decyzji na podstawie rekomendacji raportu z ewaluacji,
- wykorzystanie ewaluacji jako doświadczenia prowadzącego do lepszego zrozumienia celów działań podejmowanych przez daną jednostkę i większego utożsamienia się z tymi celami,
- traktowanie wyników ewaluacji jako jednego z wielu głosów w danej debacie, a także
- wspieranie organizacji zaangażowanej/angażującej się w badania ewaluacyjne poprzez przyswajanie przez nią specyficznej „kultury ewaluacyjnej” tj. podejścia

---

<sup>43</sup> za: [http://www.europa.eu.int/comm/budget/evaluation/Key\\_documents/evalsystem\\_regureq\\_en.htm](http://www.europa.eu.int/comm/budget/evaluation/Key_documents/evalsystem_regureq_en.htm) (stan w dniu 25.06.2004 r.)

nastawionego na systematyczną, zobiiektywizowaną analizę własnych dokonań w celu podnoszenia jakości podejmowanych w przyszłości działań.<sup>44</sup>

Jak wynika z badania, obszary, w których najczęściej wykorzystuje się studia ewaluacyjne to fundusze strukturalne i pomocowe (przy pomocy podejścia Zarządzanie Cyklem Projektu<sup>45</sup>), obszar badań naukowych (zazwyczaj przy wykorzystaniu zewnętrznych ekspertów dziedzinowych) oraz rozwój i zatrudnienie (przy pomocy szerokiego wachlarza metod łączących doświadczenie nauk społecznych i ekonomicznych<sup>46</sup>.) Jednocześnie obszary zainteresowania ewaluacją systematycznie rosną. Kierunek reform w samej KE zmierza do rozszerzenia obowiązku ewaluacji na coraz to nowe zakresy działań Komisji.

Także miejsce działań ewaluacyjnych w strukturze instytucjonalnej Komisji jest, w zależności od Dyrekcji Generalnej, silnie zróżnicowane. Funkcjonują bądź odrębne komórki organizacyjne zajmujące się koordynacją badań ewaluacyjnych (formuła zinstytucjonalizowana) bądź odpowiedzialność w tej dziedzinie rozłożona jest w ramach danej Dyrekcji na kilka jednostek jednocześnie (formuła horyzontalna).<sup>47</sup>

Z badania wynika przy tym, że im wyższy stopień instytucjonalizacji ewaluacji, tym bardziej traktowane są jako istotne w strukturze organizacyjnej odpowiednie kwalifikacje badawcze oraz doświadczenie pracowników, system procedur odnoszący się do jakości realizowanych badań jest bardziej przejrzysty i wyraźny, a także tym więcej jest sformalizowanych mechanizmów oddziaływania wyników badań na badany obiekt (politykę, program, projekt itd.).<sup>48</sup> Do Dyrekcji o wysokim stopniu instytucjonalizacji ewaluacji należą DG AGRI, DG AIDCO, DB EMPL i DG REGIO.

Dyrekcje o niskim poziomie instytucjonalizacji badań ewaluacyjnych charakteryzują się zazwyczaj niższym poziomem doświadczenia pracowników w zakresie prowadzenia takich badań, tendencją do zacierania granicy pomiędzy ewaluacją a kontrolą finansową i monitoringiem oraz niewielką kontrolą lub brakiem kontroli jakości procedur badawczych. Przykładami tej grupy Dyrekcji są, jak wynika z badania: DG ADMIN, DG COMP, DG ENTR, DG ENV, EUROSTAT, DG JAI, DG MARKET, DG TAXUD, DG TRADE, DB TREN, DG SANCI, SCIC/SDT oraz Sekretariat Generalny.

---

<sup>44</sup> The Use of Evaluation in the Commission Services. Final Report, Kevin Williams, Bastiaan de Laat, Elliot Stern, 2002, s. A3, A4; [http://europa.eu.int/comm/budget/evaluation/studies/study\\_en.htm](http://europa.eu.int/comm/budget/evaluation/studies/study_en.htm) (25.06.2004 r.)

<sup>45</sup> por. broszura pt. Zarządzanie Cyklem Projektu, Urząd Komitetu Integracji Europejskiej, 2003

<sup>46</sup> szereg tych metod zostało skrótowo opisanych w rozdziale 1.9. niniejszego opracowania

<sup>47</sup> The Use of Evaluation in the Commission Services. Final Report, Kevin Williams, Bastiaan de Laat, Elliot Stern, 2002, s. A8 [http://europa.eu.int/comm/budget/evaluation/studies/study\\_en.htm](http://europa.eu.int/comm/budget/evaluation/studies/study_en.htm) (25.06.2004 r.)

<sup>48</sup> op. cit., s. A10 - A11

W grupie Dyrekcji, gdzie rola ewaluacji ma charakter zinstytucjonalizowany, badania ewaluacyjne są częściej niż w innych Dyrekcjach konsultowane z przedstawicielami grup społecznych zainteresowanych badanym programem, częściej organizowane są też podnoszące kwalifikacje szkolenia i seminaria w zakresie typologii i metod ewaluacji. Również pracownicy niższych szczebli w tym Dyrekcjach częściej są angażowani w realizowane tam badania i lepiej poinformowani w zakresie badań już zakończonych, między innymi poprzez praktykę szerokiej prezentacji raportów końcowych. To między innymi powoduje, że w Dyrekcjach tych stopień wykorzystania zrealizowanych ewaluacji badacze oceniają jako większy. Z drugiej strony sugeruje się, że scentralizowany model obecności badań ewaluacyjnych w danej Dyrekcji może sprzyjać negatywnym zjawiskom takim jak izolacja jednostki organizacyjnej zajmującej się tą problematyką od pozostałych jednostek organizacyjnych w strukturze Dyrekcji i tworzeniu się napięć towarzyszących wykonywaniu przez nią funkcji kojarzonych z kontrolą. W przypadku modelu horyzontalnego obecność ewaluacji jest niejako bardziej naturalna, wtopiona w strukturę organizacyjną.

