

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

Wprowadzenie do ewaluacji programów i projektów społecznych
(wersja 1.0, 20.10.09)

Spis treści

Zagadnienia terminologiczne	3
Wstęp	4
Ewaluowanie jako rodzaj oceniania.....	4
Podstawowe elementy ewaluowania.....	5
Ewaluacja działania	7
Podstawowe sposoby ewaluacji działania służb społecznych	9
Cele i funkcje ewaluowania.....	11
Ewaluowanie a badania naukowe, kontrolowanie i monitorowanie działania	13
Zakończenie.....	14
Metody oceniania	16
Wstęp	17
Metody badań naukowych a metody oceniania	17
Zagadnienia metod oceny.....	18
Wybór kryteriów i ustalenie ich ważności	19
Wyznaczanie stopnia wartości i jakości programu	21
Agregowanie ocen częściowych – metodologia syntezy ewaluacyjnej	24
Zakończenie.....	28
Ewaluacja jako proces.....	29
Wstęp	30
Uzgodnienia kontraktowe.....	32
Ewaluacja kontekstu	34
Ewaluacja wkładu	35
Ewaluacja procesu	37
Ewaluacja wpływu.....	40
Ewaluacja skuteczności.....	41
Ewaluacja trwałości	42
Ewaluacja przenośności (opcjonalna)	44
Metaewaluacja	45
Raport końcowy.....	46
Zakończenie.....	48
Przykłady ewaluacji projektów	50
Wstęp	51
Ewaluacja Wolskiego Programu Reintegracji Społecznej „Druga Szansa”	51
Ewaluacja programu KE AKA HO’ONA	54
Ewaluacja projektów szkoleniowych z zakresu ICT dla osób bezrobotnych	58
Ewaluacja projektów wsparcia dla instytucji rynku pracy i pomocy społecznej.....	61
Zakończenie.....	64

Zagadnienia terminologiczne

Wstęp

W tej części książki zostanie pokazana panorama znaczeniowa pojęcia ewaluacji ze szczególnym uwzględnieniem kontekstu działania służb społecznych. Mimo, że pojęcie to pojawiło się w Polsce niedawno (upowszechnia się od lat 90. XX wieku) ma już dość złożony zakres znaczeniowy. Poza tym nie jest ono jedynym, którego używa się w tym sensie - początkowo i nadal jeszcze angielskie *evaluation* tłumaczone jest na język polski przede wszystkim jako ocena (dotyczy to także tłumaczeń oficjalnych dokumentów UE).¹ Mamy też słowa mniej lub bardziej bliskoznaczne lub blisko związane, np. weryfikacja, kontrola, audyt, monitoring. Żeby nie zagubić się w tym gąszczu słów i znaczeń potrzebny jest nam przewodnik, którego ma dostarczyć poniższy rozdział.

Po wprowadzeniu o bardziej słownikowym charakterze przedstawione zostaną zagadnienia dotyczące relacji między ewaluacją a ocenianiem, specyfiki ewaluowania ludzkich działań, jego celów i funkcji oraz związku między ewaluowaniem a kontrolą, audytem i monitoringiem.

Ewaluowanie jako rodzaj oceniania

Słowo „ewaluacja” jest zapożyczeniem z angielskiego lub francuskiego (ang. *evaluation*, franc. *évaluation*, niem. *bewertung*). Te słowa z kolei pochodzą od łacińskiego *valeo -ere* (mieć się dobrze, być coś wartym). Przedrostek e- w języku angielskim lub francuskim pochodzi najprawdopodobniej od łacińskiego *ex-*, co w znaczyło „od”, „z”. Można to interpretować tak, że ewaluacja jest wynikiem czynności (ewaluować, ewaluowanie), czymś co uzyskaliśmy dzięki niej.

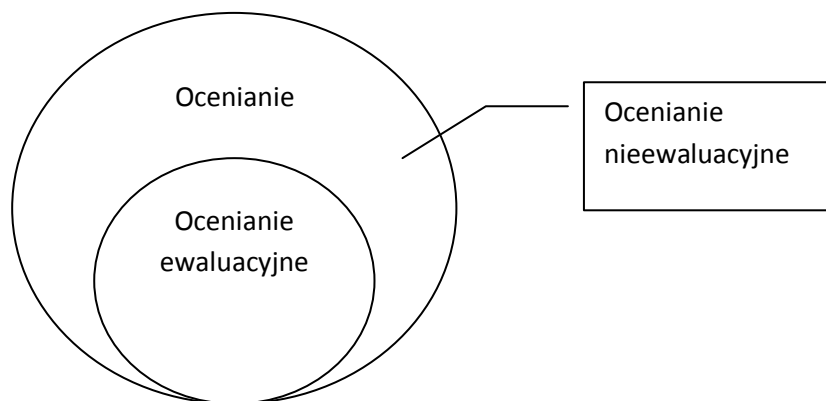
Nasze słowa najbliższe znaczeniowo łacińskiemu *valeo-ere*, angielskim *value, valuation* czy niemieckim *der Wert, bewerten* to m.in. czasowniki „wartościować”, „oceniać”, „waloryzować” i rzeczowniki „wartość”, „ocena”, „valor” („waloryzacja”), a w zastosowaniach przymiotnikowych „wartościowy” (co jednak od razu sugeruje znaczną wartość, mamy też rzeczownik „wartościowość”), „oceniony”, rzadziej „ocenny” w ogólniejszym sensie, „waloryzowany” lub „zwaloryzowany”.

Warto zaznaczyć, że angielskie *valuation* tłumaczone jest na język polski też jako „wycena”, co bezpośrednio kieruje uwagę na wartościowanie czy ocenianie czegoś w kategoriach pieniężnych. Stąd też dobrze znane pytania „ile to jest warte” w znaczeniu „jaką ma to cenę”. O ile „wycena” i „cena” są raczej pospolite w użyciu ze względu na kontekst handlowy czy rynkowy, to o „wartościach” i „wartościowaniu” trudniej powiedzieć to samo. Przypisywanie ceny temu, co uznawane jest za wartość może być nawet potępiane, np. wycenianie miłości, przyjaźni, życia, ojczyzny w pieniądzech, tak jakby były na sprzedaż.

W tej książce używamy słów „ewaluacja”, „ewaluowanie”, „ewaluacyjny” w znaczeniu szczególnego rodzaju oceniania lub wartościowania, tzn. te słowa mają szerszy zakres znaczeniowy. Stąd też ewaluowanie może być nazywane mniej poręcznie „wartościowaniem ewaluacyjnym” lub „ocenianiem ewaluacyjnym”, co sygnalizuje właśnie, że są też inne rodzaje wartościowania i oceniania niebędące jednak ewaluowaniem.

Schemat 1. Ewaluowanie a ocenianie

¹ Przykładem jest niedawno wydana książka Tadeusza A. Grzeszczyka pt. *Ocena projektów europejskich 2007-2013*, Placet, Warszawa 2009.



