

Dr Ryszard Szarfenberg
Instytut Polityki Społecznej
Uniwersytet Warszawski
www.ips.uw.edu.pl/rszarf
r.szarfenberg@uw.edu.pl

Ewaluacja strategii a polityka społeczna (wersja 1.0, 17.03.09)

Wprowadzenie

W klasycznej już koncepcji nauki dla polityki społecznej miała ona pomagać w „formułowaniu programów działania”, co z kolei wymagało „opracowania procedury diagnozowania, prognozowania, planowania, kontrolowania realizacji planów”, a to ostatnie w celu ich korygowania (Supińska, 1991, s. 10). W kolejnej wersji koncepcji tej nauki jej funkcją była racjonalizacja decyzji polityków, która miała polegać na naukowym opracowaniu „postgnozy” („jak i dlaczego doszło do tego, co jest teraz”), diagnozy, a następnie prognoz, „wewnętrznie spójnego, wykonalnego i efektywnego programu” („obmyślanie pożądanej przyszłości... i sposobów jej osiągnięcia”) oraz „ewaluacji skutków realizowanej polityki” i monitoringu reform (Supińska, 1998: 38). W ten sposób swojsko brzmiąca kontrola realizacji programów i planów przyjmuje postać ewaluacji i monitoringu.¹ W obu przypadkach przypisany został im podobny cel: korygowanie programów w trakcie realizacji. Ogólny sens tych procedur wyrażony został również tak: „ciągle ocenianie, co ta polityka naprawdę daje”, „krytyczne śledzenie skutków realizowanej polityki” (Supińska, 1998, s. 38).

Zanim idea kontrolowania programów lub ich ewaluacji i monitoringu pojawiła się w taki sposób na gruncie teorii polityki społecznej, była już dobrze zakorzeniona w polskiej refleksji socjotechnicznej inspirowanej ideą polityki prawa Leona Petrażyckiego. Już w 1957 r. Adam Podgórecki w ramach toku postępowania legislacyjnego² (Podgórecki, 1957, s. 74-75) przewidział procedury weryfikacji wstępnej na podstawie posiadanej wiedzy i eksperymentalnie³ (przed nadaniem mocy obowiązującej aktowi normatywnemu) oraz weryfikacji retrospektywnej (po nadaniu mu mocy obowiązującej, *ex retro*). Na weryfikację wtórną składały się dwa rodzaje badań: zmierzające do ustalenia „w jakim stopniu projektowane skutki zamierzone zostały faktycznie zrealizowane” oraz zmierzających do stwierdzenia „istnienia lub nieistnienia skutków niezamierzonych”, gdyż „mogą to być bowiem skutki niezamierzone pozytywne albo niezamierzone negatywne” (Podgórecki, 1957, s. 76). Wyniki weryfikacji retrospektywnej zestawione z „ocenami uznanymi przez władzę

¹ Wzajemne usytuowanie pojęć audytu (auditu) i monitoringu (a także badań naukowych) w stosunku do ewaluacji obszernie przedstawia Karol Olejniczak (2007). Interesujące byłoby włączenie do tego zestawu także kontroli, z którą z kolei zestawia się takie pojęcia jak: nadzór, rewizja, inspekcja, lustracja, wizytacja, sprawdzanie, weryfikacja, czuwanie i rozpatrywanie, analiza (zob. Kurowski, Ruśkowski i Suchacka-Krysiak, 2000, s. 14-28). Czy Najwyższa Izba Kontroli nie zajmuje się ewaluacją polityki publicznej? Z pewnością robi to jej odpowiednik z USA: Government Accountability Office. Na ten temat zob. także uwagi Tomasza Kierzkowskiego, który nie używał pojęcia audytu (2002, s. 15-17) oraz przewodniki Komisji Europejskiej (KE, 1997, s. 10; KE, 2004, s. 10-11).

² W późniejszych dziełach tego autora występuje ujęcie uogólniające dotyczące toku postępowania celowościowego (Podgórecki, 1970).

³ Podgórecki odróżniał wielką i małą legislację (regulująca konkretne sytuacje społeczne, np. zarządzenia, okólniki) i w odniesieniu do tej drugiej stwierdził: „Jest rzeczą jasną, że można akty prawne małej legislacji uzasadniać za pomocą eksperymentów sprawdzających związki przyczynowe” (Podgórecki, 1957, s. 73).

legislacyjną, tj. z celami ustawodawcy skorygowanymi przez katalog ocen”⁴ służą do wydania oceny ogólnej aktu prawnego. Jest to podejście w dużej mierze zgodne ze współczesnym rozumieniem ewaluacji (zob. Ozierański, 2006). Ponadto już wtedy Podgórecki powiązał te zagadnienia z weryfikacją teorii przyczynowo-skutkowej (określonej w uzasadnieniu aktu prawnego), co można uznać za bardzo zbliżone do współczesnej ewaluacji opartej na teorii oddziaływania programu na problem.

Tok postępowania legislacyjnego i celowościowego, który zaproponował Podgórecki miał przede wszystkim pomagać w formułowaniu aktów normatywnych i programów działania społecznego, czyli m.in. tego, co można uznać za strategię polityki społecznej. Poniżej zajmę się wyłącznie tym, co nazywał on weryfikacją wstępną i retrospektywną lub oceną skutków.

Definicje ewaluacji

W ujęciu językowym ewaluowanie, jak każda inna czynność, może być analizowane w kategoriach podmiotu (kto ewaluuje?) i przedmiotu (co jest ewaluowane?). Dodając do tego założenie o konieczności kryteriów ewaluowania (pod względem czego?) w skrócie mamy ogólną formułę: podmiot X ewaluuje przedmiot Y według kryteriów K. Podmiot ewaluowania to „ewaluator”, przedmiot ewaluowania nazywany jest czasem „ewaluandum”, ostatni element to kryteria ewaluacyjne. W formule X może być równe Y, wtedy dany podmiot ewaluuje sam siebie (auto-ewaluowanie). Ewaluowanie samo może być przedmiotem ewaluowania, sytuację taką nazywa się meta-ewaluowaniem (meta-ewaluacja). Szczególną przypadkiem mamy, gdy przedmiotem ewaluacji są bezpośrednio ludzie (np. personel realizujący jakieś zadanie), którzy mogą być podmiotami w kontekście wykonywanych przez siebie czynności, tzn. podmiot X ewaluuje podmioty innych działań i czynności. Ludzie podlegają różnym ocenom w szkole i w pracy. Ocenia się ich osiągnięcia szkolne, jakość ich pracy, a oceny te mogą mieć duży wpływ na ich życie i karierę.

Standardowy sposób czynienia ludzkich czynności zrozumiałymi polega na przypisaniu im celu, podobnie jest z ewaluowaniem – po co to robimy? Analizowane są one też poprzez pryzmat tego, co jest ich wytworem lub produktem, czyli: podmiot X ewaluuje przedmiot Y według kryteriów K w celu Z, w wyniku czego powstaje produkt P. Co może nim być? W ujęciu językowym możemy mówić tu właśnie o ewaluacji, tzn. produktem ewaluowania jest ewaluacja (pewnego rodzaju ocena), ale to drugie słowo używane jest też dla oznaczenia czynności, jak i ich bezpośrednich wytworów. Przede wszystkim są to nabyte w wyniku ewaluowania doświadczenie i wiedza ewaluatora utrwalone w różnych formach, np. w postaci raportu z przeprowadzonych czynności ewaluacyjnych. W treści raportu znajduje się ocena tego, co było przedmiotem ewaluacji.

Oczywiście ewaluowanie jednego ewaluatora może się różnić od ewaluowania innego ewaluatora nawet w przypadku tych samych przedmiotów i celów i produktów. W tym kontekście można mówić o różnych stylach, sposobach czy metodach dokonywania ewaluacji. Nasza formuła wzbogaca się wówczas następująco: podmiot X ewaluuje przedmiot Y według kryteriów K i z zastosowaniem metody M w celu Z, w wyniku czego powstaje produkt P. Czymś mniej złożonym niż metody są instrumenty, narzędzia czy techniki ewaluowania.

Jeżeli weźmiemy pod uwagę wszystkie wyróżnione elementy, to mogą się one składać na szersze podejścia ewaluacyjne czy modele ewaluacji.

⁴ Dalej podkreślał ponownie, że „ocena powinna być dokonana z punktu widzenia całego katalogu ocen, a nie z punktu widzenia ocen dotyczących danego aktu normatywnego”, przewidując możliwość kolizji ocen (tamże, s. 80).

Tabela 1. Logika ewaluowania

Ewaluowanie	Pytanie	Nazwa
Podmiot	Kto ewaluuje?	Ewaluator
Przedmiot	Co jest ewaluowane?	Ewaluandum, przedmiot ewaluacji
Kryteria	Według jakich kryteriów ewaluowane?	Kryteria ewaluacyjne
Wytwór	Co jest produktem ewaluowania?	Ewaluacja (w sensie wystawionej oceny, oszacowanej jakości i wartości)
Sposób	Jak ewaluowano?	Metoda ewaluacyjna

Źródło: opracowanie własne

Definicji ewaluacji (programów) czy badań ewaluacyjnych jest wiele, kilka z nich przedstawione zostanie w tabeli.