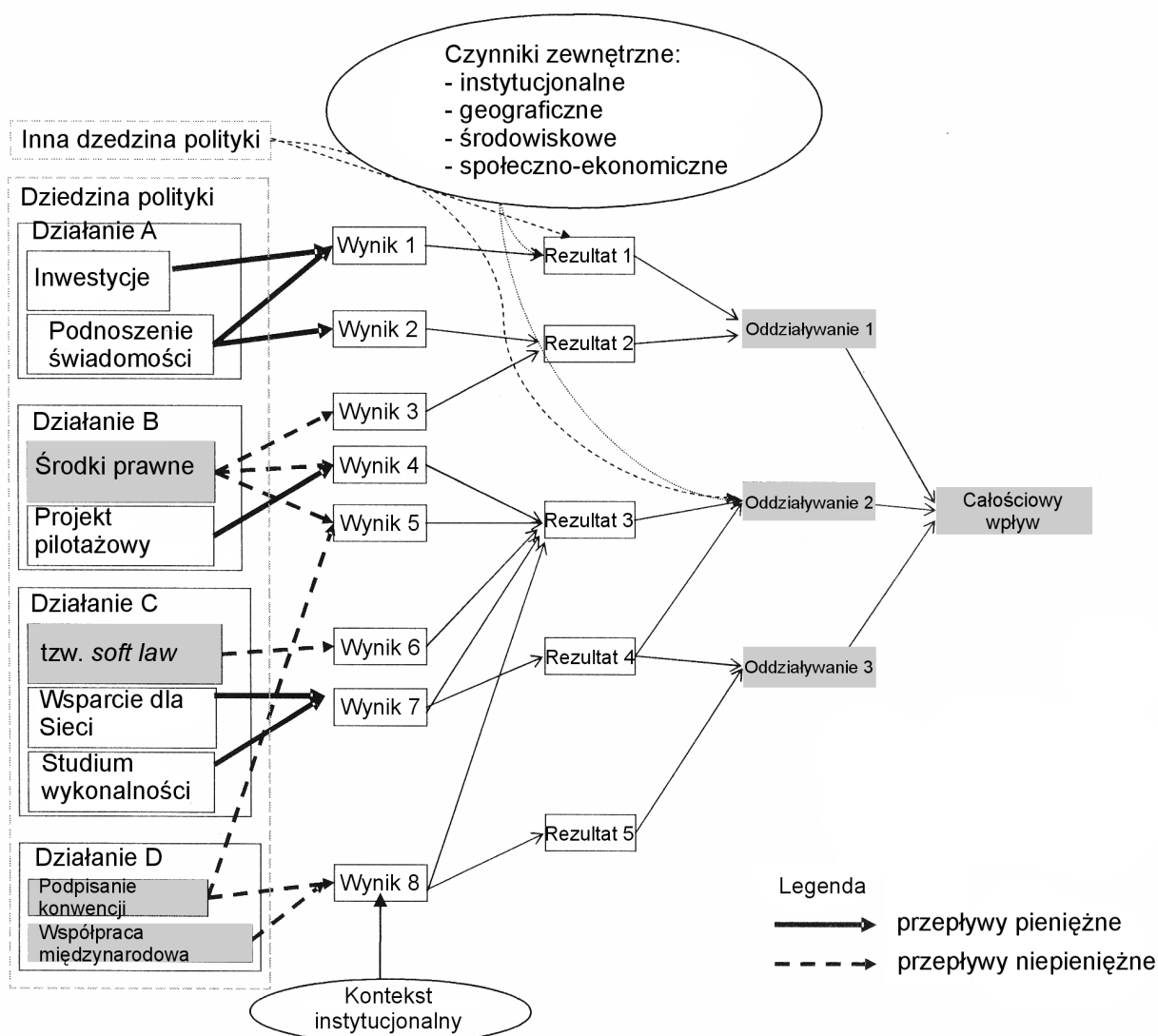
### **Warunki rozszerzania zakresu przedmiotowego badań ewaluacyjnych w ramach instytucji Unii Europejskiej**

Próby rozszerzania zakresu przedmiotowego badań ewaluacyjnych nawet na działalność prawodawczą i realizację *polityk* rozumianych jako strategie lub cele realizowane przy pomocy różnego rodzaju instrumentów kreujących zmianę społeczną można zauważyć w dokumentach Komisji Europejskiej od roku 2000, kiedy to pojawił się pierwszy odnośny dokument kierunkowy – Communication on evaluation, SEC (2000) 1051.<sup>49</sup> Podjęcie wyzwania ewaluacji polityk powoduje jednak potrzebę dokonywania analiz znacznie bardziej złożonych i wyrafinowanych niż w przypadku innych przedmiotów badania. Jako jedno z możliwych podejść do realizacji tego wyzwania Komisja proponuje tzw. analizę udziału (*contribution analysis*), która określa udział danego działania w wywołaniu określonej zmiany. Kluczowym elementem jest w tej metodzie wyznaczenie tzw. mapy logiki oddziaływania, której przykład został przedstawiony na Rysunku 2.

---

<sup>49</sup> por. Evaluating EU Activities. A Practical Guide for the Commission Service; European Commission; listopad 2003; [http://www.europa.eu.int/comm/budget/evaluation/pdf/evaluation\\_guide\\_final.pdf](http://www.europa.eu.int/comm/budget/evaluation/pdf/evaluation_guide_final.pdf) (stan z 24.06.2004 r.) oraz SEC (2000) 1051, SEC (2001) 1197/6 i 7, rozporządzenie Rady (SEC (2002) 1605) i rozporządzenie Komisji (SEC (2002) 2342)

Rysunek 2<sup>50</sup>



Rysunek 2 przedstawia ilustrację hipotetycznej dziedziny polityki wraz z jej krótko, średnio i długo terminowymi efektami.<sup>51</sup> Przyjmuje się, że polityka ta obejmuje cztery działania: pierwsze – składające się z dwóch programów finansowych, drugie - składające się z dwóch instrumentów pozafinansowych, oraz trzecie i czwarte - stanowiące połączenie formalnych instrumentów prawnych oraz tzw. *soft law* (aktów nie mających mocy wiążącej). Rysunek ukazuje schemat potencjalnych efektów polityki (w prostokątach) oraz związków pomiędzy poszczególnymi efektami (strzałki). Pokazuje także jak wraz z przejściem z poziomu efektów krótkookresowych do poziomu efektów średnio- i długookresowych coraz trudniej jest określić faktyczne zależności pomiędzy podejmowanymi działaniami

<sup>50</sup> Evaluating EU Activities. A Practical Guide for the Commission Service; European Commission; listopad 2003; [http://www.europa.eu.int/comm/budget/evaluation/pdf/evaluation\\_guide\\_final.pdf](http://www.europa.eu.int/comm/budget/evaluation/pdf/evaluation_guide_final.pdf) (stan z 25.06.2004 r.); s. 75

<sup>51</sup> op. cit., s. 76

a obserwowanymi zmianami (rozumianymi zarówno wąsko jak i jako gruntowna, wielowymiarowa zmiana w funkcjonowaniu danego systemu). Dla łatwiejszego zrozumienia treści rysunku można wyobrazić sobie sytuację wejścia w życie nowej regulacji, np. dyrektywy Rady (działanie b nr 1), która ma na celu poprawę jakości powietrza (*impact*) poprzez ograniczenie emisji trujących gazów (*result*). Wdrożenie szeregu instrumentów szczegółowych takich jak ilościowe limity zanieczyszczenia oraz nałożenie kar na łamiących nowe regulacje na rysunku określa się w tym kontekście terminem *output*.

Warto zauważyć, jak duże znaczenie przydane zostało w analizie udziału czynnikom zewnętrznym (*exogenous factors*). Rysunek sugeruje, że bardzo silnie determinują one przenikanie bezpośrednich efektów prowadzonej polityki do otoczenia społeczno-ekonomicznego. To oznacza, że przewidzenie faktycznych zależności pomiędzy działaniami podejmowanymi w ramach danej polityki a ich ostatecznymi rezultatami jest bardzo trudne. Wyniki badania ewaluacyjnego mogą być zatem w tym przypadku obciążone istotnymi błędami. Można nawet sugerować, że faktycznie oferują one jedynie *teorię* działania (*theory of action*) lub *teorię* programu (*program theory*).<sup>52</sup>

Ogólnie rzecz biorąc ewaluację polityki w stosunku do ewaluacji programu wyróżnia jednak przede wszystkim: wyższy stopień złożoności założonych celów, większa ilość adresatów działań, większy nacisk na demonstrowanie procesu osiągania celów oraz większa złożoność wzajemnych zależności pomiędzy podejmowanymi działaniami a ich kontekstem społeczno-ekonomicznym (w tym także kontekstem innych polityk).

Próby aplikowania technik i standardów ewaluacji na zjawiska tak złożone jak polityka dziedzinowa państwa każą zastanowić się z jednej strony nad perspektywami rozwoju i wpływów ewaluacji jako jednej z technik wspomagających zarządzanie organizacjami, a z drugiej - nad granicami jej stosowalności i prawomocności wyników.