Źródło: opracowanie własne

Żeby lepiej zrozumieć odróżnianie oceniania ewaluacyjnego od nieewaluacyjnego przyjrzyjmy się tabeli 1, gdzie pokazano różnicę między jednym z rodzajów tego drugiego a ewaluowaniem pod kilkoma względami.

Tabela 1. Ocenianie potoczne a ocenianie ewaluacyjne

Wymiary porównania	Ocenianie potoczne (nieewaluacyjne)	Ocenianie ewaluacyjne
Kto dokonuje oceny?	Każdy, amator	Profesjonalista
Jakie są kryteria oceny?	Nieustalone, niezdefiniowane, przypadkowe, niejasne, zmienne	Ustalone, zdefiniowane, określone przed przystąpieniem do procesu zbierania informacji i oceniania
Informacje o przedmiocie oceny	Wyrywkowe, zasłyszane, subiektywne	Rzetelne, obiektywne, zebrane metodycznie i według kryteriów oceny
W jaki sposób dokonuje się oceny?	Intuicyjny, momentalny, przypadkowy	Metodyczny, przemyślany, czasochłonny, według ustalonych procedur porównywania informacji o przedmiocie oceny z kryteriami oceny i wyciągania wniosków z takich porównań według ustalonych zasad

Źródło: opracowanie własne

Uznanie, że ewaluowanie jest jednym z rodzajów oceniania pomaga nam umieścić to obce słowo w znajomym kontekście. Wynika stąd również, że cechy definicyjne oceniania jako takiego odnoszą się również do ewaluowania.

Podstawowe elementy ewaluowania

Jak każdą inną czynność możemy ewaluowanie analizować pod względem kilku typowych charakterystyk: 1) podmiot (kto ewaluuje), 2) przedmiot (co jest ewaluowane), 3) cel (po co się ewaluuje), 4) sposób lub metodę (jak się ewaluuje), 5) narzędzia (za pomocą czego przeprowadza się ewaluację), 6) miejsce (gdzie się ewaluuje), 7) czas (kiedy się ewaluuje), 8) skutki (jakie skutki miało ewaluowanie) itd.

Szczególną rolę dla zrozumienia oceniania, a w tym oceniania ewaluacyjnego, odgrywają kryteria oceny. Załóżmy, że ktoś poprosił nas o ocenę naszego dziecka, lekarza, który nas leczył albo o ocenę naszego mieszkania czy posiłku w restauracji. Spełnienie takiej ogólnej prośby jest trudne, gdyż nie sprecyzowano, pod jakim względem mamy dokonać oceny, a inaczej mówiąc: według jakich kryteriów. Ocena naszego dziecka może dotyczyć jego postaw, poglądów, zachowań, zdrowia, sprawności itp. Z kolei zakres oceny lekarza, który nas leczył może dotyczyć jego samego, np. czy był przystojny, czy był miły, jak i usługi, której nam udzielił - jak nas leczył i co to ostatecznie dało. Kolejne przykłady nie dotyczyły ludzi, ale mieszkania, które można oceniać między innymi pod względem wielkości, wyposażenia, wewnętrznej struktury, trwałości, położenia, ceny. Posiłek mógł być zaś oceniony pod względem smaku, kaloryczności, wyglądu, zapachu, ceny itp.

Skąd się biorą kryteria tego rodzaju, skąd je czerpiemy? Najprostsza odpowiedź na to pytanie - z naszych własnych i/lub ogólnych przekonań dotyczących tego, jakie są pożądane i niepożądane cechy ocenianego człowieka lub przedmiotu, z wzorów i antywzorów odnoszących się do przedmiotu oceny, np. wzór dobrego dziecka czy lekarza, albo mieszkania z powyższych przykładów. Wzory takie bywają zmienne i zależą od kontekstu dokonywania oceny, np. inny wzór dobrego lekarza może mieć na względzie pacjent, a inny ktoś, kto decyduje o wydatkach na ochronę zdrowia, w tym na pensje lekarzy.

Konieczność sprecyzowania kryteriów oceny nie prowadzi do wniosku, że nie jesteśmy w stanie dokonać oceny ogólnej (tzn. bez wskazania kryteriów), gdy się nas o to prosi. Im jednak bardziej złożony przedmiot oceny, tym mniej informacji niesie ze sobą jego ogólna ocena i tym mniej jest ona użyteczna. Należy odróżniać od siebie ocenę ogólną od oceny łącznej, która wydawana jest na podstawie ocen pod względem poszczególnych kryteriów (oceny cząstkowe). Procedura uzyskiwania oceny łącznej nazywana jest niekiedy agregowaniem lub syntezą ewaluacyjną (patrz rozdział o metodach oceniania).

Ocenianie jest rodzajem ludzkich czynności, a ewaluowanie jest rodzajem oceniania. Może być więc analizowane tak jak inne działania ludzi oraz tak jak ocenianie, ale ma też cechy specyficzne w stosunku do oceniania jako takiego. Z kolei samo ewaluowanie może dotyczyć różnych przedmiotów, mieć różne cele i odbywać się w różnych sferach i kontekstach, np. w sferze zainteresowań służb społecznych, co nadaje mu kolejne cechy specyficzne (patrz rozdział o typach ewaluowania). Część powyższych informacji została uporządkowana i podsumowana w tabeli 2.

Tabela 2. Podstawowe elementy ewaluacji

Ewaluowanie	Pytanie	Nazwa
Podmiot	Kto ewaluuje?	Ewaluator
Przedmiot	Co jest poddawane ewaluowaniu?	Ewaluandum, przedmiot lub obiekt ewaluacji
Kryteria	Według jakich kryteriów ewaluowane?	Kryteria ewaluacyjne
Cel	W jakim celu ewaluowano?	Cele ewaluowania
Sposób, styl	Jak ewaluowano?	Metoda ewaluacji
Narzędzia	Za pomocą czego ewaluowano?	Instrumenty ewaluowania
Wytwór	Co jest produktem ewaluowania?	Ewaluacja (w sensie wystawionej oceny, oszacowanej jakości i wartości czegoś)
Wpływ, konsekwencje, skutki	Jaki wpływ miało ewaluowanie i jego wytwory na....?	Wpływ ewaluowania na podmiot, przedmiot, otoczenie ewaluowania

Źródło: opracowanie własne

Czytelnik powinien pamiętać, że ewaluowanie jest złożonym działaniem o charakterze procesu (trwa w czasie, składa się z różnych czynności, może być dzielone na fazy lub etapy, patrz rozdział dotyczący ewaluacji jako procesu), którego bezpośrednim wytworem czy produktem jest ocena ewaluacyjna tego, co było przedmiotem ewaluowania. Wiemy już, że kryteriów ewaluowania może być wiele, więc i ocen wydanych na ich podstawie może być więcej niż jedna. Możemy jednak próbować poddać te oceny agregacji, dokonać ich syntezy i uzyskać w jej wyniku ocenę łączną.