Tabela 2. Definicje ewaluacji

Definicja	Źródło
„zastosowanie metod naukowych do pomiaru realizacji i wyników programów dla celów związanych z podejmowaniem decyzji”	Rutman, za: Kosłowski, 2000, s. 861
„...proces określania jakości, wartości lub ważności czegoś, albo rezultat tego procesu [...] Nazwa autonomicznej dyscypliny... co oznacza badania i zastosowania procedur prowadzenia obiektywnej i systematycznej ewaluacji w pierwszym sensie”	Scriven, 1991, s. 139, 141
„Ewaluacja... to systematyczne zbieranie informacji o działaniach, charakterystykach i wynikach programów w celu dokonania ich oceny, poprawienia skuteczności i/lub wspomagania decyzji o ich przyszłości”	Patton, 1997, s. 23
„Ewaluacja składa się ze zbioru działań pomocnych w korygowaniu, wspieraniu i rozszerzaniu indywidualnego lub kolektywnego dokonywania potocznych ocen (<i>natural sensemaking</i> , ew. „czynienia zrozumiałymi”, „nadawania sensu”) polityk i programów realizowanych w celu zaspokajania ludzkich potrzeb... Ostatecznym celem ewaluacji jest poprawa społeczna (<i>social betterment</i>) [„ograniczenie problemów społecznych lub zapobieganie im, poprawa warunków społecznych, złagodzenie ludzkiego cierpienia”], poprzez pomaganie instytucjom demokratycznym w lepszej selekcji, nadzorze, usprawnianiu i ocenie programów i polityk społecznych”	Mark, Henry i Julnes, 2000, s. 3
„Ewaluacja jest procesem badawczym nastawionym na praktykę i polegającym na zbieraniu i dokonywaniu syntezy dowodów, który kulminuje się we wnioskach o sytuacji, wartości, jakości, znaczeniu programu, produktu, personelu, polityki, propozycji lub planu”	Fournier, 2005, s. 139
„Ewaluacja to systematyczny proces określania, uzyskiwania, referowania i stosowania informacji opisowych i ocennych o jakości, wartości, rzetelności, wykonalności, bezpieczeństwie, ważności i/lub sprawiedliwości (<i>equity</i>) danego przedmiotu”	Stufflebeam, Shinkfield, 2007, s. 18
„Ewaluacja to systematyczne badanie, prowadzone z użyciem zróżnicowanych metod, złożone ze zbierania danych, analizy, oceny oraz informowania o wynikach. Jego celem jest oszacowanie (w odniesieniu do jasno sformułowanych kryteriów) jakości i wartości procesu oraz efektów wdrażania interwencji publicznych”	Olejniczak, Kozak, Ledzion, s. 22
„Ewaluacja to systematyczne zbieranie, analiza i interpretacja danych w celu określenia wartości polityki społecznej lub programu pomocowego. Dane te mogą być wykorzystane przy podejmowaniu decyzji dotyczących owej polityki czy programu. Uzyskane informacje powinny być przydatne dla planowania, rozwoju i realizacji aktualnie prowadzonego programu”	J. Krzyszkowski, za: Ornacka, 2003, s. 125

Po pierwsze, możemy zastanawiać się, czy ewaluacja i badania ewaluacyjne są tym samym. Przyjmijmy najpierw, że ewaluacja to pojęcie o szerszym zakresie znaczeniowym niż badania ewaluacyjne, a ściślej: ewaluacja z zastosowaniem metod naukowych i/lub też w celach naukowych (o naukowości może decydować jeden lub kilka elementów każdej ewaluacji, np. naukowość podmiotu, przedmiotu, kryteriów, a w szczególności metod). Takie stanowisko w latach 60. przyjmował E. Suchman odróżniając ewaluację utożsamioną z ostatecznym wnioskiem ewaluacyjnym (*judgement*) oraz badania ewaluacyjne, czyli taki wniosek oparty na badaniach naukowych (za Stufflebeam... 2007, s. 276). To drugie pojęcie w amerykańskiej literaturze ewaluacyjnej nie jest już stosowane, mowa tam o ewaluacji lub ewaluacji programów (tamże, s. 277). Bardzo obszerne zestawienie podobieństw i różnic między badaniami ewaluacyjnymi i badaniami naukowymi przedstawia K. Ornacka (2003, s. 74-77), ale można je też interpretować, jako wskazanie różnic między ewaluacją a nauką. Oczywiście sama teoria i praktyka ewaluacji poddawana jest naukowym badaniom teoretycznym, empirycznym i stosowanym, co w dłuższym okresie doprowadziło do powstania całej społeczności badaczy i autonomicznej dyscypliny naukowej, którą również nazywa się ewaluacją (jest to drugie znaczenie tego słowa, które wyróżnił Scriven). Podobnie polityka społeczna rozumiana jest jako pewna praktyka, jak i dyscyplina nauki, która uznaje tę praktykę za główny przedmiot swoich badań i ewentualnie naukowo uzasadnionych oddziaływań.

Po drugie, w przytoczonych definicjach mowa o różnych przedmiotach ewaluacji: „interwencja publiczna”, „program, projekt lub polityka”, „program, produkt, personel, polityka, propozycja lub plan”, „program interwencji społecznej”, „polityka społeczna lub program pomocowy”. Jedyne w definicji Scrivena przedmiot ewaluacji pozostaje niedookreślony, czyli jest to definicja ewaluacji niezależnie od tego, co ma być jej przedmiotem. Autor ten jednak pisał o „wielkiej szóstce” w odniesieniu do rodzajów ewaluowanych przedmiotów: program, produkt, personel, osiągnięcia (*performance*), propozycja, polityka (1991, s. 141). Przy czym produkt i personel wydają się być nieodłącznymi składowymi każdego programu, tzn. trudno wyobrazić sobie program bez ludzi, którzy go realizują, a także bez bezpośrednich wytworów ich działań, z których korzystają klienci programu. Jeżeli polityka społeczna zdefiniowana zostanie w kategoriach interwencji publicznych w mniej lub bardziej ograniczonym obszarze potrzeb, problemów, kwestii czy spraw społeczno-socjalnych, to właśnie wyznaczy nam ogólne ramy przedmiotu lub przedmiotów, o których ewaluowanie nam tu chodzi. Szerzej o przedmiocie i obszarach ewaluacji w polityce społecznej piszę dalej.

Po trzecie, ewaluowanie ma prowadzić do pewnego bezpośredniego rezultatu, którym jest „ocena” lub „określenie jakości, wartości, rzetelności, wykonalności, bezpieczeństwa, ważności i/lub sprawiedliwości” pewnego przedmiotu (programu). W kontekście polityki społecznej chodzi m.in. o określenie oceny, jakości czy wartości programu lub projektu społeczno-socjalnego, czy ogólniej całej polityki społecznej, na której podstawie powstaje wiele takich programów i projektów. Ewaluowanie ma nam dać odpowiedź na pytania w rodzaju: jakiej jakości i wartości jest ten program, projekt czy ta polityka społeczna?

Najwięcej teoretycznych i praktycznych aspektów wydaje się wiązać z pojęciem jakości oraz ilościowym określaniem jej poziomu (stopnie, skale, klasy i stany jakości, a ogólniej wartościowanie jakości oraz nauka o jakości, czyli kwalitologia, np. Kolman, 1973). Odróżnienie jakości od wartości polega na tym, że ta pierwsza odpowiada na pytanie, jak dobrze dany program robi to, do czego go powołano, a druga – na ile program wysokiej jakości zaspokaja potrzeby, które uzasadniały jego uruchomienie. Przykładowo, program może być wysokiej jakości, ale jego produkt przestał być już potrzebny lub stał się mało istotny, bo wymyślono nowe programy inaczej i lepiej zaspokajające dany zestaw potrzeb.

W kontekście produkcji rynkowej bardzo rozpowszechniły się rozmaite metody, normy i standardy zarządzania, które mają na celu zapewnienie wysokiej jakości rozumianej zarówno jako „te cechy produktu, które odpowiadają potrzebom konsumenta i przez to zapewniają jego satysfakcję”, jak i jego „wolność od braków – wolność od błędów, które powodują konieczność powtórzenia pracy lub odpowiadają za jego porażki na rynku, czyli niezadowolenie konsumentów, ich skargi...” (Juran, 1998, s. 2.1, 2.2). Zagadnienie ceny i jakości towarów jest chyba znane każdemu konsumentowi, np. racjonalne zakupy polegają na uzyskaniu jak najlepszej jakości za daną cenę (kierowanie się porównaniami jakości do ceny). Dla producentów ma to też istotne znaczenie, gdyż towary z wieloma łatwymi do zidentyfikowania wadami (wadliwe, czyli niskiej jakości) słusznie nie cieszą się popularnością, a i mogą być zakazane, jeżeli wadliwość jest groźna dla konsumentów. Odniesienie tego rodzaju języka do sfery polityki społecznej, czyli mówienie o cenie i jakości usług społecznych i socjalnych, np. edukacji publicznej, publicznej ochrony zdrowia, usług z zakresu pomocy społecznej, usług pośrednictwa i doradztwa zawodowego, jest obecnie coraz bardziej popularne.