### **Formy wzmocnienia roli ewaluacji w Komisji Europejskiej**

Działania Komisji Europejskiej na rzecz wzmocnienia roli badań ewaluacyjnych obejmują między innymi szereg praktyk włączających je do grupy jej działań rutynowych. Komisja promuje wykorzystanie wyników ewaluacji prowadzonych i zleczanych przez poszczególne Dyrekcje Generalne przy tworzeniu cyklu Strategicznego Planowania i Programowania (*Strategic Planning and Programming cycle*), w którym ujęte są strategiczne priorytety i cele KE. Zgromadzony materiał ewaluacyjny również jest bazą

informacyjną przy corocznym opracowywaniu tzw. Rocznej Strategii Działań (*Annual Policy Strategy*) Komisji, a także przy dokonywaniu modyfikacji w budżecie UE. Charakteru strategicznego nabiera także planowanie badań ewaluacyjnych. Jak wynika z zapisu wystąpienia przedstawiciela DG Budżet na Spotkaniu Ekspertów w zakresie Ewaluacji – przedstawicieli krajów członkowskich, z marca 2003, na rok 2004 zaplanowano 43 badania tego typu, na rok 2005 – 30, na 2006 r. – już 11, a na 2007 r. – 5.<sup>53</sup>

Komisja Europejska jest również w znacznej mierze odpowiedzialna za uwrażliwienie administracji nowych krajów członkowskich na kwestię ewaluacji. Dla ewaluacji programów realizowanych w ramach funduszu PHARE powstała na Węgrzech Krajowa Jednostka Ewaluacyjna (*National Evaluation Unit*), podobne jednostki z identycznego powodu powstały m.in. w Czechach i w Estonii.<sup>54</sup> Polski odpowiednik - Krajowa Jednostka Oceny - powstała w 2004 r.

Komisja inicjuje również wiele działań zmierzających do popularyzacji badań ewaluacyjnych wśród swych agend i swych beneficjentów. Największy udział w tym procesie wydaje się mieć DG Budżet. Prowadzi bogatą w informacje podstronę internetową poświęconą obecności ewaluacji w strukturach Komisji<sup>55</sup> i promuje „dobre praktyki” ewaluacyjne. DG Budżet koordynuje również działanie zespołu ewaluacyjnego (*evaluation team*), który czuwa nad realizacją badań zamawianych/realizowanych przez Komisję, udziela merytorycznego wsparcia w trakcie realizacji badań i działa na rzecz wzmocnienia wizerunku badań ewaluacyjnych w administracji UE. Całość tych działań bywa nawet traktowana jako odrębna polityka Komisji (nazywana *policy of evaluation*).

### **3.2. Ewaluacja polityk państwa we Francji na przykładzie badania stanu realizacji polityki w zakresie bezpieczeństwa drogowego**

Tendencją panującą w zarządzaniu finansami publicznymi we Francji jest stopniowe odchodzenie od prowadzenia polityk dziedzinowych w trybie ciągłym i na sposób resortowy na rzecz operowania w ramach programów – wyodrębnionych jednostek realizacji polityki państwa w danej dziedzinie, zakładających osiągnięcie w danym czasie szczegółowo określonych oraz możliwych do zmierzenia celów. Podejście takie zakłada konieczność

---

<sup>52</sup> tamże

<sup>53</sup> *Meeting of the Evaluation Experts of Member States Administration; Brussels, 27 March 2003;* [http://europa.eu.int/comm/budget/evaluation/externalnetwork/minutes\\_2003\\_03\\_27.pdf](http://europa.eu.int/comm/budget/evaluation/externalnetwork/minutes_2003_03_27.pdf), s. 3 (stan z 25.06.2004 r.)

<sup>54</sup> ibidem

dokonywania systematycznego monitorowania przebiegu realizacji tych programów, co z kolei wymaga opracowania kompletnego zestawu odnośnych wskaźników. Aczkolwiek wspomniane warunki organizacyjne stanowią swego rodzaju typ idealny, nieobecny w rzeczywistości, a jedynie szkicujący schemat procedury, zarządzanie poprzez realizację programów zyskuje na popularności, i bywa aplikowane także na poziomie realizacji polityk na poziomie krajowym.<sup>56</sup> W ten sposób wzmocniona zostaje rola badań ewaluacyjnych, które przyjmują bardzo ważną instytucjonalną rolę kontrolną.

Przykładem zastosowania badań ewaluacyjnych w takim kontekście są kolejne edycje ewaluacji działań administracji państwowej w dziedzinie bezpieczeństwa drogowego we Francji. Pierwsze badanie ewaluacyjne w tym zakresie podjęte zostały we Francji w 1968 r. w ramach racjonalizacji budżetu państwa wzorowanej na systemie planowania budżetu Stanów Zjednoczonych.<sup>57</sup> Nowe badanie uruchomiono w październiku 2000 r. na podstawie decyzji międzyresortowego Komitetu ds. reformy państwa w dobrym klimacie instytucjonalno-organizacyjnym. Komitet kierujący realizacją badania (tzw. *le comite de pilotage*) składał się z przedstawicieli administracji (reprezentantów Generalnego Komisariatu Planu, Krajowej Rady ds. Ewaluacji) oraz niezależnego Krajowego Centrum Badań Naukowych (CNRS). Badanie zaprojektowano jako obejmujące jednocześnie krajowy oraz lokalny poziom zjawiska wypadków drogowych i wykorzystujące zarówno metody badań jakościowych jak i ilościowych. Na potrzeby studium przeprowadzono m.in. ok. 200 wywiadów indywidualnych z dobranymi celowo przedstawicielami wybranych grup zawodowych i społecznych (analizy socjologiczne i psychologiczne) oraz przeanalizowano ok. 5000 przypadków naruszenia prawa (od momentu naruszenia prawa do wymierzenia kary). Szereg analiz na potrzeby raportu wykonali eksperci zewnętrzni (agencja Deloitte & Touche). W realizacji badania wzięło udział również ok. 120 ekspertów różnego typu, m.in. członków tematycznych grup roboczych, badaczy i ekspertów indywidualnych związanych z różnymi tematycznymi grupami refleksyjnymi, sprawozdawców odnośnego stanowiska rządowego itd. Celem było w równej mierze zbadanie efektywności realizowanej polityki

---

<sup>55</sup> [http://www.europa.eu.int/comm/budget/evaluation/index\\_en.htm](http://www.europa.eu.int/comm/budget/evaluation/index_en.htm) (stan z 25.06.2004 r.)

<sup>56</sup> por. tekst wystąpienia p. Serge Arnaud, delegowanego w zakresie modernizacji zarządzania usługami publicznymi i reformy państwa, podczas *European Learning Laboratory. The Reasoned Use of Performance Indicators in the Evaluation of Public Policies (application to the subject of Road Traffic Safety)*, 12 grudnia 2003 r.

<sup>57</sup> por. tekst wystąpienia p. Michela Ternier, inżyniera generalnego dróg i mostów podczas *European Learning Laboratory. The Reasoned Use of Performance Indicators in the Evaluation of Public Policies (application to the subject of Road Traffic Safety)*, 12 grudnia 2003 r.; zob. również *La politique de securite routiere. Les systemes locaux de controle- sanction*, President Michel Ternier, La documentation Francaise 2003

(badanie o charakterze *ex post* jak i sformułowanie rekomendacji dla zwiększenia efektywności aktualnie prowadzonych działań w tej dziedzinie.