Co to znaczy wydać ocenę ewaluacyjną kogoś lub czegoś? Najprościej wyobrazić sobie to jako wyznaczenie ewaluowanemu przedmiotowi pozycji na skali oceny. Jedną z bardziej znanych skal tego rodzaju jest skala ocen stosowana w wielu polskich szkołach: celujący, bardzo dobry, dobry, dostateczny, mierny, niedostateczny (ewentualnie z plusami i minusami, co nie dotyczy oceny ostatecznej o szczególnym znaczeniu). Może być ona łatwo przekształcona w skalę liczbową od 6 do 1, gdzie liczby odpowiadają kolejnym ocenom. Liczby poddają się operacjom matematycznym, takim jak uśrednianie, co ułatwia analizę ilościową.

Ewaluowanie prowadzi więc do przypisania temu, co oceniane oceny ewaluacyjnej. Ważne jest, abyśmy pamiętali, że jest to złożony i metodyczny proces wymagający wielu rodzajów zasobów, w tym czasu, wiedzy i umiejętności. W kontekście edukacyjnym jest to czas, wiedza i umiejętności ewaluacyjne nauczycieli, którzy oceniają postępy uczniów w nauce. Analogicznie w działaniu służb społecznych przedmiotem ewaluowania mogą być postępy klientów na drodze do integracji społecznej. Na podstawie osiągnięć edukacyjnych uczniów i osiągnięć integracyjnych wspomaganym klientom, dokonuje się ocen pracy nauczycieli i pracowników służb społecznych oraz instytucji, które ich zatrudniają. Im lepsza praca nauczyciela czy pracownika socjalnego, tym lepsze powinny być osiągnięcia tych, którym udzielają oni swoich usług.

Ewaluacja działania

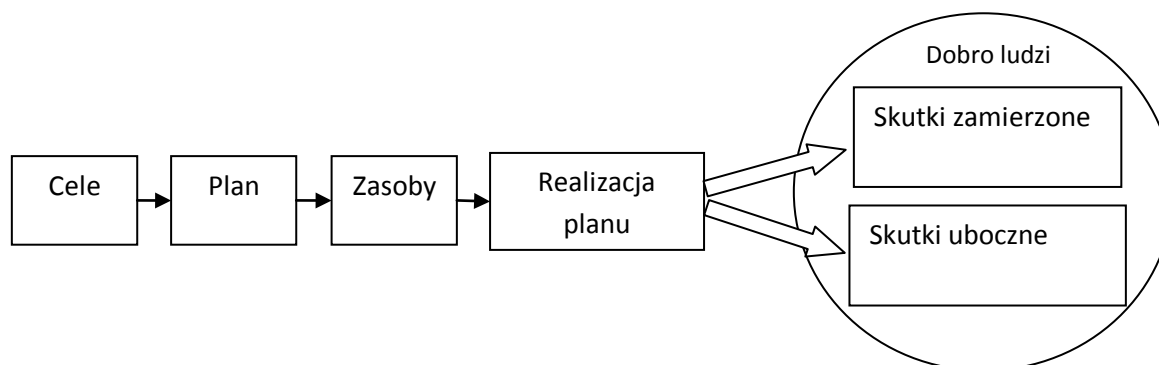
Specyficznym przedmiotem ewaluacji są ludzkie działania, czyli czyjeś świadome i celowe wysiłki, które mają doprowadzić do osiągnięcia czegoś, co zostało zamierzone (cele) i czego osiągnięcie nie jest łatwe i natychmiastowe. W tej książce głównie o takich przedmiotach mowa, w szczególności chodzi o ewaluowanie działań służb społecznych.

Świadomość i celowość zakładają, że na początku są mniej lub bardziej jasne zamierzenia w postaci celów, jakie ktoś chce osiągnąć za pomocą działania (inaczej mówiąc są to skutki zamierzone działania). W przypadku służb społecznych, udzielających pomocy i wsparcia, chodzi o cele tych działań, np. uzyskanie integracji społecznej wspomaganym, zwiększenie poziomu zaspokojenia ich potrzeb. Są to bardzo ogólne określenia, które wymagają doprecyzowania, aby stać się użytecznymi w praktyce działania, np. w danym przypadku integracja społeczna będzie polegać na odzyskaniu wiary w siebie, pogodzeniu się z rodziną, wyleczeniu się z choroby, ukończeniu szkoły, znalezieniu i utrzymaniu (samo)zatrudnienia, spłaceniu długów itd. Im trudniejszy przypadek, tym trudniejsze jest to zadanie. W praktyce trzeba znaleźć jego rozwiązanie (a może być ich kilka, więc dochodzi jeszcze problem wyboru najlepszego z nich) i jeszcze wdrożyć je w życie. Wymaga to planowania i uwzględnienia koniecznych zasobów, np. czas, energia i informacje konieczne dla czynności służb społecznych.

Realizacja planów przy pomocy posiadanych i/lub dodatkowo zdobytych zasobów ma dać w efekcie osiągnięcie celów (czyli skutków zamierzonych), ale przynosić może też skutki uboczne (są to niezamierzone skutki działania) – wielu krytyków działania służb społecznych zatrudnionych w

publicznej pomocy społecznej opiera się na założeniu, że mają one szkodliwe skutki uboczne, które są większe niż osiągnięte skutki zamierzone.

Schemat 2. Logika świadomego działania



Źródło: opracowanie własne

W skrócie mówiąc, chcemy coś osiągnąć i określamy to w pewien sposób (formułowanie celów), a następnie zastanawiamy się, jak to osiągnąć (wybór strategii, konstruowanie planu). W kolejnych krokach zdobywamy konieczne do tego zasoby (jeżeli ich nie mamy) i przystępujemy do działania, które przyczynia się do powstawania skutków, o które nam chodziło. Zdajemy sobie przy tym sprawę, że możliwe są też skutki uboczne, których możemy świadomie unikać (w szczególności, gdy są one negatywne).

Zastosujmy ten schemat w kontekście działania służb społecznych. Chcemy pomóc konkretnej osobie (rodzinie, społeczności) w zaspokojeniu jej potrzeb, a więc konkretyzujemy, o jakie potrzeby chodzi (diagnoza luki między tym, co jest, a tym co może i powinno być) i zastanawiamy się, jak pomóc, co prowadzi do sformułowania planu pomocy. Jego realizacja wymagać będzie zaangażowania samego wspomaganego, innych ludzi, a także środków i zasobów, które powinny być również w miarę precyzyjnie określone.

Podkreślenia wymaga, że służby społeczne są to ludzie zatrudnieni przez społeczeństwo do realizowania pewnych jego celów (społecznych). Cele te mogą być dość ogólnie sformułowane (np. w ustawach), a swoboda w określaniu tego, jak je można osiągać może być duża, a więc autorami i realizatorami konkretnych działań są służby społeczne. Jest to przypadek dużej autonomii służb, im jej mniej tym bardziej służby są tylko realizatorami działań zaprojektowanych przez kogoś innego.