Cele i funkcje ewaluacji

W niektórych definicjach określono nie tylko bezpośredni wynik ewaluowania, ale także to, czemu taki wynik miałby służyć. Weźmy jakikolwiek produkt będący przedmiotem wymiany handlowej, np. telewizor. Określenie poziomu jego jakości może być wykorzystane do różnych celów przez różnych ludzi i organizacje czy instytucje. Organizacje konsumentów przedstawiają ranking jakości telewizorów podobnych rodzajów, ale wytworzonych przez różnych producentów. Załóżmy, że jesteśmy producentem tego, który otrzymał najgorsze oceny, do czego nam taka informacja może się przydać? Najprostsza odpowiedź: do tego, aby coś zmienić w procesie produkcji naszych telewizorów, żeby poprawić ich miejsce w rankingu produktów tego rodzaju. Analogicznie informacja o jakości programów, projektów czy polityk społeczno-socjalnych służyć może do tego, aby ich projektanci i wykonawcy poprawiali je w trakcie ich trwania, a w przyszłości projektowali i wykonywali programy o wyższej jakości, z których będą bardziej zadowoleni ci, którzy je zamawiają i finansują oraz ci, którzy z nich mają korzystać. Innymi słowy ewaluacja w polityce społecznej powinna przyczyniać się do takich jej zmian, które sprawiają, że jej poszczególne części, jak i ona, jako całość, zajmą wyższe miejsca w rankingach ewaluacyjnych, tzn. zostaną lepiej ocenione.

Pytanie z wyższego poziomu dotyczy tego, po co nam polityka społeczna lepszej jakości. Producenci zapewne chcą być lepsi od konkurentów na rynku, a robią to po to, aby mieć większe zyski. Jeżeli wygrywają poprzez najlepszy stosunek ceny do jakości to jest to z korzyścią dla konsumenta. W przypadku programów czy projektów społecznych projektodawcy też często muszą konkurować ze sobą, np. w ramach konkursów grantowych, a wówczas ostateczny cel jest podobny – dobro lub korzyść klienta. Trzeba jednak pamiętać, że klientem w takim przypadku nie są ci, którzy otrzymują produkt programu, ale ktoś inny (grantodawca). Jest to jedna z zasadniczych różnic między typowym programem rynkowym (produkt kupowany jest przez tego, kto go zużyje lub będzie użytkował) a typowym programem społecznym (produkt kupowany jest dla kogoś trzeciego, kto w domyśle nie jest w stanie za niego zapłacić ceny rynkowej).

W tabeli przytoczone zostanie kilka przykładów wycień celów ewaluowania w rozumieniu takim, jakie przedstawiono wyżej.

Tabela 3. Cele ewaluacji

Cele	Autor i źródło
------	----------------

<ul style="list-style-type: none"> • Oszacowanie jakości i wartości polityki lub programu społecznego • Usprawnienie programu i jego realizacji • Nadzór i stosowanie się do przepisów • Rozwój wiedzy 	Mark, Henry i Julnes, 2000, s....
<ul style="list-style-type: none"> • Wsparcie dla funkcji kontrolnej sprawowanej przez demokratyczne instytucje • Tworzenie wiedzy, która będzie lepszą podstawą projektowania polityki i programów • Pomaganie agencjom publicznym w zwiększaniu ich zdolności planistycznych, realizacyjnych i analitycznych • Informowanie opinii publicznej o działaniach władz publicznych poprzez upowszechnianie wyników ewaluacji 	Chelimsky, 2006, s. 39
<ul style="list-style-type: none"> • Rozliczenie z efektywnego wykorzystania zasobów i osiągnięcia oczekiwanych wyników (<i>accountability</i>) • Uczenie się • Zarządzanie programami działania i ich rozwój • Wypełnienie obowiązków etycznych 	Green, South, 2006, s. 5
<ul style="list-style-type: none"> • Wszechstronna i całościowa ocena programu • Uczenie się i poprawianie funkcjonowania programu • Rozliczenie z efektywnego wykorzystania zasobów i osiągnięcia oczekiwanych wyników (<i>accountability</i>) • Monitoring działania programu • Rozwój i zmiana sposobu działania - dostosowanie do złożonych, nowych i zmiennych warunków • Tworzenie wiedzy 	Patton, 2008, s. 139-141

W polskiej literaturze) zamiast o celach ewaluacji częściej pisze się o jej funkcjach (Olejniczak, Kozak, Ledzion, 2008, Ornacka, 2008). Czasem zrównuje się je ze sobą: „Ta szeroko pojęta rola może być realizowana przez szereg funkcji zwykle określanych mianem celów ewaluacji” (Olejniczak, 2008, s. 22). Różnica między celem a funkcją jest taka, że pierwsze pojęcie oznacza tylko zamierzone przez kogoś skutki ewaluowania, drugie natomiast – ogół wszystkich skutków użytecznych z jakiejś perspektywy (zwykle związane jest to z pytaniem: do czego i komu to może się przydać, jakie i czyje potrzeby to zaspokajają). Cele i funkcje mogą mieć wydźwięk pozytywny lub negatywny. W większości analiz mowa tylko o pozytywnych czy uprawnionych celach i funkcjach, ale socjologowie już dawno odróżnili eufunkcje (w pozytywnym sensie) i dysfunkcje (w negatywnym sensie). Niewiele dotąd pisano o dysfunkcjach ewaluacji, ale z pewnością nadmiar aktywności ewaluacyjnej w społeczeństwie może doprowadzić do skutków, które przez większość znaczną być odczuwane jako przynajmniej dokuczliwe.

Jeżeli głównym odbiorcą ewaluacji mają być agencje rządowe, to w zależności od tego, jak postrzega się ich rolę i działania (interwencje) w społeczeństwie, pozytywne nacechowanie naszego pojęcia może zanikać, np. ewaluacje dla służby bezpieczeństwa, w szczególności, jeżeli ewaluatorzy przyjmują za kryteria ewaluacji wyłącznie cele takich instytucji lub ich zarządów (Frieske, 2008). W jednoznacznie pozytywnym ujęciu polityki społecznej, to co służy jej poprawie może być też oceniane pozytywnie. Politykę tę postrzega się jednak nie tylko pozytywnie, może być krytykowana za zmniejszanie efektywności działania gospodarki, ograniczanie wolności i podważanie znaczenia naturalnych wspólnot, czemu przypisuje się mniej lub bardziej katastrofalne konsekwencje (patrz: Szarfenberg, 2008). W związku z tym ewaluowanie i ewaluacje można interpretować jako dobre o ile przyczyniają się do zmniejszania lub usuwania tych problemów, ale jeżeli są one uznawane za nieusuwalne (jako nieodłączne właściwości polityki społecznej jako takiej), to i rola ewaluacji staje się dwuznaczna.

Tabela 4. Funkcje ewaluacji

Funkcje	Autor i źródło
<ul style="list-style-type: none"> • Formatywna lub usprawniająca służąca poprawianiu, doskonaleniu • Konkluzywna lub podsumowująca związana z selekcją, certyfikacją, rozliczaniem • Psychologiczna lub społeczno-polityczna związana ze wzbudzaniem motywacji do działania i świadomości • Administracyjna związana ze sprawowaniem władzy 	Nevo 1983, s. 119
<ul style="list-style-type: none"> • Wspomaganie decyzji związanej z zarządzaniem programami społecznymi • Określenie jakości i wartości programu dla tych, którym ma on służyć 	Scriven, 1993, s. 9 (autor argumentował, że główną funkcją powinna być druga z nich)
<ul style="list-style-type: none"> • Opisuje mechanizm interwencji oraz oszacowuje... wpływ danej interwencji – poznawcza i sprawozdawcza (informowanie o ustaleniach ewaluacji) • Ocenia zasadność i sens podjętych działań publicznych - normatywna • Przedstawia rekomendacje i wnioski - techniczna 	Olejniczak, 2004, s. 4.
<ul style="list-style-type: none"> • Budowanie partnerstwa i współwłasności • Poprawa wdrażania i bieżącej jakości • Racjonalizowanie planowania • Funkcja poznawcza • Rozliczanie i wzmocnienie odpowiedzialności 	Olejniczak, Kozak, Ledzion, 2008, s. 30-31
<ul style="list-style-type: none"> • Poprawa planowania przez zagwarantowanie racjonalnego uzasadnienia dla danej interwencji • Poprawa wdrażania i kontrola jakości m.in. istniejącej struktury, procedur i procesu zarządzania interwencją publiczną • Wzmocnienie odpowiedzialności przez zwiększenie skuteczności i wydajności działań publicznych • Wspieranie procesów uczenia się m.in. przez budowanie zdolności ewaluacyjnych, weryfikowanie teorii leżących u podstaw danych polityk czy programów • Wzmacnianie partnerstwa i poczucia współwłasności przez realizowanie zasady <i>empowerment</i>, aktywizowanie i wzmacnianie powiązań, relacji, współpracy, interakcji między uczestnikami określonego przedsięwzięcia 	Ornacka, 2008, s. 28

Olejniczak (2008) umieszcza wszystkie funkcje między dwoma biegunami: orientacją operacyjną (na procesy zarządzania programem) oraz orientacją strategiczną (na ocenę efektów programów). Przypomina to nieco rozróżnienie Scrivena na ewaluację zorientowaną na zarządzanie (lub zarządców i ich potrzeby) i ewaluację zorientowaną na konsumenta i jego potrzeby, tzn. na tego, komu program ma służyć czy pomóc. Ta druga niekoniecznie musi się pokrywać z orientacją strategiczną wyróżnioną przez Olejniczaka.