Charakter badania określały cztery główne założenia. Miało być ono autonomiczne (wobec realizatorów badanej polityki oraz pozostałych grup interesu) i użyteczne (stąd eksploatowano techniki służące budowaniu zaufania społecznego oraz sprzyjającego klimatu w zaangażowanych w ewaluację instytucjach wokół jego realizacji). Po trzecie, zakładano uwzględnienie opinii i ocen szerokiego spektrum grup zainteresowanych realizacją polityki, tak by uzyskać pełny i obiektywny obraz zjawiska. Po czwarte, dla zapewnienia komfortu pracy badaczy oraz zapobieżenia negatywnemu wpływowi opinii publicznej na prowadzenie studium realizacja badania miała charakter poufny – pierwszy raport z badań powstał 2 lata od uruchomienia programu badawczego. Taka konstrukcja badania pozwalała ewaluatorom na dużą elastyczność pod względem analizowanych zagadnień oraz doboru metod badawczych. Elastyczność zadecydowała też o jakości badania. Z punktu widzenia realizacji badanej polityki był to bowiem okres bardzo dynamiczny. W lipcu 2002 r. nowo wybrany prezydent Francji Jacques Chirac określił bezpieczeństwo drogowe jako jeden z trzech priorytetów jego rozpoczynającej się pięcioletniej kadencji. Deklaracja wywołała szeroką debatę w mediach i sprowokowała do systematycznego badania związku poziomu publicznego zainteresowania tą tematyką z faktycznym poziomem bezpieczeństwa drogowego, mierzonym w liczbie wypadków drogowych, ich ofiar itd. Znacznie zaostrzono wtedy również przepisy ruchu drogowego. Jednocześnie zwiększyła się też ilość kontroli kierowców.

W badaniu wykorzystano rozbudowaną baterię wskaźników obejmującą kontekst sytuacyjny występowania wypadków, wpływ wdrażania nowych regulacji prawnych (odnośnie obowiązywania pasów bezpieczeństwa, ograniczeń prędkości czy dopuszczalnej zawartości alkoholu we krwi kierowcy) na zachowania społeczne oraz wskaźniki o charakterze gospodarczo – ekonomicznym (m.in. analizujące wpływ stanu bezpieczeństwa drogowego na stan gospodarki). Analizowano również m.in. funkcjonowanie systemu ubezpieczeniowego, odnośne procedury policji, system ścigalności kar czy prawdopodobieństwo poddania kierowcy kontroli trzeźwości. Ponadto, mierzono wynikami badań opinii publicznej wskaźniki społecznego zadowolenia z podejmowania przez rząd i administrację wymienionych powyżej działań. Końcowy raport ewaluacyjny opublikowano w kwietniu 2003 r. Odpowiednio do szeroko zakrojonych analiz, zamieszczone tam wnioski oraz rekomendacje dotyczyły nie tylko bezpośrednich działań systematycznie realizowanych w celu zapewnienia bezpieczeństwa drogowego, ale także właśnie sfery regulacji prawnych,

otoczenia instytucjonalnego i systemu kompetencji, warunków egzekwowania i skuteczności obowiązujących regulacji karnych i systemu działań komunikacyjnych.

Wraz z postępami badania ewaluacyjnego systematycznie zaczęła we Francji spadać liczba śmiertelnych ofiar wypadków drogowych. Podczas gdy w roku 2001 było ich 7720, rok później było ich o 6,2 p.p. mniej (7242). W 2003 roku osiągnięto aż 20 punktowy spadek tego wskaźnika (20,9 p.p.), a poziom szacunkowy wyniku dla roku 2004, obliczony na podstawie okresu od stycznia do kwietnia br. zakłada kolejny osiemnastopunktowy spadek.<sup>58</sup> Aczkolwiek nie można łączyć tego sukcesu z równoległą sytuacją badania lecz z podjętymi przez administrację środkami dyscyplinującymi kierowców, przyjaznym klimatem politycznym wokół zaostżenia przepisów oraz z ożywioną debatą publiczną, analitycy z Commissariat General du Plan twierdzą, że realizowane badanie ewaluacyjne miało silny wpływ na zaangażowanie wszystkich istotnych grup społecznych w zaplanowane w badanej dziedzinie działania, i to właśnie ta mobilizacja faktycznie tę debatę podtrzymywała drażąc niejako zmianę społeczną.<sup>59</sup> Analizy zastanych schematów organizacyjnych zmuszały do ich rewizji, inicjowały refleksje nad możliwościami poprawy efektywności już podejmowanych działań. Stworzyły przyjazny klimat w szerokich sferach zarządzania badaną polityką.

### **3.3. Ewaluacja jako standard zarządzania w administracji na przykładzie Wielkiej Brytanii**

Wielka Brytania jest krajem, w którym zarówno pod względem ilości studiów teoretycznych jak i praktyki badawczej ewaluacja wydaje się być zdecydowanie najbardziej w Europie popularna. Z drugiej strony uważa się, że istotne wzmocnienie znaczenia ewaluacji w praktyce administracji publicznej nastąpiło dopiero wraz z dojściem do władzy Partii Pracy w maju 1997 r.<sup>60</sup> Hasłem wyborczym tej partii była wtedy obietnica poprawy efektywności systemu usług publicznych. Jej realizacja poskutkowała gwałtownym wzrostem wydatków publicznych na modernizację placówek edukacyjnych, systemu opieki zdrowotnej, usług publicznych, a także samej administracji centralnej. Z drugiej strony sposób wdrażania reform

---

<sup>58</sup> La securite routiere en France. Bilan de l'anne 2002; La documentation Francaise, 2003; s. 52

<sup>59</sup> wystąpienie Claudine Perez – Diaz, badaczka Krajowego Centrum Badań Naukowych podczas spotkania *European Learning Laboratory. The Reasoned Use of Performance Indicators in the Evaluation of Public Policies (application to the subject of Road Traffic Safety)*, 26 kwietnia 2004 r.; zob. również La politique de securite routiere. Les systemes locaux de controle- sanction, President Michel Ternier, La documentation Francaise 2003; s. 197 - 201

był próbą realizacji kluczowych zasad tzw. *New Public Management (NPM)*. (Doktryna ta zakłada decentralizację administracji, przy jednoczesnym zwiększeniu kontroli merytorycznej i finansowej państwa, rozumianej między innymi jako kontrola sprawowana bezpośrednio przez obywateli-konsumentów. Zadaniem ministrów jest w niej jedynie wyznaczenie strategicznych, a jednocześnie wyraźnie określonych za pomocą wskaźników, celów poszczególnych polityk. Powinni być oni natomiast rzecznikami owych polityk wobec opinii publicznej. *New Public Management* zakłada również, że usługi, za których dostarczenie obywatelom państwo jest odpowiedzialne, miałyby zostać częściowo urynkowane i pozostawione w sferze kompetencji menedżerów „liniowych” – tj. wykonawczych<sup>61</sup>).<sup>62</sup> W przypadku Partii Pracy charakterystyczne było uzupełnienie modelu *NPM* o specyficzne podejście zakładające ścisłą współpracę rządu z partnerami lokalnymi (m.in. realizatorami, odbiorcami), realizowane poprzez uruchamianie licznych projektów pilotażowych, powoływanie grup konsultacyjnych itd. (*collaborative approach*). To stworzyło nowe wyzwania dla całej administracji. Konieczne stały się większa otwartość i kreatywność, a z drugiej strony ściśle związanie podejmowanych działań z wynikami dokonywanych kontroli (*evidence based policy making*).<sup>63</sup> Zmienić miał się przy tym charakter dotychczasowych kontroli - z działań mierzących stosunek wkładu (*input*) do bezpośredniego wyniku (*output*) danego działania przekształcić się miały w studia nad szeroko pojętym oddziaływaniem tych działań (na sferę społeczną, gospodarczą itd.).