Z drugiej strony mamy adresatów działań, czyli tych, którym mają one bezpośrednio przynieść korzyść (pozytywny wpływ na ich dobro), stąd też nazwa „beneficjent”. Określenie „adresat” jest ogólniejsze, można je stosować nawet dla takich programów, które mają na celu zmniejszenie czyjegoś dobra, np. działania z zakresu karania, trudno bowiem beneficjentem nazwać kogoś, kto został słusznie ukarany. Nie zakłada ono również, że ktoś rzeczywiście odniósł jakieś korzyści – to, że coś w założeniach ma powiększać czyjeś dobro, nie znaczy, że w efekcie się do tego przyczyni – beneficjent założony niekoniecznie staje się beneficjentem faktycznym. Poza adresatami mamy też ich bliższe i dalsze otoczenie społeczne, czyli tych, których dobro jest powiązane z dobrem bezpośrednich adresatów (rodzina, przyjaciele, znajomi, współpracownicy, współlokatorzy, sąsiedzi itp.). Można ich nazywać adresatami lub beneficjentami pośrednimi.

Podstawowe sposoby ewaluacji działania służb społecznych

Na podstawie schematu zamieszczonego w poprzednim punkcie można omówić kilka najważniejszych podejść do oceny ewaluacyjnej działania służb społecznych.

Po pierwsze, działania ludzi ocenia się na podstawie oceny ich celów (intencji, zamierzeń). Najważniejszym kryterium oceny w tym przypadku jest dobro ludzi, w szczególności dobro tych, którzy są klientami czy podopiecznymi służb społecznych (ogólniej – tych, na których dobro działanie to ma mieć i ma wpływ). Jeżeli będziemy je rozumieli w kategoriach zaspokojenia potrzeb, to cele powinny być takie, aby zamierzone skutki przyczyniały się do tego w możliwie największym stopniu. W tym przypadku mowa o ocenianiu działania pod względem jego **użyteczności** dla klienta służb społecznych lub dla adresata usług przez nie udzielanych. Zasadniczym pytaniem ewaluacyjnym w tym przypadku jest, jak należy ocenić wpływ działania o danych celach na dobro tych, którym ma ono służyć (dobro to może być różnie rozumiane, np. jako zaspokojenie potrzeb, integracja społeczna)? Ewaluowaniu pod względem użyteczności mogą też podlegać osiągnięte skutki działania bezpośrednio lub pośrednio. Może się tak zdarzyć, że mimo wysokiej oceny użyteczności celów, ocena użyteczności skutków jest dużo niższa. Dzieje się tak z wielu względów, np. w procesie realizacji jedynie połowicznie osiągnięto to, co zamierzano, albo uwzględnienie skutków ubocznych wpłynęło na obniżenie oceny użyteczności.

Po drugie, jeżeli cele działania ocenimy jako użyteczne w co najmniej odpowiednim stopniu, to pozostaje osiągnięcie określonych przez te cele skutków (wkładanie wysiłku w osiąganie celów nieużytecznych jest marnotrawstwem). W praktyce jest to trudniejsza część zadania – łatwiej jest ambitne, szlachetne i użyteczne cele formułować na papierze niż potem osiągnąć zbliżone do nich skutki. Stąd też istotnym sposobem oceniania działania jest porównywanie osiągniętych skutków z tymi, które zostały zamierzone i określone w celach.

Zauważmy, że porównywanie osiągniętych skutków (w tym ubocznych) z dobrem adresatów działania może się odbywać z pominięciem tego, co zamierzali jego autorzy (ocena użyteczności skutków). Prowadzi to niekiedy do argumentacji, iż można ewaluować działania abstrahując od ich celów (zamierzeń), które w dodatku mogą być ogólnikowe, sprzeczne ze sobą itd. Jeżeli skutki osiągnięte znacząco odbiegają od celów to działanie ocenimy jako mało skuteczne lub nieskuteczne (niska ocena na skali), ale nadal jego ocena pod względem użyteczności może być wysoka.

Załóżmy, że celem działania służb społecznych było usamodzielnienie wspomaganego, czego praktycznym wyrazem miało być znalezienie przez niego pracy i utrzymanie jej w okresie co najmniej pięciu lat. Jeżeli w wyniku działania doszło do tego, to działanie powinno być ocenione jako maksymalnie skuteczne. Możliwe jednak, że został on zatrudniony, ale trwało to półtora roku, a nie pięć lat. Czy oznacza to, że działanie to należy ocenić na niedostateczny, jako całkowicie nieskuteczne? Odpowiedź na to pytanie zależy już od tego, czy zamierzone pięcioletnie zatrudnienie było progiem minimalnym dla oceny pozytywnej tego działania pod względem jego **skuteczności**, czyli poziomu zgodności osiągniętych skutków ze skutkami zamierzonymi określonymi poprzez cele. Zasadniczym pytaniem ewaluacyjnym w tym przypadku jest, jak należy ocenić poziom osiągnięcia celów działania?

Po trzecie, osiągnięcie zamierzonych skutków wymaga zużycia zasobów pracy, narzędzi i infrastruktury, które kosztują. Jeżeli cele i osiągnięte skutki (z uwzględnieniem ubocznych) były użyteczne (zgodne z dobrem jego adresatów), a działanie skuteczne (osiągnięto skutki zgodne z użytecznymi celami) to pozostaje jeszcze ocena tego działania pod względem zużycia zasobów lub jego kapitałochłonności czy kosztowności. Tu również możemy pominąć cele działania i odnosić

ocenę tego rodzaju do użyteczności – ile kosztowało nas osiągnięcie określonego wpływu na dobro danej osoby czy grupy.

Zalóżmy, że celem służb społecznych było usamodzielnienie wspomaganego, czego wyrazem miało być znalezienie przez niego pracy i trwałe jej utrzymanie. Można to osiągnąć działając w różny sposób, co najmniej część z tych sposobów będzie się różniła kosztownością (ceną), czyli jedne będą droższe, np. długotrwała psychoterapia połączona ze zdobywaniem kwalifikacji w szkole, inne – tańsze, np. porada i nieformalne szkolenie. Przy takiej samej ocenie użyteczności i skuteczności lepiej powinniśmy oceniać działania tańsze (mniej zasobochłonne), gdyż oszczędzamy wtedy ograniczone zasoby ludzkiego czasu i energii, a także pieniędzy publicznych. W taki sposób ocenia się działania służb społecznych pod względem **ekonomiczności**, ewentualnie oszczędności (koszty w porównaniu ze skutecznością lub użytecznością). Częściej używanym słowem jest zapewne „efektywność”. Może być ono jednak mylące, gdyż skutek i efekt to słowa bliskoznaczne. W ocenie ekonomiczności chodzi o coś innego niż o użyteczność czy skuteczność, a mianowicie o osiągnięcie wpływu na dobro ludzi możliwie najmniejszym kosztem, co jest charakterystyczne dla podejścia ekonomicznego. W tym przypadku główne pytanie ewaluacyjne brzmi, jak należy ocenić koszty osiągnięcia celów lub koszty poprawy dobra adresatów działania?