Obok pozytywnie ocenianych funkcji ewaluacji, które uznać można też nazywać funkcjami jawnymi warto wspomnieć o jej funkcjach nieuprawnionych lub ukrytych, uważanych za niezgodne z tym do czego ewaluacja powinna służyć. Należą do nich (Weiss, 1998, s. 22):

- odwołanie ważnych decyzji;
- uzasadnianie decyzji, które zostały już podjęte;
- unikanie lub rozpraszanie odpowiedzialności za podjęte decyzje;

- poprawa wizerunku osoby lub instytucji zamawiającej badania (ewaluowanie tylko wybranych aspektów programu, żeby ukazać go w lepszym świetle, unikanie ewaluacji pewnych aspektów programów, aby ukryć jego porażkę).

Wyliczenia pozytywnych funkcji tabeli mają charakter normatywny, tzn. wyrażają przekonanie, czemu ewaluowanie i ewaluacja powinny służyć. Jeżeli już zgodzimy się na normatywną wizję ewaluacji, to z empirycznego punktu widzenia bardzo ważna jest odpowiedź na pytanie, czy rzeczywiście dobrze ona tej wizji służy. Pod koniec lat 80. zaczęto powątpiewać w to, czy nadzieje pokładane w ewaluacji są uzasadnione, m.in. ze względu na problemy z określeniem, komu ma ona służyć (np. wyzyskiwaczom czy wyzyskiwanym, biurokratom czy obywatelom), problemy z zastosowaniem jej wyników i problemy o charakterze metodologicznym (Rossi, 1987).

Przedmiot i kryteria ewaluacji w polityce społecznej

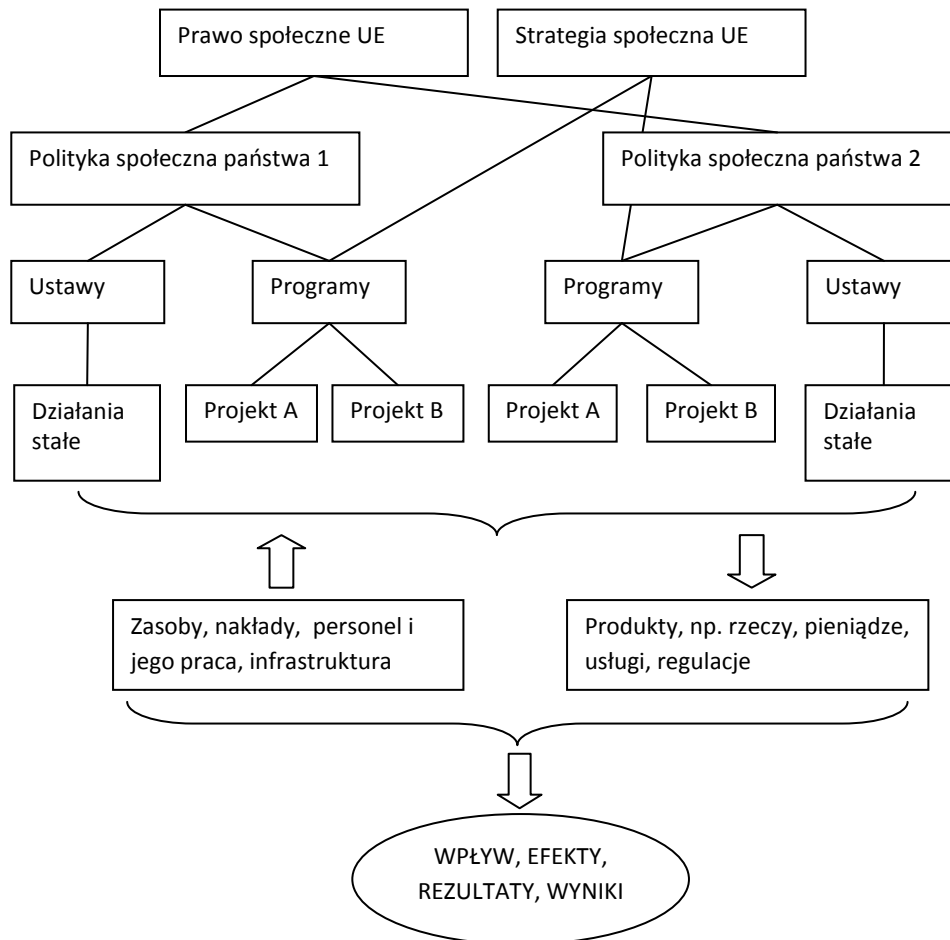
W amerykańskim dyskursie akademickim o ewaluacji najczęściej za jej przedmiot w interesującym nas kontekście uznaje się programy (społeczne), które z kolei rozkłada się na mniejsze części składowe, np. zasoby programu, działania programu, personel programu, klienci programu, produkty programu itp. (są to podstawowe elementy tzw. logiki programu). Z kolei w Polsce upowszechniła się nazwa projekt (w kontekście wyboru działań, które mają być finansowane m.in. z Europejskiego Funduszu Społecznego) lub projekt socjalny (w kontekście pracy socjalnej i pedagogiki społecznej), mają one podobne znaczenie do amerykańskiego terminu „program społeczny”.

Jeżeli przedmiotem ewaluacji jest pewien program lub projekt społeczny, to ewaluowanie jego poszczególnych elementów będzie składało się na ewaluację całości. Dlatego też ewaluowanie i jego wyniki dla niektórych tylko elementów programu lub projektu nie są tożsame z jego całościowo rozumianym ewaluowaniem. Przykładowo, ewaluowanie personelu programu nie jest ewaluowaniem programu, podobnie trudno uznać określenie jakości personelu (personel programu X był wysokiej jakości) za równoznaczne ustaleniu jakości programu (program X był wysokiej jakości), gdyż nie tylko z personelu się on składa i nie tylko personel decyduje o osiągnięciach i innych pożądanach cechach programów społeczno-socjalnych. Czy ewaluowanie tylko niektórych części składowych programu albo tylko w niektórych jego fazach, np. tylko po jego zakończeniu ma sens?

Politykę społeczną można potraktować analogicznie, jak traktuje się programy społeczne. Jej zasoby i nakłady na nią mierzone są m.in. udziałem wydatków społecznych w PKB, jej produkty mają postać regulacji, rzeczy, pieniędzy, usług itp., jej personel nazywany jest służbami społecznymi, przypisuje się jej też cele i instrumenty, a także podmioty i przedmiot, jak każdemu działaniu. Jak należy rozumieć relacje między polityką społeczną a programami i projektami społecznymi? Dotychczas, gdy mowa była o polityce społecznej w poszczególnych państwach uwaga obserwatorów skupiała się głównie na stałych elementach tej polityki, takich jak świadczenia społeczne w postaci pieniędzy, rzeczy i usług. Prawo do nich i ich kształt instytucjonalny zostały określone w konstytucjach, ustawach i aktach prawnych niższego rzędu. W porównaniu z typowym projektem społecznym nie określa się terminu ich zakończenia, są to działania o charakterze stałym i w różny sposób gwarantowane i wymagane przez państwo. Z ewaluacyjnego punktu widzenia sprawia to pewien kłopot, gdyż trudno mówić w takich przypadkach o ewaluacji po zakończeniu działania programu, czego nie przewidziano (rzadko ustawy mają określony czas do którego obowiązują, może poza aktami o charakterze wyjątkowym, ale już nie tak rzadko ogłasza się utratę ich mocy). Gdy jednak weźmiemy pod uwagę ogólną logikę programów społecznych, działania na podstawie prawa, np. ustawy o emeryturach i rentach z FUS, i działania na podstawie projektu

przewidującego pakiet świadczeń dla jakiejś zbiorowości ludzi w wieku emerytalnym nie różnią się od siebie i w zasadzie mogą być ewaluowane w podobny sposób. Uproszczony schemat wielopoziomowej polityki społecznej z uwzględnieniem twardych (dyrektywy UE, ustawy krajowe) i miękkich instrumentów (strategie, programy) oraz podziału na ogólniejsze strategie, programy (operacyjne) oraz projekty i działania stałe.