Przykładem aplikacji wspomnianego nowego podejścia jest realizacja pilotażowego programu *Best Value* (1998 - 2000 r.), który w miejsce wcześniejszych regulacji przewidujących uniwersalną procedurę rozpisywania przetargów na usługi zamawiane przez administrację lokalną wprowadził w wybranych jednostkach administracji niższego szczebla obowiązek dokonywania gruntownych i systematycznych analiz realizowanych usług poprzez:

- badanie przesłanek do podjęcia danego działania,
- porównywanie jakości wykonania usługi z tą oferowaną przez innych dostawców obecnych na rynku,
- porównanie konkurencyjności cenowej usługi w stosunku do pozostałych ofert oraz

---

<sup>60</sup> por. S. Martin, I. Sanderson, *Evaluating Public Policy Experiments. Measuring Outcomes, Monitoring Processes or Managing Pilots?*, Evaluation. SAGE Publications, Vol 5(3), 1999; s. 245-246

<sup>61</sup> w odróżnieniu od „sztabowych”, których zadaniem jest planowanie działań, ich koordynacja itd.

<sup>62</sup> *Public Management Reform. A Comparative Analysis.*, Christopher Pollit and Geert Bouckaert, Oxford University Press, 2000 s. 137, 140

<sup>63</sup> *Evaluating Public Policy Experiments. Measuring Outcomes, Monitoring Processes or Managing Pilots?*, Steve Martin, Ian Sanderson, Evaluation 5 (3), SAGE Publications, 1999 s. 246

- systematyczne konsultacje odnośnych potrzeb zgłaszanych przez bezpośrednich adresatów usługi, innych zainteresowanych grup oraz szeroką opinię publiczną.

Założenia programu przypominały w swym ogólnym zarysie elementy badania ewaluacyjnego – jego realizacja stanowiła więc faktycznie stworzenie obowiązku dokonywania jakiejś formy ewaluacji *wszystkich* usług zamawianych przez te jednostki.

Także sam program poddany został ewaluacji. Przed badaniem, mającym charakter kroczący – tzn. postępujący równocześnie z realizacją badanego programu, postawiono dwa zadania. Miało ono zidentyfikować zakres faktycznie osiągniętej poprawy jakości i efektywności dostarczanych usług (o ile miała ona miejsce), ale także zidentyfikować najlepsze przykłady aplikacji nowej regulacji w celu określenia warunków sprzyjających osiąganiu najlepszych wyników, co zakładało konieczność dokonania pogłębionych analiz jakościowych. Bez względu na liczne trudności metodologiczne omawianego badania, przeprowadzona ewaluacja zdołała nie tylko wąsko określić rachunek podjętych w ramach programu działań (opierając się na schemacie tzw. trzech E – *economy* – rachunek ekonomiczny, *efficiency* – efektywność oraz *effectiveness* - skuteczność), ale także dostarczyć przykładów pozytywnych wyników zastosowania nowej regulacji (niższe koszty usług dla budżetu przy utrzymaniu ich dotychczasowej jakości) i przedstawić je w formie możliwej do zaprezentowania szerokiej opinii publicznej (realizując zestaw tzw. trzech D – *developing* – rozwój, *demonstrating* - prezentacja i *disseminating* - promocja).<sup>64</sup> Widać wyraźnie, że tradycyjnemu podejściu do kontroli realizowanych projektów towarzyszyły tu elementy podejścia konsultacyjnego, dialogicznego, promującego specyficzną „kulturę ewaluacyjną” wśród instytucji, grup i poszczególnych osób zaangażowanych w badanie.

Tendencja, jaką jest wzrost znaczenia badań ewaluacyjnych w praktyce administracji publicznej w Wielkiej Brytanii w ostatnich latach, wykorzystywanie jej na wielu poziomach procesu zarządzania, obecnie jest także stymulowana poprzez opublikowanie przez brytyjskie Ministerstwo Skarbu w 2003 r. *Zielonej Księgi. Ocena i Ewaluacja w Administracji Centralnej*.

Dokument ten zakłada, że „*Wszystkie nowe polityki, programy oraz projekty, czy to dotyczące kwestii skarbowych, związanych z wydatkowaniem pieniędzy publicznych, czy prawnych, powinny być poddawane ogólnej, acz proporcjonalnej ocenie, tam gdzie to możliwe, w celu*

---

<sup>64</sup> op. cit., 254-255

*promowania interesu publicznego.*<sup>65</sup> Kładzie przy tym nacisk na prowadzenie badań ewaluacyjnych w charakterze dialogu z grupami zainteresowanymi programem (choć wartość podejścia dialogicznego pojawia się tam przede wszystkim w kontekście wspomagania procesu zbierania danych niezbędnych do przeprowadzenia badania). Jako istotny problem postrzega się w dokumencie ponadto dostarczenie metodologicznych instrumentów i mechanizmów instytucjonalnych zapewniających wykorzystanie wyników badań ewaluacyjnych. Dla wzmocnienia pozycji wyników ewaluacji zaleca m.in.: tworzenie odrębnych jednostek organizacyjnych zajmujących się ewaluacją w ramach większych instytucji, konsultowanie uruchomienia danego studium w gronie zainteresowanych badaniem resortów i grup oraz tworzenie archiwów raportów z badań już zrealizowanych.

Zielona Księga podkreśla również kluczowe znaczenie skrupulatnego formułowania celów planowanych przez administrację działań. Podkreśla się, że powinny być one „wyraziste” (org. *SMART* – akronim pochodzący od słów *specific* - szczegółowe, *measurable* - mierzalne, *achievable* - osiągalne, *relevant* – stosowne, trafne i *time-bound* – określone w czasie). Także skutki realizowanych programów powinny być obserwowane w nowej perspektywie. Księga dzieli je na bezpośrednie (*outputs*) i te rozumiane bardziej szeroko (*outcomes*) i w ten sposób sugeruje potrzebę ich dokładniejszego pomiaru i analizy.<sup>66</sup>

Z drugiej strony, w opisywanym dokumencie utrzymuje się błędnie, że ocena (o charakterze finansowym) oraz ewaluacja posługują się podobnymi metodami badawczymi, a podstawową różnicą między oboma rodzajami analiz jest ta, że ocena (kontrola) finansowa zawsze poprzedza implementację programu, a ewaluacja implementacji towarzyszy bądź miejsce już po jej zakończeniu (co negowałoby niejako zasadność ewaluacji *ex ante*).<sup>67</sup> Wyraźny to dowód na to, że choć świadomość znaczenia systematycznego dokonywania ewaluacji realizowanych działań kreowana jest na standard w administracji Wielkiej Brytanii, to nieporozumienia wokół charakterystyki zadań i funkcji, które może pełnić nie znikają, a nawet, paradoksalnie, mnożą się. Techniczne podejście do delikatnej materii analiz o charakterze społecznym, zawarte w Zielonej Księdze powoduje, że tracą one stanowiącą o ich wartości elastyczność, a obecne w postrzeganiu ewaluacji nieporozumienia definicyjne i metodologiczne pozostają w dużej mierze niewyjaśnione.

---

<sup>65</sup> The Green Book. Appraisal and Evaluation in Central Government. Treasury Guidance, HM Treasury, 2003, s. 1; [http://www.hm-treasury.gov.uk/economic\\_data\\_and\\_tools/greenbook/data\\_greenbook\\_index.cfm](http://www.hm-treasury.gov.uk/economic_data_and_tools/greenbook/data_greenbook_index.cfm) (stan z dnia 01.07.2004 r.)