Warto podkreślić, że ocena użyteczności i ekonomiczności działania może się odbywać niezależnie od celów, jakie mieli sponsorzy, autorzy czy wykonawcy działania. Jeżeli mamy wykonać zadanie dokonania ewaluacji pewnego działania wystarczy nam informacja o tym, jaki miało ono wpływ na dobro adresatów działania (informacja o skutkach działania niezależnie od tego, które były zamierzone, a które uboczne) oraz informacja o zasobach jakie zaplanowano i/lub zużyto podczas działania.

Ocenianie działania za pomocą tylko jednego z powyższych sposobów jest błędem. Duża użyteczność może być osiągnięta nieoszczędnie – w skrajności: osiągnęliśmy zasadniczą poprawę dobra klientów, ale zużyliśmy całe zasoby, czyli nie możemy dalej pomagać kolejnym potrzebującym lub ta pomoc jest zasadniczo gorszej jakości. Z kolei duża skuteczność może być mało użyteczna – w skrajności, osiągnięte skutki działania są zgodne z celami, ale nie wpłynęło to pozytywnie na dobro klienta (cele nie były użyteczne). Duża ekonomiczność czy oszczędność działania może być również mało użyteczna, jeżeli brano pod uwagę jedynie koszty osiągnięcia skuteczności – działaliśmy oszczędnie i osiągnęliśmy cele, ale to, co osiągnęliśmy miało niewielki wpływ na dobro adresatów działania. Ponadto zbyt niski poziom zasobów w stosunku do uzasadnionych potrzeb programu może odpowiadać za jego słabe wyniki lub porażkę.

Trzy powyżej omówione sposoby ewaluowania działania służb społecznych nie są jedynymi, ale wydają się być najważniejsze dla zrozumienia tego, czym jest ewaluowanie i ewaluacja. Można te działania oceniać również pod względem zakresu, w jakim ci, dla których dobra je uruchomiono uczestniczyli w ich projektowaniu i realizowaniu. Nie trzeba chyba wspominać o tym, że działania oceniane są również pod względem zgodności z prawem (patrz rozdział o kryteriach ewaluacji).

Tabela 3. Podstawowe podejścia do ewaluacji działania

Co porównujemy?	Z czym porównujemy?	Co jest oceniane na podstawie porównania?
Cele	Spodziewane zaspokojenie potrzeb (ogólniej - spodziewany wpływ działania na dobro zamierzonych i pośrednich	Użyteczność ex ante (przed przystąpieniem do realizacji)

	adresatów)	
Skutki (zamierzone i niezamierzone)	Zaspokojenie potrzeb (ogólniej - osiągnięty wpływ na dobro zamierzonych i pośrednich adresatów)	Użyteczność ex post (po zakończeniu działania)
Cele	Skutki zamierzone	Skuteczność
Zaspokojenie potrzeb osiągnięte dzięki działaniu, wpływ na dobro ludzi	Zasoby zużyte podczas działania	Ekonomiczność

Źródło: opracowanie własne

Cele i funkcje ewaluowania

W zrozumieniu czym jest ewaluowanie pomaga nam również wskazanie celów i funkcji, jakie ma to działanie i jego bezpośredni produkt, czyli ocena ewaluacyjna. Może ona być zastosowana w pracy służb społecznych, analogicznie do tego, jak jest stosowana w szkolnictwie. Pomyślmy o ocenie postępów klienta w procesie integracji społecznej oraz o przydatności tej oceny. Z kolei ocena pracy służb społecznych może posłużyć do celów związanych z planowaniem i zarządzaniem.

Załóżmy, że postępy w integracji społecznej czy działania pomocowe służb społecznych zostały ocenione na dostateczny. Do czego i komu może się taka informacja przydać? Najprostsza odpowiedź jest taka: do tego, aby zmienić w ewaluowanym działaniu to, co odpowiada za jego niską ocenę ewaluacyjną. Innymi słowy, chodzi o doskonalenie pomocy udzielanej przez służby społeczne, o poprawę jakości ich usług.

W tabeli przytoczone zostanie kilka przykładów wyliczeń celów ewaluowania w takim rozumieniu.

Tabela 4. 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 	Melvin M. Mark i in. ²
<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 	Eleanor Chelimsky ³

² M. M. Mark, G. T. Henry, G. Julnes *Evaluation: An Integrated Framework for Understanding, Guiding, and Improving Policies and Programs*, Jossey-Bass, San Francisco 2000, s. 54-58.

³ E. Chelimsky *The Purposes of Evaluation in a Democratic Society*, w: I. Shaw, J. C. Greene, M. M. Mark (red.) *The SAGE Handbook of Evaluation*, SAGE Publications, Thousand Oaks 2006, s. 39.

<ul style="list-style-type: none"> • Wszechstronna i całościowa ocena programu • Uczenie się i poprawianie funkcjonowania programu • Rozliczenie z 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 	Michael Q. Patton ⁴
---	--------------------------------

W polskiej literaturze) zamiast o celach ewaluacji pisze się też o jej funkcjach. 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 czy pragnienia to zaspokaja.

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, pozytywne skutki i konsekwencje istnienia czegoś) i dysfunkcje (w negatywnym sensie, negatywne skutki i konsekwencje). Niewiele dotąd wiemy o dysfunkcjach ewaluacji, ale z pewnością nadmiar aktywności ewaluacyjnej w społeczeństwie może doprowadzić do skutków, które przez większość zaczną być odczuwane jako przynajmniej dokuczliwe.

Przykładowe zestawienia głównych funkcji ewaluacji zostały przedstawione w kolejnej tabeli.

Tabela 5. Wybrane zestawienia funkcji ewaluacji

Funkcje	Autor
<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 	David Nevo ⁵
<ul style="list-style-type: none"> • Wspomaganie decyzji związanej z zarządzaniem programami społecznymi • Określenie jakości i wartości programu dla tych, którym ma on służyć 	Michael Scriven ⁶ (autor argumentował, że główną funkcją powinna być druga z nich)
<ul style="list-style-type: none"> • Budowanie partnerstwa i współwłasności • Poprawa wdrażania i bieżącej jakości • Racjonalizowanie planowania • Funkcja poznawcza • Rozliczanie i wzmocnienie odpowiedzialności 	Karol Olejniczak ⁷
<ul style="list-style-type: none"> • Poprawa planowania przez zagwarantowanie racjonalnego uzasadnienia dla danej interwencji 	Katarzyna Ornacka ⁸

⁴ M. Q. Patton *Utilization-Focused Evaluation*, SAGE Publications, Thousand Oaks 2008, s. 139-141.