Schemat 1. Wielopoziomowa i wielonarodowa struktura polityki społecznej



Źródło: opracowanie własne

Wydaje się, że trudno dokonać ewaluacji tak złożonych systemów rozmaitych działań podejmowanych w różnych dziedzinach. Przykładów ewaluacji tego rodzaju dostarcza komparatystyka ewaluacyjna, w której jednostkami analizy jest polityka społeczna poszczególnych krajów (np. Goodin i in., 1999). Wymaga to najpierw wyróżnienia uniwersalnych kryteriów ewaluacyjnych polityki społecznej. Jednym z nich może być zakres, w którym polityka społeczna (zwykle ujmowana operacyjnie jako system podatkowo-transferowy) przyczynia się do zmniejszania ubóstwa, czyli jej skuteczność w tym względzie (por. Smeeding, 2005) lub skuteczność i efektywność (por. Herrmann i in., 2008; Szarfenberg, 2008).

Na podstawie przeglądu definicji polityki społecznej wyróżniłem w tym kontekście trzy często przywoływane wartości: wszechstronny dobrobyt, sprawiedliwość społeczna i integracja społeczna (Szarfenberg, 2008). Paul Spicker dodaje jeszcze kondycję gospodarki, która zawsze była istotna dla ekonomistów (2000), a Robert E. Goodin i in. wymieniają także stabilność społeczną i autonomię (1999). Z kolei Jolanta Supińska twierdziła, że z pojęciem

polityki społecznej zrosły się takie wartości, jak kolektywizm, równość, bezpieczeństwo, życie, do czego dodawała dwie kolejne – wolność i efektywność (1991).

Najbardziej złożoną oraz bardziej specyficzną propozycję modelu ewaluacyjnego dla polityki społecznej z uwzględnieniem jej poszczególnych elementów (np. cele, formy świadczeń, warunki uprawniające, sposób finansowania) zaproponował Donald E. Chambers (2005). Dla każdego z nich wyróżnione zostały kryteria specyficzne, a ponadto każdy oceniany był pod względem dwóch rodzajów kryteriów uniwersalnych: 1) konsekwencje elementu polityki społecznej dla klientów lub użytkowników programu pod względem adekwatności, sprawiedliwości i efektywności; 2) dopasowanie do wyników analizy problemu społecznego (definicja problemu, analiza przyczynowa, analiza wartości i ideologii oraz wygranych i przegranych).

Tabela 5. Kryteria ewaluacyjne specyficzne dla części składowych polityki społecznej

Podstawowy element polityki	Specyficzne kryteria ewaluacyjne
Cele i zadania	<ol style="list-style-type: none"> 1. Nie tylko dostarczanie świadczeń i usług, ale również ich wpływ na grupy docelowe; 2. jasność, mierzalność, manipulowalność; 3. włączenie standardów jakości i wyszczególnienie adresatów (<i>target</i>).
Formy świadczeń i usług	<ol style="list-style-type: none"> 1. Stygmatyzacja; 2. docieranie do adresatów; 3. koszty-skuteczność; 4. zastępowalność; 5. suwerenność konsumenta; 6. sprzeczności (<i>trade-offs</i>); 7. przymusowość/naruszanie prywatności; 8. złożoność i koszty administracyjne; 9. przystosowanie do różnych użytkowników; 10. ryzyko polityczne.
Warunki uprawniające	<ol style="list-style-type: none"> 1. nadmierne/niedostateczne wykorzystanie; 2. zbyt duże koszty; 3. stygmatyzacja/alienacja; 4. demobilizacja do pracy; 5. bodźce do prokreacji i rozwiązywania małżeństw i/lub dziedziczna zależność od świadczeń.
Administracja i dostarczanie usług	<ol style="list-style-type: none"> 1. czy istnieje wyartykułowana forma polityki/programu; 2. integracja/trwałość; 3. dostępność; 4. odpowiedzialność; 5. upodmiotowienie klienta/konsumenta; 6. uczestnictwo konsumenta w podejmowaniu decyzji; 7. radzenie sobie z rasową, płciową i etniczną różnorodnością.
Finansowanie	<ol style="list-style-type: none"> 1. trwałość finansowania; 2. stabilność mimo zmian ekonomicznych (inflacja/recesja) i demograficznych (wyże i niżej, zmiany w strukturze wiekowej).

Źródło: skrócona i uproszczona tabela za Chambers i Wedel, 2005, s. 56-57.

Kryterium oceny oznacza, że polityka społeczna mająca większe osiągnięcia w oznaczonym przez nie obszarze w porównaniu z inną polityką społeczną (np. w innym kraju lub w przeszłości, czy też dopiero planowaną), powinna być uznana za lepszą lub mającą większą wartość. Informacja tego rodzaju jest bardzo istotna dla celów praktycznych, gdyż może być podstawą do uzasadnienia reform, które miałyby poprawić wyniki gorszej z polityk pod względem danego kryterium. Poza tym w teorii decyzji wielokryterialna ocena

alternatywnych działań (a w naszym przypadku alternatywnych polityk społecznych) jest podstawą do podejmowania decyzji, tzn. należy wybrać to działanie czy politykę, które charakteryzuje się lepszym profilem w przestrzeni kryteriów ewaluacyjnych. Modele ewaluacyjne (zestawy kryteriów ewaluacyjnych) umożliwiają nam ocenę polityki społecznej lub jej elementów, co ma zasadnicze znaczenie dla dokonywania racjonalnych wyborów w tym zakresie.

Rodzaje i modele ewaluacji

Działania ewaluacyjne mogą być dzielone na szersze klasy według wielu różnych kryteriów podziału. Najczęściej stosowany podział opiera się na tym, po której fazie w procesie „od pomysłu do przemyślu” dokonuje się ewaluacji programu. Proces ten można podzielić na dwa główne etapy: 1) przygotowanie zakończone dokumentami takimi, jak m.in. plan działania; 2) realizację, podczas której za pomocą zasobów, personelu i działań wytwarzany i dostarczany jest produkt. Ewaluacja pierwszego etapu może dotyczyć przebiegu samego procesu przygotowań oraz jego wyników, czyli treści dokumentów takich jak diagnozy, strategie, plany, harmonogramy, budżety itd. Proces dochodzenia do nich mógł być niedemokratyczny, przeprowadzony bez dyskusji i konsultacji, pełen konfliktów i napięć, a ich treści mogą być niejasne, niespójne, nielogiczne, nierealistyczne, ogólnikowe i posiadać wiele innych luk i braków. Ewaluacja może je wykryć i dać szansę na poprawienie zarówno procesu przygotowawczego, jak i samych jego wytworów. Tego rodzaju dociekania nazywa się ewaluacją programu *ex ante*, czyli przed przystąpieniem do jego realizacji.

Po przystąpieniu do realizacji programu, a przed zakończeniem jego działania przeprowadza się ewaluację nazywaną *interim* lub *in media res*, czyli w trakcie realizacji (określenie „*mid-term*” dotyczy ewaluacji dokonanej w połowie okresu działania programu, a *on-going* – ewaluacji bieżącej lub okresowej). Służy ona ocenie programu jeszcze przed jego zakończeniem i może być przydatna do poprawienia jego działania lub wcześniejszego zakończenia w razie wystąpienia istotnych problemów realizacyjnych. Gdy działanie programu zostanie zakończone przeprowadza się ewaluację nazywaną *ex post*. Oczywiście im później zostanie ona przeprowadzona tym większe szansa na uwzględnienie długookresowych wyników programu. Stąd też istotną rolę odgrywają dociekania ewaluacyjne, które wykonuje się periodycznie przez kilka lat od zakończenia programu.

Ewaluacja całościowa (*ex ante*, *interim* i *ex post*) jest ważna z perspektywy troski o to, aby programy społeczne działały sprawnie i zaspokajały uzasadnione potrzeby.

Inny ważny podział ewaluacji ma u podstaw różne cele, jakie się jej przypisuje. Celem ewaluacji podsumowującej jest oszacowanie jakości i wartości programu społecznego. Jakość dotyczy samego programu i jego charakterystyk, np. jakość dostarczanych usług może być wysoka lub niska, natomiast wartość ma on z perspektywy zewnętrznej, np. wartość dla klienta, dla społeczności lokalnej. W trakcie tego rodzaju ewaluacji ocenia się m.in. wyniki programu (zamierzone i niezamierzone, pozytywne i negatywne) oraz inne jego charakterystyki.