<sup>66</sup> op. cit., s. 13-14

<sup>67</sup> op. cit., s. 48

### 3.4. Wnioski dla Polski

1. Administracje krajów Unii Europejskiej, a także administracja instytucji Wspólnoty przypisują badaniom ewaluacyjnym coraz większe znaczenie. Stają się one powoli standardowym elementem procesu zarządzania. W przypadku polskiej administracji wiedza w zakresie technik ewaluacji niezbędna jest już teraz wszystkim jednostkom administracji zajmującym się obsługą unijnych funduszy strukturalnych. W sytuacji nasilania się dążeń do standaryzacji warunków życia w krajach UE (np. jakości usług publicznych, bezpieczeństwa itd.) znaczenie tej wiedzy będzie systematycznie wzrastało.
2. Ewaluacja programów administracji publicznej zaczyna zyskiwać w Unii Europejskiej stabilne umocowanie prawne. Wynika to zapewne z wieloletnich doświadczeń instytucji z obszaru UE związanych z funkcjonalnością badań ewaluacyjnych w sferze polityki wewnętrznej, europejskiej i z praktyki działania organizacji trzeciego sektora. Rozwiązania prawne legitymizujące badania ewaluacyjne w administracji mają zatem określony społeczny kontekst, który nie występuje na razie w przypadku Polski.
3. Z opisanych przykładów stosowania ewaluacji w krajach UE i w Komisji Europejskiej wynika, że rola badań ewaluacyjnych jest różnie pojmowana. Jej praktyczne formy rozciągają się na skali od kontroli finansowej do eksperymentu społecznego. Istnieje zatem potrzeba porządkowania obecnych już w krajowej administracji definicji roli i zadań tych badań. Błędem byłoby negowanie, jako nieprawidłowe, któregokolwiek z podejść już aplikowanych. Chodzi raczej o uświadamianie, jak różne zakresy wiedzy o badanym obiekcie ewaluacja może dostarczyć, i jaką zmianę społeczną spowodować.
4. Przykłady wskazują, że bardzo istotnym, jeśli nie niezbędnym, warunkiem efektywnej ewaluacji jest przyjazne otoczenie społeczne badania (związane np. z poparciem dla polityki rządu). Zaufanie do wyników ewaluacji, do wniosków jakie z niej wynikają, jest funkcją zaufania do grupy realizującej badanie i instytucji, które takie badanie zamówiło.
5. Poddanie ewaluacji projektów wzbudzających społeczne zainteresowanie sprzyja budowaniu korzystnego klimatu wokół nich. Poddanie programu ewaluacji bywa traktowane jako symbol uznania dla jego ważności, rodzi także przekonanie, że jego realizatorzy/koordynatorzy dbają o to, by zapewniał on produkty jak najwyższej jakości.

#### **4. Wybrane aspekty obecności badań ewaluacyjnych w polskiej administracji**

Zapisy o potrzebie prowadzenia badań ewaluacyjnych pojawiły się w polskiej administracji publicznej przede wszystkim poprzez odwzorowanie schematów programowania europejskich funduszy strukturalnych. Jak dotąd, zapisy takie funkcjonują w niewielu dokumentach rządowych. O ile się pojawiają, występują zazwyczaj w formie lakonicznej, są niedookreślone co do zakresu badania, jego celu, wykonawcy i sposobu późniejszego wykorzystania uzyskanych wyników. To, biorąc pod uwagę, że skrupulatne zaplanowanie ewaluacji, w tym uzyskanie dla niej szerokiego poparcia zwierzchników i innych zainteresowanych, jest niezbędnym warunkiem późniejszego wykorzystania jej wyników, oznacza potencjalne kłopoty z jej faktycznym przeprowadzeniem.

Jednocześnie, wraz z coraz szerszym zjawiskiem współfinansowania projektów, polityk, czy strategii rządowych przez fundusze Unii Europejskiej spodziewać się należy, że standard planowania ewaluacji wszystkich istotnych projektów realizowanych przez administrację publiczną będzie coraz częściej przenoszony na działania niezwiązane finansowo z polskim członkostwem we Wspólnocie. Administracja zyska nowe zadania, wymagające od niej mobilizacji, odpowiedniego przygotowania merytorycznego i elastyczności.

##### **4.1. Obecność ewaluacji w praktyce programowania działań w administracji centralnej w Polsce. Analiza narzędzi i technik ewaluacji w strategiach i programach resortowych opracowanych w latach 2000 – 2004**

W ciągu ostatnich kilku lat wzrosła liczba opracowywanych strategii i resortowych programów działania obejmujących różnorodne horyzonty czasowe (podczas gdy w roku 2001 opracowano 10 dokumentów tego rodzaju, dwa lata później już 43)<sup>68</sup>. Są one silnie zróżnicowane pod względem charakteru (kierunkowe, wykonawcze, dotyczące polityki wewnętrznej, programujące europejskie fundusze strukturalne itd.) i stopnia szczegółowości celów i zadań. Są zresztą niejednorodne już na poziomie konstrukcji – wiele zawiera niemal wyłącznie analizy bieżącego stanu realizacji danej polityki, a faktyczne cele działań w danej dziedzinie mają często charakter dezyderatów, których realizacja jest praktycznie niemierzalna (np. „budowa społeczeństwa obywatelskiego” czy „przywrócenie autorytetu

---

<sup>68</sup> por. Wieloletnie strategie i programy resortowe powstałe w Polsce w latach 1999-2003, RCSS, 2003

państwa jako konstytutywnego gwaranta bezpieczeństwa obywateli”)<sup>69</sup>. Właśnie ze względu na często jedynie kierunkowy i ogólny charakter zapisanych w strategiach i programach celów nie dziwi fakt, że tylko niecałe 30 z ponad 130 zanalizowanych dokumentów<sup>70</sup> zawiera rozdziały, podrozdziały lub akapity odnoszące się bezpośrednio do jakichkolwiek form ewaluacji realizacji określonych tam zadań (przy użyciu słów „ocena”, „kontrola”, „monitorowanie”, „monitoring” lub „ewaluacja”). Ponieważ brak szczegółowo zdefiniowanych celów wyklucza określenie wskaźników mierzących postępy w ich realizacji, ewaluacja postępów we wdrażaniu programów staje się niemożliwa nawet do zaprojektowania.

Co interesujące, choć bywa że nie są precyzyjnie nazwane, stosunkowo często pojawiają się w badanych strategiach wybrane elementy ewaluacji. Pojawiają się już na poziomie konstrukcji dokumentu<sup>71</sup> bądź, uelastyczniając kryteria analizy, jako formę ewaluacji można traktować te części strategii, które stanowią analizę wskaźników bieżącego stanu spraw w danej dziedzinie<sup>72</sup>.

W bardzo niewielu dokumentach pojawia się wątek konsultacji społecznych w odniesieniu do założeń samego dokumentu i sposobów jego wdrażania (monitoringu). Zazwyczaj jest to wzmianka wymieniająca z nazwy grupy interesu biorące udział w przygotowaniu programu lub informacja, że program został poddany konsultacjom międzyresortowym. Nie jest jednak opisany ani przebieg tych konsultacji, ani ich zakres i wynik. W kilku analizowanych dokumentach założono, że funkcję społecznych konsultacji mogą pełnić badania opinii publicznej<sup>73</sup>. W wyjątkowych przypadkach jako część realizacji programu założono zainicjowanie społecznej debaty nad problemem będącym przedmiotem dokumentu (działanie takie może skądinąd paradoksalnie obniżyć wizerunek samego programu – i ile debata taka,

---

<sup>69</sup> *Bezpieczna Polska. Program Poprawy Bezpieczeństwa Obywateli*, Ministerstwo Spraw Wewnętrznych i Administracji, sierpień 2002, s. 5

<sup>70</sup> do połowy czerwca 2004 r.