⁵ D. Nevo *The Conceptualization of Educational Evaluation: An Analytical Review of the Literature*, Review of Educational Research, tom 53, nr 1, 1983, s. 119.

⁶ M. Scriven *Hard-won lessons in program evaluation*, Jossey-Bass, San Francisco 1993, s. 9.

⁷ K. Olejniczak, M. Kozak, B. Ledzion (red.) *Teoria i praktyka ewaluacji interwencji publicznych. Podręcznik akademicki*, Wydawnictwa Akademickie i Profesjonalne, Akademia im. Koźmińskiego Warszawa 2008, s. 30-31

⁸ K. Ornacka *Rola ewaluacji w pomocy społecznej*, w: J. Starega-Piasek, A. Hryniewicka (red.) *O potrzebie ewaluacji w pomocy społecznej*, Instytut Rozwoju Służb Społecznych, Warszawa 2008, s. 28.

<ul style="list-style-type: none"> • 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 	
---	--

Obok pozytywnie ocenianych funkcji ewaluacji, które 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 nich⁹:

- odwlekание 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 w powyższej 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.

Ewaluowanie a badania naukowe, kontrolowanie i monitorowanie działań

Badania naukowe tym się różnią od ewaluowania, że ich celem podstawowym nie jest ocena działania służb społecznych, które jest ich przedmiotem. Służą one przede wszystkim opisowi, klasyfikacji oraz wyjaśnianiu poddanego badaniu działania, a w skrócie chodzi o jego poznanie. Stosuje się w tym celu rozmaite metody i techniki zbierania i analizowania danych o różnym charakterze. Ze względu na to, że ewaluacja ma być obiektywnym sposobem określania oceny działań, wykorzystuje się w niej naukowe metody, w szczególności do zbierania i analizowania informacji o ewaluowanym przedmiocie. Stanowią one potem podstawę informacyjną wystawiania ocen, czemu służą już metody ściśle ewaluacyjne (patrz rozdział o metodach oceniania). Przykładowo, żeby określić wpływ danego programu działań służb społecznych na dobro klientów należy stwierdzić, czy i jaki to działanie wywarło wpływ, jakie skutki dla dobra klientów miało (badanie zależności przyczynowo-skutkowych).

Niektóre z dyscyplin nauki bardzo blisko wiążą się z ideą ewaluacji. W klasycznej już koncepcji nauki dla polityki społecznej, w tym dla programów działania służb społecznych, miała ona

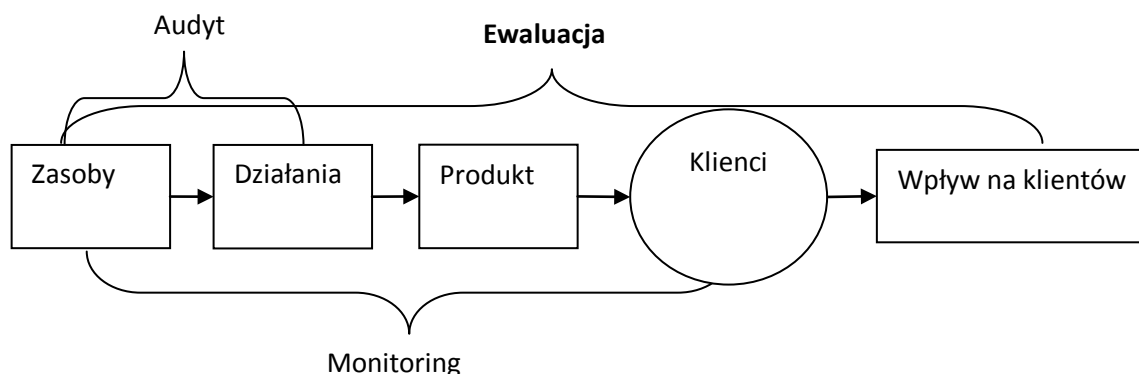
⁹ Carol H. Weiss *Evaluation*, Prentice Hall, Upper Saddle River 1998, s. 22.

pomagać w ich formułowaniu. Wymagało to opracowania procedury diagnozowania, prognozowania, planowania oraz kontrolowania realizacji planów¹⁰. Swojskie kontrolowanie, które miało służyć korygowaniu programów i planów, później znika i zostaje zastąpione zapożyczeniami: ewaluacją i monitoringiem. W obu przypadkach przypisany został im podobny cel: korygowanie programów w trakcie realizacji.¹¹

Produktem działalności Najwyższej Izby Kontroli są raporty, w których opisowo ocenia się działania podmiotów sektora publicznego pod względem celowości, gospodarności, legalności i rzetelności. Jest to więc działalność podobna do tego, co składa się na ewaluację.

Słowo „audyt” (lub w kontekście norm ISO – „audit”) to kolejne zapożyczenie, które zaczęło zastępować nasze pojęcie kontroli. W ujęciu sugerowanym przez Komisję Europejską (KE), audyt dotyczy głównie zgodności realizacji programu z formalnymi procedurami i prawem, monitoring obejmuje również bezpośredni produkt realizacji działań lub udzielane usługi (np. komu, ile i jakich usług udzielono). Ewaluowanie jest działaniem o szerszych celach, które wymaga zarówno czynności audytorskich, jak i polegających na monitoringu, ale także uwzględnia wpływ działań na tych, którzy mieli z nich korzystać (to decyduje o specyfice ewaluacji). Szczególnie użyteczny dla działalności ewaluacyjnej jest monitoring rozumiany jako dokumentowanie tego, co się działo podczas realizacji programu, a także, jako ustanowienie systemu wskaźników, za pomocą których monitorowane są wyniki programu. Uprozczone zależności między audytem, monitoringiem i ewaluacją w ujęciu KE przedstawione zostały na schemacie.

Schemat 3. Ewaluacja a audyt i monitoring



Źródło: opracowanie własne w na podstawie KE

Zakończenie

Tłumaczenie angielskiego słowa *evaluation* poprzez polską „ocenę” ma podstawy w tym, że ewaluowanie jest rodzajem oceniania. Jest to działalność, która wykonywana metodycznie i przez biegłych w niej ewaluatorów ma dostarczyć rzetelnej i wiarygodnej oceny programów działań służb społecznych, Jest ona potrzebna między innymi dla doskonalenia pracy tych służb, zwiększania jej wartości i jakości dla tych, którzy mają być jej głównymi beneficjentami i dla społeczeństwa, które zdecydowało się za to płacić.

¹⁰ Jolanta Supińska *Dylematy polityki społecznej*, Instytut Polityki Społecznej, Warszawa 1991, s. 10

¹¹ Jolanta Supińska *Rozwój i zadania nauki o polityce społecznej*, w: Antoni Rajkiewicz, Jolanta Supińska, Mirosław Książkowski (red.) *Polityka społeczna. Materiały do studiowania*, Wydawnictwo „Śląsk”, Katowice, 1998, s. 38.