Z kolei ewaluacja usprawniająca nastawiona jest na poprawianie programu i jego realizacji. Celem jest uzyskanie informacji, które pomogą zmodyfikować i ulepszyć działanie programu. Ważne są w tym kontekście aktualne informacje o produktach i o procesie realizacji służące do zidentyfikowania elementów programu, które nie spełniają standardów i formułowanie zaleceń naprawczych lub usprawniających. Wykorzystuje się w tego rodzaju ewaluacji m.in. porównywanie działań programu do najlepszych praktyk (podobne programy, które zostały uznane za sukces i wzór do naśladowania). Wyróżnia się też kilka innych celów ewaluacji, które mogą być podstawą do wyróżniania dodatkowych jej rodzajów.

W zależności od tego, kto dokonuje ewaluacji można podzielić ją na ewaluację wewnętrzną i zewnętrzną. Pierwszej dokonuje wyodrębniona jednostka, stanowiąca jednak część samego programu lub organizacji, która odpowiada za jego działanie. Druga przeprowadzana jest przez ewaluatorów, pracujących dla innej organizacji, która nie ma żadnych związków z administratorami programu, ani z jego personelem. Czasem wyróżnia się też autoewaluację, gdy ewaluator należy do personelu programu, czyli w jego planie przewidziano takie stanowisko. Autoewaluacja bez zatrudnienia ewaluatora może polegać na tym, że każdy członek personelu ma obowiązek przeprowadzać ewaluację swojej własnej pracy lub pracy własnej jednostki organizacyjnej.

Udział personelu programu, a także jego klientów w zarządzaniu lub w konsultowaniu planu i działań ewaluacyjnych służy do ich podziału na współzarządzane, uczestniczące i nieuczestniczące.

Źródłem podziału może być również to, czy ocenie podlega proces programu (ewaluacja procesu), czy też jego rezultaty (ewaluacja wyniku). Biorąc pod uwagę podstawowe elementy składowe każdego programu ewaluacji mogą podlegać zasoby, działania i produkt i stąd możliwość wyróżnienia kolejnych rodzajów. Wyróżnione i dodatkowe możliwe podziały przedstawia tabela.

Tabela 6. Rodzaje ewaluacji

Kryterium podziału	Podział ewaluacji na rodzaje
Fazy procesu programu	Ex ante / interim (w tym <i>mid-term</i> , <i>on-going</i>) / ex post
Cel	Podsumowująca / usprawniająca / kontrolna / poznawcza
Pozycja ewaluatora wobec programu	Wewnętrzna / zewnętrzna
Udział personelu, klientów	Współzarządzana / uczestnicząca / nieuczestnicząca (administracyjna)
Liczba kryteriów	Jednokryterialna / wielokryterialna
Rodzaj kryteriów	Skuteczności / efektywności / użyteczności / trafności itd.
Źródło kryteriów	Kryteria określone z góry (m.in. poprzez cele programu) / kryteria odkrywane w trakcie
Składniki programu	Planu / zasobów (w tym personelu, kosztów itp.) / działań / produktu (np. dobra, usługi)
Rodzaj wyników	Wyników w krótkim okresie / w średnim / w długim Wyniki bezpośrednie / wyniki pośrednie
Wykorzystywana metoda	Metodami eksperymentalnymi / metodami nieeksperymentalnymi / metodami mieszanymi Metodami ilościowymi / metodami jakościowymi / metodami mieszanymi
Kontekst	W polityce społecznej / w pracy socjalnej / w edukacji / w ochronie zdrowia / w kulturze / w ochronie środowiska / w polityce rynku pracy itd.

Źródło: opracowanie własne

W ciągu XX wieku tylko w USA ukształtowało się wiele szkół i podejść do ewaluacji opartych na różnych założeniach filozoficznych, metodologicznych i różnym rozumieniu misji społecznej. Najszersze ujęcie tej problematyki przedstawiają Daniel L. Stufflebeam i Anthony J. Shinkfield (2007). Uwzględnili oni również takie podejścia, które są lub mogą być mało wiarygodne – pseudo-ewaluacyjne (pięć podrodzajów, np. reporty z zakresu *public relations*, raporty o charakterze politycznym, programy upodmiotowienia pod przykryciem ewaluacji); lub też wybiórczo przedstawiają jakość i wartość programów społecznych – podejścia nastawione na konkretne pytania lub konkretną metodę (quasi-ewaluacyjne,

czternaście podrodzajów, np. ewaluacja skoncentrowana na celach programu, ewaluacja dająca informację o relacji rezultatów do nakładów, testowanie osiągnięć, analiza koszty-korzyści, badania z zastosowaniem metody eksperymentów kontrolowanych, ewaluacje oparte na teorii programu, ewaluacyjne studia przypadków). Właściwe podejścia ewaluacyjne (tzn. bez przedrostków pseudo- lub quasi-) autorzy podzielili najpierw na takie, które nastawione są na podnoszenie jakości programów w trakcie ich działania i pełną ocenę po ich zakończeniu, czyli spełniają funkcję usprawniającą i podsumowującą ewaluacji (trzy podrodzaje: zorientowane na podejmowanie decyzji i określenie jakości programu – model CIPP Stufflebeam’a, badania zorientowane na konsumenta, studia podejmowane dla celów akredytacji i certyfikacji). Kolejny rodzaj podejść ewaluacyjnych ma cele związane z reformą społeczną i rzecznictwem (trzy podrodzaje: zorientowane na klienta, ewaluacja konstruktywistyczna, podejście z zastosowaniem zasad demokracji deliberatywnej). Ostatni typ podejść ewaluacyjnych został nazwany eklektycznym, zaliczono tam tylko jedno stanowisko – rozwijany od lat przez Michaela Q. Pattona model ewaluacji nastawionej na zastosowanie.

Arbitralne stwierdzenie które z podejść ewaluacyjnych są pseudo lub quasi ewaluacjami może być zakwestionowane, jeżeli samo nie jest poparte rzetelnie przeprowadzoną ewaluacją. Jak przystało na środowisko ewaluacyjne podjęto więc próbę oszacowania wartości poszczególnych podejść według standardów ewaluacji programów, uzgodnionych przez środowisko ewaluatorów amerykańskich. Najlepsze wyniki łączne uzyskały podejścia reprezentowane przez model ewaluacji CIPP (skrót od słów angielskich *context, input, process, product*) – 92 punkty; model zorientowany na zastosowania Pattona – 86 punktów; trzecie miejsce z takim samym wynikiem zajęły ewaluacja zorientowana na konsumenta (rozwijana przez Michaela Scrivena) oraz podejście zorientowane na klienta (reprezentowane głównie przez Roberta E. Stake’a) – po 84 punkty. Wysokie wyniki miały też ewaluacja konstruktywistyczna (reprezentantami tego podejścia są Egon Guba i Yvonna Lincoln) oraz jedno z podejść skoncentrowanych na metodzie, a mianowicie ewaluacyjne studia przypadków (po 81 punktów). Zdecydowanie niższe wyniki uzyskały podejście oparte na celach programu (62 pkt.) i z zastosowaniem metody eksperymentów kontrolowanych (56 pkt.). W praktyce możliwości zastosowania najlepszej metody mogą być ograniczone z wielu względów.

Rozwój i perspektywy ewaluacji

Ewaluacja rozwija się już od co najmniej kilku dziesięcioleci w USA i innych krajach. Stała się już nie tylko działalnością naukową, ale profesjonalną. W USA jednym z głównych bodźców do tego rozwoju był wzrost zaangażowania rządu federalnego w rozwiązywanie problemów i kwestii społecznych, któremu towarzyszyło szybkie zwiększanie wydatków na programy społeczne. Zdecentralizowane państwo i kultura pragmatyzmu oraz wzrost wydatków społecznych, który zaczął być postrzegany jako problem sam w sobie, a także rosnące wątpliwości dotyczące skuteczności programów społecznych sprzyjały rozwojowi dociekań nastawionych na dostarczanie informacji o nich, ocen ich działania oraz rekomendacji, które miały je usprawnić lub pokazać alternatywne sposoby rozwiązywania problemów społecznych. Dodatkowym bodźcem było wprowadzenie prawnego obowiązku przeprowadzania „analiz wpływu regulacji” (*regulatory impact analysis, RIA*) dla inicjatyw legislacyjnych o kosztach powyżej 100 mln \$⁵.

Wśród wymogów narzuconych w ten sposób agencjom rządowym były m.in. takie: „działania regulacyjne nie powinny być podejmowane o ile potencjalne korzyści płynące dla

⁵ Rozporządzenie wykonawcze 12291 podpisane przez R. Reagana w 1981 r. W. Clinton podpisał w 1993 r. rozporządzenie o planowaniu i przeglądzie regulacji nr 12866 rozwijające zasady RIA.

społeczeństwa z regulacji nie przeważają nad ich potencjalnymi kosztami”, „cele regulacyjne powinny być wybierane tak, aby maksymalizować korzyści netto dla społeczeństwa”, „spośród alternatywnych podejść do realizacji każdego z celów regulacyjnych wybrane powinny być te, które nakładają najmniejsze koszty netto dla społeczeństwa”.⁶

W propagowanie szeroko rozumianego RIA zaangażowała się również Organizacja Współpracy Gospodarczej i Rozwoju (OECD). W 1995 Rada OECD przyjęła zalecenie w sprawie poprawy jakości regulacji rządowych, w którym m.in. zaprasza rządy państw członkowskich do składania co trzy lata raportów w tym zakresie.