<sup>71</sup> np. pkt. 3.2. Analiza SWOT przed restrukturyzacją w: *Restrukturyzacja i rozwój hutnictwa żelaza i stali w Polsce do 2006 r.*, Ministerstwo Gospodarki, Pracy i Polityki Społecznej, 2003 r., pkt. 7. Analiza SWOT sektora elektronicznego w Polsce w: *Strategia dla przemysłu elektronicznego do roku 2010*, Ministerstwo Gospodarki, 2002 r.

<sup>72</sup> np. pkt. 3.1. Ocena realizacji głównych celów polityki energetycznej w: *Założenia Polityki Energetycznej Polski do 2020 roku*, Ministerstwo Gospodarki, 2000 r., pkt. 2. Stan obecny przemysłu drzewnego i pkt. 3. Główne bariery rozwoju i wzrostu konkurencyjności i perspektyw jego rozwoju w: *Strategia dla przemysłu drzewnego do 2006 roku*, Ministerstwo Gospodarki, Pracy i Polityki Społecznej, 2003 r.

<sup>73</sup> m.in. *Program uporządkowania i ograniczenia wydatków publicznych*, Ministerstwo Gospodarki, Pracy i Polityki Społecznej, 2003 r., *Założenia strategii działań informacyjnych i edukacyjnych w pierwszych latach członkostwa polski w Unii Europejskiej*, Urząd Komitetu Integracji Europejskiej, 2004 r.

mogąca mieć rozmaity przebieg, jest prowadzona w trakcie jego realizacji).<sup>74</sup> W obu jednak opisanych rodzajach zapisów brakuje informacji, w jaki sposób wyniki badań lub społecznej debaty mogłyby praktycznie wpływać na sposób realizacji programu. Mechanizmy takie wydają się nie być stworzone, co w praktyce oznacza, że wspomniane zapisy mogą mieć charakter jedynie symboliczny (świadczy o tym np. brak precyzji w planowaniu wspomnianych działań „konsultacyjnych” – „*W trakcie trwania debaty prowadzone będą badania opinii publicznej (np. w październiku i styczniu), celem sprawdzenia akceptowalności proponowanych oraz oceny skuteczności informowania opinii publicznej o prowadzonej debacie.*”<sup>75</sup> (tekst oryginalny).

W tych strategiach i programach, w których jakakolwiek forma szeroko pojętej ewaluacji ich realizacji jest uwzględniona *explicite*, rola tych elementów, mierzona szczegółowością odnośnych zapisów, jest również bardzo zróżnicowana. W zdecydowanej większości dokumentów problem potraktowany jest w sposób wysoce ogólny. Zakłada się np. powołanie bliżej nieokreślonego Zespołu mającego zajmować się opracowaniem harmonogramu realizacji zadań programu, prowadzeniem monitoringu zapisanych w programie działań oraz składaniem sprawozdań ze stanu ich wykonania (np. *Narodowa Strategia Rozwoju Dostępu Szerokopasmowego do Internetu „Siećpospolita” na lata 2004 – 2006*, Ministerstwo Infrastruktury, Ministerstwo Nauki i Informatyzacji, 2003 r.) lub nakłada na danego ministra działowego obowiązek przedkładania Radzie Ministrów corocznych raportów z realizacji programu, nie uściślając jednak, jakie metody badawcze powinny zostać wykorzystane do sporządzania takich opracowań, ani kto miałby je przygotowywać (m.in. *Program rozwoju energetyki wiatrowej w Polsce na lata 2002 – 2005*, Ministerstwo Środowiska, 2001 r.). Zazwyczaj jednak zakłada się jako pilne zadanie „wypracowanie skutecznych mechanizmów realizacji funkcji kontrolnych”, „stworzenie ram prawnych dla funkcjonowania systemu kontroli” lub „poprawę jakości dostępnych danych”. Ostateczne mechanizmy dokonywania ewaluacji pozostają więc nieznane, a ich adekwatność czy jakość trudno określić.

Pod względem szczegółowości zapisów dotyczących ewaluacji realizacji programów wyróżniają się przede wszystkim dokumenty będące częścią systemu programowania

---

<sup>74</sup> *Program uporządkowania i ograniczenia wydatków publicznych*, Ministerstwo Gospodarki, Pracy i Polityki Społecznej, 2003 r., s. 54, por. także *Program działań MENiS w latach 2002-2010*, Ministerstwo Edukacji Narodowej i Sportu, 2002 r.

<sup>75</sup> *Program uporządkowania i ograniczenia wydatków publicznych*, Ministerstwo Gospodarki, Pracy i Polityki Społecznej, 2003 r., s. 55

funduszy strukturalnych Unii Europejskiej: *Narodowy Plan Rozwoju*, 2003 r., *Podstawy Wsparcia Wspólnoty*, 2003 r., *EQUAL. Jednolity Dokument Programowy dla Polski*, Ministerstwo Gospodarki, Pracy i Polityki Społecznej, 2003 r., *Sektorowy Program Operacyjny. Rozwój Zasobów Ludzkich 2004-2006*, Ministerstwo Gospodarki, Pracy i Polityki Społecznej, 2004 r., *Sektorowy Program Operacyjny Rybołówstwo i przetwórstwo ryb 2004-2006*, Ministerstwo Rolnictwa i Rozwoju Wsi, 2004 r., *Sektorowy Program Operacyjny: Transport na lata 2004-2006*, Ministerstwo Infrastruktury, 2003 r., *Sektorowy Program Operacyjny „Restrukturyzacja i modernizacja sektora żywnościowego i rozwój obszarów wiejskich”*, Ministerstwo Rolnictwa i Rozwoju Wsi, 2002 r., *Sektorowy Program Operacyjny. Wzrost konkurencyjności przedsiębiorstw, lata 2004-2006*, Ministerstwo Gospodarki, Pracy i Polityki Społecznej, 2004 r., *Program Operacyjny Pomoc techniczna 2004-2006*, Ministerstwo Gospodarki, Pracy i Polityki Społecznej, 2003 r., *Zintegrowany Program Operacyjny Rozwoju Regionalnego 2004-2006*, Ministerstwo Gospodarki, Pracy i Polityki Społecznej, 2004 r., oraz *Strategia Wykorzystania Funduszu Spójności*, Ministerstwo Gospodarki, Pracy i Polityki Społecznej, 2003 r. Zazwyczaj wprost odwołują się do rozporządzenia Rady 1260/1999 i zachowują zapisane tam standardy. Wyraźnie określone są w nich cele i zadania programu, zarysowane są podstawowe wskaźniki opisujące stan realizacji tych celów i zadań na danym etapie programu. Regulują kwestie monitorowania postępów realizacji programów oraz sposoby pozyskiwania niezbędnych danych. Osobno określają sposób realizacji, zakres oraz wykonawców poszczególnych etapów badań ewaluacyjnych (*ex ante*, *mid term*, *ex post*). Co istotne, oddzielnie projektują schemat prowadzenia kontroli finansowej (audytu), określając tym samym ewaluację jednoznacznie jako badanie o charakterze społecznym, a nie jako część studium efektywności ekonomicznej.