Należy przewidywać, że kultura ewaluacyjna będzie się dynamicznie rozwijać w Polsce w najbliższych latach. Jest to związane z wieloma procesami reformatorskimi jakim podlega polityka społeczna, mająca decydujący wpływ na praktykę działań służb społecznych. Jednym z najważniejszych jest coraz większy pluralizm pomiotów zapewniających usług społeczne obywatelom, przy zachowaniu ich finansowania publicznego. Powoduje to, że wzrasta znaczenie konkursów i przetargów, gdzie ewaluacji wymagają składane oferty, a następnie realizowane w ich wyników programy działań społecznych. W coraz większym stopniu będzie dotyczyło to również usług związanych z pracą socjalną i działaniami służb społecznych.

Lista haseł do skorowidza przedmiotowego

Użyteczność

Skuteczność

Ekonomiczność

Audyt

Monitoring

Cele ewaluacji

Funkcje ewaluacji

Funkcje ukryte ewaluacji

Skutki zamierzone

Skutki uboczne

Adresaci, beneficjenci zamierzeni

Metody oceniańia

Wstęp

To co odróżnia ewaluację od innych rodzajów systematycznych dociekań związane jest z ocenami i wartościami. Jest to szczególnie interesujące zagadnienie w naukach społecznych i politycznych, których przedmiotem są również działania służb społecznych. Czy badacz powinien je oceniać? Czy z badań naukowych nastawionych na stwierdzanie faktów mogą wynikać oceny i wskazówki, co i jak należy zmienić w praktyce? Odpowiedź na te pytania może być różna i stąd spór o znaczenie wartości w nauce, a ściślej wartości innych niż naukowe (np. prawda, rzetelność, bezstronność, obiektywność), czyli między innymi wolności, równości czy braterstwa (solidarności), ale też skuteczności, efektywności, oszczędności.

Ewaluowanie to rodzaj oceniania, ale z wykorzystaniem całego arsenału metod i technik ściśle naukowych. W związku z tym jest to działalność znajdująca się pomiędzy biegunem całkowicie niezaangażowanej i nieoceniającej nauki a ocenianiem zupełnie pozbawionym podstaw w rzetelnej i obiektywnej analizie.

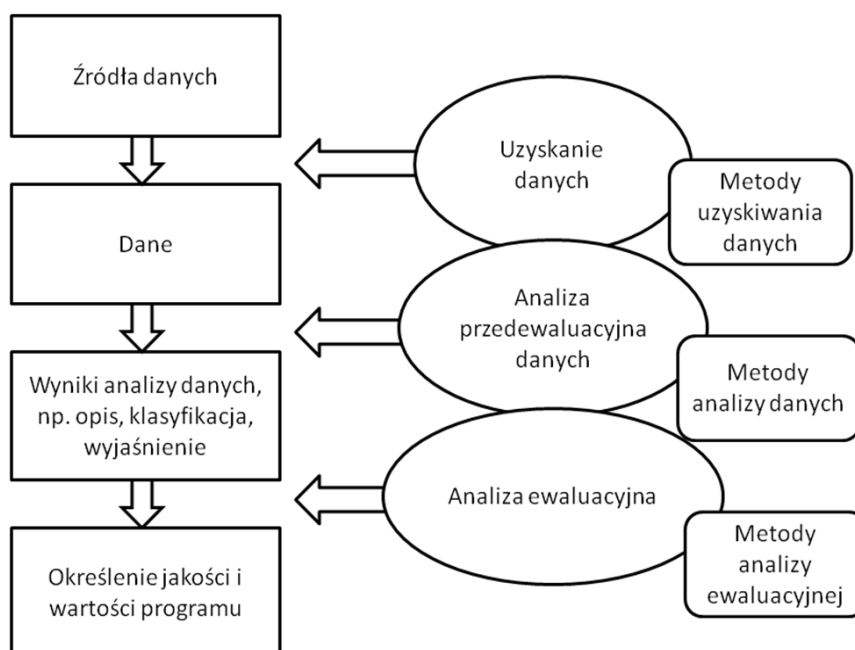
Poniżej zostaną przedstawione podstawowe kroki, które powinny zostać wykonane, gdy mamy dokonać ewaluacji w pełnym tego słowa znaczeniu. Ze względu na ograniczoną objętość prezentacja musiała być uproszczona i mało szczegółowa. Mamy jednak nadzieję, że poniższy materiał pozwoli na osiągnięcie podstawowej orientacji w tym zakresie. Niektóre z narzędzi mogą być bezpośrednio wykorzystywane w praktyce ewaluacyjnej.

Metody badań naukowych a metody oceniania

Skoro ewaluowanie jest czymś innym niż badanie naukowe, to również metody i techniki w obu przypadkach nie powinny być utożsamiane. Nie wystarczy opanowanie sztuki zbierania i analizowania danych o działaniach służb społecznych, aby zostać ewaluatorem, potrzebna jest jeszcze wiedza i umiejętności dotyczące czynności stricte ewaluacyjnych.

Wzajemne powiązania między różnymi poziomami analizy w działalności ewaluacyjnej zostanie pokazane na schemacie.

Schemat 1. Od danych o programie do wniosków o jego jakości i wartości



Źródło: opracowanie własne

Odróżnienie analizy przedewaluacyjnej od stricte ewaluacyjnej jest bardzo istotne. Analiza pierwszego rodzaju wystarcza do odpowiedzi na pytanie, czy dane działanie służb społecznych przyczyniło się (było przyczyną) określonej zmiany w sytuacji grupy klientów objętych usługą. Odpowiedź tego rodzaju jest jednak dopiero wstępem do określenia oceny ewaluacyjnej dotyczącej poziomu wartości i jakości tego działania.

Załóżmy, że stwierdzono, iż pewien program usługowy wywołał zmiany w życiu większości tych, którzy z niego skorzystali. W życiu jednych były to zmiany większe, innych – mniejsze, w życiu pozostałych klientów takich zmian nie zanotowano. W perspektywie ocen mamy tu dwa zagadnienia, po pierwsze, ocena tych zmian (czy były to zmiany na lepsze, na gorsze, czy może neutralne pod tym względem), po drugie, ocena samego programu na podstawie stwierdzonych i ocenionych zmian w życiu klientów. Może być tak, że duże pozytywne zmiany zanotowano u 30% klientów, małe i pozytywne u kolejnych 30%, żadnych zmian u następnych 35%, a 5% pozostałych program przyniósł małe zmiany na gorsze. Jaką na tej podstawie powinniśmy wystawić ocenę ewaluacyjną?