Unia Europejska w połowie lat 90. zamówiła badania na temat praktyk ewaluacyjnych swoich własnych organów, a głównie Komisji Europejskiej i jej Dyrekcji Generalnych wyspecjalizowanych w prowadzeniu poszczególnych działów polityki wspólnotowej. Efektem tego był komunikat skierowany do Komisji, który miał poprawić sytuację w tej dziedzinie. Rozpoczyna go następujące zdanie: „rosnące rozmiary budżetu i troska podatkowników UE o to, aby za swój wkład do niego otrzymali odpowiednie usługi (*value*), a także rozwój w Państwach Członkowskich i gdzie indziej w kierunku zarządzania sektorem publicznym nastawionego na wyniki, stworzyły potrzebę przeglądu i unowocześnienia podejścia Komisji do zarządzania budżetem” (KE, 1996).

Jedno z zaleceń sformułowane w tym kontekście zobowiązywało wszystkie dyrekcje generalne do wyodrębnienia własnych funkcji ewaluacyjnych. Jak widać, głównym impulsem do rozwijania ewaluacji na poziomie UE były względy finansowe. Wspomniany komunikat powstał w kontekście inicjatywy „zdrowe i efektywne zarządzanie finansowe” zwanej w skrócie SEM 2000. Jej wyniki zostały uwzględnione w głównym akcie regulującym kwestie ogólnego budżetu UE z 2002 i tam też mowa o ewaluacji w art. 27 pkt. 4: „W celu poprawy procesu podejmowania decyzji instytucje powinny przeprowadzać ewaluacje *ex ante* i *ex post* zgodnie ze wskazówkami Komisji. Powinny być one stosowane do wszystkich programów i działań, które wiążą się ze znacznym wydatkowaniem środków, a wyniki ewaluacji powinny być rozpowszechniane...” (RUE, 2002).

Ewaluacja programów i projektów społecznych w Polsce zaczęła się dynamicznie rozwijać w związku z włączeniem Polski do UE.⁷ Społeczny wymiar integracji był obecny w największym stopniu w Sektorowym Programie Operacyjnym Rozwój Zasobów Ludzkich 2004-2006 oraz w Programie Inicjatywy Wspólnotowej EQUAL, które finansowane były z Europejskiego Funduszu Społecznego (EFS). W dokumencie będącym podstawą działania pierwszego z tych programów zapisano m.in.: „wzrost potencjału ewaluacyjnego polskiej administracji jest jednym z głównych celów w okresie programowania 2004-2006” (MGiP, 2004a). W instrukcji dla projektodawców, którzy chcieli korzystać ze środków EFS znalazł się natomiast taki punkt: „należy opisać, czy i jak będą badane rezultaty projektu, w szczególności efekty długoterminowe; w jaki sposób i kiedy zostanie przeprowadzona ewaluacja projektu po zakończeniu wsparcia, jeśli tak, to należy opisać sposób oraz termin przeprowadzenia oceny” (MGiP, 2004b, s. 76).

W PIW EQUAL zadeklarowano, że Polska zapewni, aby „wszystkie Partnerstwa na rzecz Rozwoju [organizacje realizujące projekty] opracowały metodologię i zastosowały odpowiednie mechanizmy dla oceny bieżącej swoich działań i osiągnięć, łącznie z opisem możliwych do weryfikacji wskaźników, które ilustrują, w jaki sposób cele i rezultaty będą mierzone i oceniane, a także odpowiednią metodologię i mechanizmy monitoringu i oceny wspólnych działań w ramach współpracy ponadnarodowej” (MGPiPS, 2004, s. 181). W

⁶ Rozporządzenie 12291.

⁷ Generalnie już od wczesnych lat 90. wraz z zaangażowaniem w Polsce amerykańskich funduszy pomocowych, stawiane przez grantodawców wymogi ewaluacyjne wytwarzały popyt na wiedzę i umiejętności w tym zakresie. Pierwszy zbiór tłumaczeń amerykańskich tekstów o ewaluacji został wydany w Polsce pod tytułem „Ewaluacja w edukacji” (Korporowicz, 1997). Rozwój ewaluacji amerykańskiej też pierwotnie był związany z edukacją.

innym dokumencie interpretuje się ten zapis jako obowiązek prowadzenia przez partnerstwa autoewaluacji. W PIW EQUAL znajdujemy jeszcze jedno istotne zdanie o wykorzystaniu okresu 2004-2006 na „stworzenie i rozwinięcie systemów ewaluacji programów oraz możliwość zastosowania ich w kolejnym okresie programowania” (tamże). W związku z tym zarówno na poziomie administracji centralnej, jak i w organizacjach z różnych sektorów powstał popyt na usługi ewaluacyjne. Szybko zaczęły powstawać podyplomowe studia ewaluacyjne⁸, kursy i szkolenia w zakresie ewaluacji, coraz więcej instytucji i firm zaczyna wchodzić na stworzony w ten sposób rynek, co dobrze wróży rozwojowi całej branży oraz profesjonalizacji tej działalności.

Przejdźmy teraz do początków RIA w Polsce. Obowiązują już przepisy, które wprowadzają obowiązek „oceny skutków regulacji” (OSR). Pojawiły się one w uchwale Rady Ministrów o regulaminie jej pracy z 1997 r. w wyniku poprawki z 2001 r. W obecnie obowiązującej wersji tej uchwały z 2002 r. mowa o tym, że każdy projekt aktu normatywnego powinien mieć w uzasadnieniu odrębną część z OSR, czyli oceną „przewidywanych skutków (kosztów i korzyści) społeczno-gospodarczych regulacji”. Ma ona zawierać „wskazanie podmiotów, na które oddziałuje akt normatywny”, „wyniki przeprowadzonych konsultacji” jeżeli są one wymagane, a także „przedstawienie wyników analizy wpływu aktu normatywnego” na: sektor finansów publicznych, w tym budżet państwa i budżety jednostek samorządu terytorialnego; rynek pracy; konkurencyjność gospodarki i przedsiębiorczość, w tym na funkcjonowanie przedsiębiorstw; sytuację i rozwój regionalny (RM, 2002).

Ministerstwo Gospodarki opublikowało w 2006 r. dokument pt. „Program reformy regulacji”, w tym samym roku powstało opracowanie pt. „Wytyczne do oceny skutków regulacji”. Znalazło się tam następujące zdanie: „Ocena skutków regulacji (OSR) stanowi instrument, który pozwala na określenie konsekwencji wprowadzania regulacji. Należy przy tym podkreślić, że OSR przygotowany jest w przypadku podjęcia decyzji o konieczności podjęcia działań interwencyjnych przez państwo i jest dokonywana przed opracowaniem projektu aktu prawnego” (MG, 2006a, s. 9). W takim znaczeniu OSR jest bardzo bliskie temu, co nazywa się ewaluacją *ex ante*. Miała to być odpowiedź na problemy z dotychczasowym systemem, który „Nie funkcjonował efektywnie ze względu na mało użyteczne procedury, które powodowały, że system OSR nie miał istotnego wpływu na proces podejmowania decyzji” (MG, 2006b).

Inicjatywa ta nawiązuje do komunikatu Komisji Europejskiej z 2005 r. „Lepsza regulacja dla wzrostu i miejsc pracy w UE”. Oznacza to, że OSR czyli RIA w polskim wydaniu ma również dobre perspektywy na rozwój i profesjonalizację (zob. Springer, Rogowski, 2007). Dotyczyć to będzie również regulacji w zakresie polityki społecznej.

Wnioski

Opracowywanie strategii, programów, planów rozwoju polityki społecznej lub rozwoju społecznego może być interpretowane jako przejaw ogólniejszej tendencji do racjonalizacji działania administracji publicznej w sferze społeczno-socjalnej. Poza tym potrzeba wprowadzenia miękkich sposobów zarządzania publicznego związana jest z procesem międzynarodowej regionalizacji gospodarczej (UE) i decentralizacji polityki społecznej w ramach państw oraz zwiększenia poziomu jej wielosektorowości. W wyniku tego środki publiczne wydawane są w coraz większym stopniu na realizowanie społecznych zadań publicznych według zasad konkursów o granty czy konkurencyjnych przetargów. Strategie i programy wyznaczają cele, priorytety i kierunki działań, które są wskazówką dla projektodawców i stanowią podstawę do ewaluacji *ex ante* projektów wybieranych na jej

⁸ W 2000 r. powstało Polskie Towarzystwo Ewaluacyjne. W 2008 r. środowisko to przedstawiło do dyskusji „Standardy ewaluacji”, dostępne na stronach PTE: www.pte.org.pl.

podstawie. Poziom kontroli państwa nad podmiotami spoza sektora publicznego realizującymi zadania publiczne za środki publiczne jest ograniczony, stąd wymogi ewaluacyjne w zakresie ewaluacji *ex post*, które są jednym z instrumentów pośredniej kontroli sensowności wydatkowania pieniędzy publicznych. To samo dotyczy prywatnych instytucji grantodawczych, których fundatorzy także nie chcą, aby ich pieniądze były marnowane na marne programy i projekty społeczne.