Konstrukcję podobną do Narodowego Planu Rozwoju 2004-2006, prawdopodobnie ze względu na kontekst, w jakim powstawał, ma *Plan Rozwoju Obszarów Wiejskich na lata 2004-2006* (Ministerstwo Rolnictwa i Rozwoju Wsi, 2004 r.). Wyraźnie świadczy to o tym, że perspektywa zaangażowania funduszy UE w realizację danej strategii wymusza wzmocnienie roli badań ewaluacyjnych. Jako niezbędny pojawia się w takim programie przejrzysty system wskaźników i schemat organizacyjny działań kontrolujących. Ewaluacja jest tam traktowana jako integralna część realizacji programu, jako jego wsparcie.

Nacisk na potrzebę systematycznej realizacji badań ewaluacyjnych (choć słowo „ewaluacja” nie zostało użyte) położono również w *Krajowym Programie Działań na Rzecz Kobiet – II etap wdrożeniowy na lata 2003 – 2005* (Kancelaria Prezesa Rady Ministrów, 2003 r.) oraz *Krajowym Programie Przeciwdziałania Narkomanii na Lata 2002-2005* (Ministerstwo Zdrowia, 2002 r.) W obu przypadkach wyraźnie określone zostały zadania programu, zostały im przypisane organy odpowiedzialne za ich wykonanie, termin realizacji oraz wskaźniki postępów w ich realizacji.

Paradoksalnie, spośród analizowanych strategii i programów resortowych ostatnich lat dokumenty przywołane w dwóch poprzednich akapitach są wyjątkami. Treść pozostałych sugeruje, że w obowiązującej w polskiej administracji konwencji słowo „strategia” lub „program” używane jest dla określenia dokumentu niemal wyłącznie ramowego, kierunkowego, wyznaczającego raczej cele niż sposób ich osiągnięcia, dopasowanego niejako do zmiennej rzeczywistości politycznej. Ewaluacja jawi się w tym kontekście jako praktyka wyrafinowana, dyscyplinująca dokument pod względem formy i treści, zazwyczaj zresztą faktycznie narzucona na podstawie zewnętrznych przepisów prawnych (w przypadku *Narodowego Planu Rozwoju, Podstaw Wsparcia Wspólnoty czy Sektorowych Programów Operacyjnych*) lub wymuszona porównaniami z podobnymi dokumentami realizowanymi przez administracje centralne innych krajów (jak w przypadku wspomnianego *Krajowego Programu Działań na Rzecz Kobiet – II etap wdrożeniowy na lata 2003 – 2005*).

#### **4.2. Warunki obecności ewaluacji w programowaniu europejskich funduszy strukturalnych**

Spośród wszystkich polskich ministerstw najwięcej uwagi na swoich stronach internetowych poświęca ewaluacji Ministerstwo Gospodarki i Pracy. Posługuje się przy tym wąską jej definicją, jako działania wyłącznie *oceniającego*, a więc przede wszystkim kontrolnego („ocena/oszacowanie stopnia jakości (stopnia) realizacji programu, tzn. jego faktycznych efektów, w stosunku do wcześniejszych założeń, tzn. oczekiwanych efektów).(…)”). Przydaje się przy tym ewaluacji funkcję melioracyjną („ulepszanie skuteczności i efektywności programów”) oraz porządkującą („zwiększenie przejrzystości

działań”). Uznana zostaje także jej niezależność wobec jednostki zarządzającej badanym programem.<sup>76</sup>

Obowiązkowi tak pojętej ewaluacji podlegają w Polsce w chwili obecnej dokumenty programujące fundusze Unii Europejskiej – tj. *Narodowy Plan Rozwoju*, jego siedem programów operacyjnych (*Program Operacyjny Pomocy Technicznej*, *Zintegrowany Program Operacyjny Rozwoju Regionalnego*, *Sektorowy Program Operacyjny Rozwój Zasobów Ludzkich*, *Sektorowy Program Operacyjny Wzrost Konkurencyjności Przedsiębiorstw*, *Sektorowy Program Operacyjny Restrukturyzacja i Modernizacja Sektora Żywnościowego*, *Sektorowy Program Operacyjny Rybołówstwo i Przetwórstwo Ryb* oraz *Sektorowy Program Operacyjny Transport*), uzupełnienia tych programów oraz kontrakty wojewódzkie.

Za przeprowadzenie ewaluacji Narodowego Planu Rozwoju odpowiada Krajowa Jednostka Oceny, wykonująca swoje obowiązki we współpracy z instytucjami zarządzającymi, instytucjami pośredniczącymi oraz zarządami województw. Powstała ona jako Zespół ds. Oceny w ramach Departamentu Koordynacji Polityki Strukturalnej Ministerstwa Gospodarki i Pracy. Do jej zadań należy:

- rozpowszechnianie wiedzy na temat stosowanych w Unii Europejskiej standardów ewaluacji wykorzystania funduszy strukturalnych Wspólnoty,
- organizowanie i nadzór nad przebiegiem ewaluacji Narodowego Planu Rozwoju oraz
- koordynacja innych działań z zakresu ewaluacji realizowanych przez administrację publiczną.

Jednocześnie, docelowo, w każdej z instytucji zarządzających poszczególnymi programami operacyjnymi mają powstać odrębne zespoły do spraw ewaluacji.

„Wzmocnieniu zdolności w zakresie zarządzania, monitorowania, oceny i kontroli Podstaw Wsparcia Wspólnoty” poświęcony jest Priorytet 1 Uzupełnienia Programu Operacyjnego Pomoc Techniczna 2004-2006. Dokument zakłada, że „kompetencje osób zaangażowanych w proces realizacji Podstaw Wsparcia Wspólnoty (CSF) oraz wypracowane metody i procedury zarządzania, monitorowania, oceny i kontroli wdrażania CSF będą miały decydujące znaczenie dla sukcesu lub porażki pierwszego kontaktu (Polski – przyp. aut.) z funduszami strukturalnymi.”<sup>77</sup> Realizacja Priorytetu 1 zakłada gruntowne przeszkolenie w zakresie technik zarządzania, monitorowania, oceny i kontroli realizacji Podstaw Wsparcia Wspólnoty szerokiego grona osób związanych z administrowaniem funduszami UE.

<sup>76</sup> <http://www.fundusze-strukturalne.gov.pl/fundusze.php?dzial=1418> (stan z dnia 28.06.2004 r.)

<sup>77</sup> Program Operacyjny Pomoc Techniczna 2004-2006. Uzupełnienie Programu, Warszawa 2004 r.; s. 13

W szkoleniach, seminariach, treningach mają wziąć udział: pracownicy Instytucji Zarządzającej CSF, Instytucji Zarządzającej PO Pomoc Techniczna, jednostek monitorująco – kontrolnych Europejskiego Funduszu Rozwoju Regionalnego, Europejskiego Funduszu Społecznego, Europejskiego Funduszu Orientacji i Gwarancji Rolnej, Finansowego Instrumentu Wspierania Rybołówstwa, Krajowej Jednostki Ewaluacyjnej i jednostek audytu wewnętrznego, a także członkowie Komitetu Monitorującego CSF, Komitetu Monitorującego i Sterującego PO Pomoc Techniczna, a nawet pracownicy Urzędów Kontroli Skarbowej. Oznaczałoby to, że wiedza na temat funkcji, metod i znaczenia ewaluacji wśród szerokiej grupy pracowników administracji publicznej i szeregu innych osób aktywnie uczestniczących w procesie absorpcji europejskich funduszy strukturalnych w okresie do 2008 roku (zgodnie z zasadą n+2) wzrośnie bardzo znacząco. To spowoduje, że popularność, i praktyczne wykorzystanie ewaluacji przekroczy kontekst polityki strukturalnej Wspólnoty, i przeniknie do większości obszarów aktywności administracji publicznej.

*opracowała: Katarzyna Ekiert, Departament Spraw Społecznych i Ustrojowych*