Literatura

- Chambers, D. E., Wedel, K. R. (2005). *Social Policy and Social Programs: A Method for the Practical Public Policy Analyst*. Allyn & Bacon.
- Chelimsky, E. (2006). The Purposes of Evaluation in a Democratic Society. W I. Shaw, J. C. Greene, M. M. Mark (red.) *The SAGE Handbook of Evaluation*. Thousand Oaks: SAGE Publications.
- Fournier, D. M. (2004). Evaluation. W S. Mathison (red.) *Encyclopedia of Evaluation*. Thousand Oaks: SAGE Publications.
- Frieske, K. W. (2008). *Ewaluacje w demokratycznym społeczeństwie*. Referat wygłoszony na IV Konferencji Ewaluacyjnej "Ewaluacja - kluczowy instrument poprawy jakości działań administracji publicznej", 17 październik.
- Goodin, R. E., Headey, B., Muffels, R., Dirven, H.-J. (1999). *The Real Worlds of Welfare Capitalism*. Cambridge: Cambridge University Press.
- Green, J., South, J. (2006). *Evaluation*. Open University Press.
- Herrmann, P. i in. (2008). *Efficiency and Effectiveness of Social Spending*. IZA Working Paper nr 3482, maj.
- Juran, J. M., Godfrey, A. B. (1998). *Juran's Quality Handbook*. McGraw-Hill Professional.
- KE (1996). *Concrete steps towards best practice across the Commission*. Komisja Europejska, SEC 96/659, maj.
- KE (1997). *Evaluating EU Expenditure Programmes: A Guide*. XIX/02 - Budgetary overview and evaluation Directorate-General XIX – Budgets, European Commission, styczeń.
- KE (2004). *Evaluating EU Activities: A Practical Guide for the Commission Services*. DG Budget – Evaluation unit, European Commission, lipiec.
- Kierzkowski, T. (2002). *Ocena (ewaluacja) programów i projektów o charakterze społeczno-gospodarczym w kontekście przystąpienia Polski do Unii Europejskiej*. Warszawa: Polska Agencja Rozwoju Przedsiębiorczości.
- Korporowicz L. (red.) (1997), *Ewaluacja w edukacji*, Warszawa, Oficyna
- Kolman, R. (1973). *Ilościowe określanie jakości*. Warszawa: Państwowe Wydawnictwo Ekonomiczne.
- Korporowicz, L. (1997). *Ewaluacja w edukacji*. Warszawa: Oficyna Naukowa.
- Kosloski, K. (2000). Evaluation research. W E. F. Borgatta, R. J. V. Montgomery (red.) *Encyclopedia of Sociology*. Macmillan Reference.
- Kurowski, L., Ruśkowski, E., Sochacka-Krysiak, H. (2000). *Kontrola finansowa w sektorze publicznym*. Warszawa: Polskie Wydawnictwo Ekonomiczne.
- Mark, M. M., Henry, G. T., Julnes, G. (2000). *Evaluation: An Integrated Framework for Understanding, Guiding, and Improving Policies and Programs*. San Francisco: Jossey-Bass.
- MG (2006a). *Wytyczne do oceny skutków regulacji*. Ministerstwo Gospodarki, październik
- MG (2006b). *Przedsiębiorca ma prawo - Program Reformy Regulacji*. Informacja prasowa z 28 sierpnia, Ministerstwo Gospodarki.

- MGiP (2004a). Sektorowy Program Operacyjny Rozwój Zasobów Ludzkich 2004-2006. Załącznik do Rozporządzenia Ministra Gospodarki i Pracy z dnia 1 lipca 2004 r. w sprawie przyjęcia Sektorowego Programu Operacyjnego Rozwój Zasobów Ludzkich 2004-2006 r. (Dz. U. 166 poz. 1743),.
- MGiP (2004b). *Europejski Fundusz Społeczny w Polsce - Poradnik dla projektodawców*, Ministerstwo Gospodarki, Pracy i Polityki Społecznej, luty.
- MGPiPS (2004). *Program Inicjatywy Wspólnotowej EQUAL dla Polski*. Ministerstwo Gospodarki, Pracy i Polityki Społecznej, kwiecień.
- Nevo, D. (1983). The Conceptualization of Educational Evaluation: An Analytical Review of the Literature. *Review of Educational Research*, tom 53, nr 1.
- Olejniczak, K. (2008). *Mechanizmy wykorzystania ewaluacji*. Warszawa: Wydawnictwo Naukowe Scholar.
- Olejniczak, K., Kozak, M., Ledzion, B. (red.) (2008). *Teoria i praktyka ewaluacji interwencji publicznych. Podręcznik akademicki*. Warszawa: Wydawnictwa Akademickie i Profesjonalne, Akademia im. Koźmińskiego.
- Olejniczak K. (2007). Ewaluacja jako narzędzie zarządzania w sektorze publicznym. W G. Gorzelak, A. Tucholska (red.) *Rozwój, region, przestrzeń*. Warszawa: Ministerstwo Rozwoju Regionalnego.
- Olejniczak, K. (2004). *Proces przeprowadzenia ewaluacji ex-ante i on-going w ramach Narodowego Planu Rozwoju z uwzględnieniem etapów prac: planowanie, projektowanie i realizacja ewaluacji - wkład do procedur operacyjnych ewaluacji w Polsce*. Ekspertyza dla Ministerstwa Gospodarki i Pracy.
- Ornacka, K. (2003). *Ewaluacja między naukami społecznymi i pracą socjalną*. Kraków: Wydawnictwo UJ.
- Ornacka, K. (2008). Rola ewaluacji w pomocy społecznej. W J. Staręga-Piasek, A. Hryniewicka (red.) *O potrzebie ewaluacji w pomocy społecznej*. Warszawa: Instytut Rozwoju Służb Społecznych.
- Ozierański, P. (2006). O związkach ewaluacji z socjotechniką. *Polityka Społeczna* nr 1 (382).
- Patton, M. Q. (1997, 2008). *Utilization-Focused Evaluation*. Thousand Oaks: SAGE Publications.
- Podgórecki, A. (1957). *Założenia polityki prawa. Metodologia pracy legislacyjnej i kodyfikacyjnej*. Warszawa: Wydawnictwo Prawnicze.
- Podgórecki, A. (1970). Logika praktycznego działania. W A. Podgórecki (red.) *Socjotechnika – jak oddziaływać skutecznie*. Warszawa: Książka i Wiedza.
- RM (2002). Uchwała Rady Ministrów nr 49 z dnia 19 marca 2002 r. Regulamin pracy Rady Ministrów (M. P. nr 13, poz. 221).
- Rossi, P. H. (1987). Iron Law of Evaluation and Other Metallic Rules. W J. Miller, M. Lewis (red.) *Research in Social Problems and Public Policy*, tom IV, Greenwich, CT: JAI Press.
- RUE (2002). Rozporządzenie Rady (WE, Euroatom) nr 1605/2002 z 25 czerwca 2002 w sprawie Rozporządzenia Finansowego mającego zastosowanie do ogólnego budżetu Wspólnot Europejskich.
- Scriven, M. (1991). *Evaluation Thesaurus*. Thousand Oaks: SAGE Publications.
- Scriven, M. (1993). *Hard-won lessons in program evaluation*. San Francisco: Jossey-Bass.
- Smeeding, T. (2005). *Poor People in Rich Nations: the United States in Comparative Perspective*. Luxembourg Income Study Working Paper nr 419.
- Spicker, P. (2000). *Welfare State: A General Theory*. Thousand Oaks: SAGE Publications.
- Stufflebeam, D. L., Shinkfield, A. J. (2007). *Evaluation Theory, Models, and Applications*. San Francisco: Jossey-Bass.

- Supińska, J. (1991). *Dylematy polityki społecznej*. Warszawa: Instytut Polityki Społecznej UW.
- Szarfenberg, R. (2008). *Krytyka i afirmacja polityki społecznej*. Warszawa: Wydawnictwo Instytut Filozofii i Socjologii PAN.
- Szpringer, W., Rogowski, W. (red.) (2007). *Ocena skutków regulacji – poradnik OSR, doświadczenia, perspektywy*. Warszawa: C.H. Beck.
- Weiss, C. H. (1998). *Evaluation*. Upper Saddle River: Prentice Hall.