Może pozytywną, gdyż większość zanotowała zmiany na lepsze? Ale program mógł być pod tym względem jeszcze lepszy, więc może pozytywną, ale nie najwyższą. Wynika stąd, że oceny należy stopniować. Skala ocen, jaką wszyscy znamy ze szkoły ma taki charakter w przypadku ocen pozytywnych (od celującej do miernej z ewentualnymi plusami i minusami). Jak więc ocenić nasz przykładowy program na tej skali? Gdzie wyznaczyć próg, poniżej którego zaczyna się ocena niedostateczna? Co zrobić, gdy kryteriów ewaluacyjnych jest więcej niż jedno? Na te pytania odpowiada metodologia wystawiania oceny ewaluacyjnej.

Zagadnienia metod oceny

Rozważmy typowy i znany wszystkim przykład oceniania szkolnego na podstawie testu z dwudziestoma pytaniami zamkniętymi; za trafne zakreślenie odpowiedzi można otrzymać 1 punkt, za nietrafne 0. Wiadomo, że maksymalna liczba punktów za test to 20, a minimalna 0. Wyniki uczniów będą się układały na tej skali, jedni dostaną więcej, inni mniej punktów. Samo stwierdzenie wyniku uczniów nie ma jeszcze charakteru oceny, np. wiemy, że uczeń X dostał 15 punktów, a uczeń Y 5 punktów na skali od 0 do 20. Jak od takiej informacji przejść do wydawania ocen za pomocą sześciostopniowej skali od celującej do niedostatecznej? Sprawę komplikuje fakt, że w ocenianiu szkolnym i pozaszkolnym bierze się pod uwagę nie tylko jeden wynik, np. testów może być kilka, ważne mogą być też inne kryteria, np. aktywność na lekcjach, udział w życiu społecznym w szkole, zachowanie, obecność na zajęciach itd.

Każda z wyliczonych czynności wymaga metodycznego podejścia, co związane jest z metodologią oceniania w ewaluacji:

- 1) wybór kryteriów ewaluacyjnych odpowiednich dla danego przedmiotu (odpowiadamy na pytanie, jakie kryteria oceny powinien mieć dany program działań służb społecznych, pod jakimi względami mamy go ocenić);
- 2) określenie ważności poszczególnych kryteriów (np. kryterium użyteczności wydaje się ważniejsze od kryterium skuteczności, ale o ile ważniejsze? Na jakiej skali mamy to oceniać?);
- 3) określenie miar lub skal potrzebnych do pomiaru wyników ewaluowanego przedmiotu (np. miara skuteczności, miara poziomu zaspokojenia potrzeb);
- 4) podział miar lub skal na przedziały i przypisanie im ocen, ustalenie progów minimalnych ocen pozytywnych (np. wynikowi od 50 do 60 na skali od 0 do 100 odpowiadać będzie ocena dostateczny, wynikowi poniżej 50 na tej skali przypisujemy ocenę niedostateczny itd.);
- 5) po uzyskaniu ocen cząstkowych agregowanie ich do oceny łącznej z uwzględnieniem ustalonych wcześniej wag (np. działanie służb społecznych otrzymało dostateczny pod względem użyteczności i dobry pod względem skuteczności, przy czym waga użyteczności ustalona została jako dwa razy większa od wagi skuteczności – zadanie polega na ustaleniu oceny łącznej tego działania).

Wybór kryteriów i ustalenie ich ważności

Prześledźmy tę problematykę najpierw na przykładzie ewaluacji programu działań służb społecznych. Pierwsze pytanie dla metodyka ewaluacji – w jaki sposób powinniśmy określić kryteria oceny tego programu? Najpierw należy zwrócić uwagę na cele tego programu, żeby stwierdzić, co chciano dzięki niemu osiągnąć. Jeżeli w dokumentacji programu te cele są zapisane, to należy rozważyć, czy kryteria oceny nie powinny być opracowane właśnie na ich podstawie. Może się jednak okazać, że cele nie zostały nigdzie wyraźnie sformułowane, albo, że zostały określone ogólnikowo i nie w odniesieniu do dobra klientów, tylko na poziomie usługi, np. celem było przeszkolenie 100 osób bezrobotnych, zapewnienie opieki dla 20 dzieci.

Czy w takich wypadkach nie można dokonać ewaluacji? Oczywiście można, gdyż program niezależnie od tego, jakie cele mu stawiano musiał wywrzeć pewien wpływ na swoich klientów, np. biorący udział w szkoleniu czegoś się dowiedzieli i coś otrzymali, co z kolei miało (lub nie) dalej idące konsekwencje dla ich życia zawodowego. W razie problemów z oparciem kryteriów na celach, ich źródłem powinna być diagnoza potrzeb bezrobotnych i dzieci z naszych przykładów. Czy i jakich potrzebują szkoleń, czy nie potrzeba im czegoś jeszcze, aby sobie poradzili? Czy i jakiej potrzebują opieki, czy nie potrzeba im czegoś jeszcze, aby dobrze się rozwijały? Na te potrzeby mają odpowiadać działania służb społecznych, staje się więc jasne, że są to kryteria oceny tych działań.

Poza celami i potrzebami należy rozważyć również inne źródła kryteriów. Znajdziemy je w odpowiedziach na pytania: 1) co to znaczy ukończyć szkolenie, albo otrzymać opiekę; 2) jakie są niezbędne składniki dobrego szkolenia lub właściwej opieki; 3) jakiego rodzaju wyniki w sytuacji czy w zachowaniach bezrobotnych i dzieci powinny być poddane ocenie? Inne źródła kryteriów w tym przypadku to: wymogi prawne, etyczne (ogólne, np. sprawiedliwe traktowanie, i zawodowe np. Kodeks Etyczny Polskiego Towarzystwa Pracowników Socjalnych), specyfikacja programu (zgodność z programem szkoleń i opieki), standardy profesjonalne (dotyczące szkoleń czy opieki), logiczne (program szkolenia lub opieki powinien być logicznie skonstruowany), rynkowe (atrakcyjność dla szkolonych i podopiecznych), oceny ekspertów (co do tego, jakie cechy powinno mieć dobre szkolenie, dobra opieka).

Ze względu na to, że jakość kryteriów jest kluczowa dla dokonania dobrej ewaluacji w tabeli został przedstawiony zestaw uniwersalnych i formalnych cech (prakryteriów), które powinny mieć kryteria dobrej jakości.

Tabela 1. Prakryteria dla kryteriów ewaluacyjnych

Nazwa prakryterium	Interpretacja
Jednoznaczność	ma mieć jedno znaczenie, dopuszczające tylko jednoznacznie zrozumiałą interpretację
Swoistość znaczeniowa	nie może być identyczne znaczeniowo z jakimś innym kryterium (czyli ma być nietożsamościowe); nie może być synonimem innych kryteriów
Niezależność	nie może ulegać oddziaływaniu innych kryteriów
Realność	powinno możliwie dokładnie odzwierciedlać rzeczywistość (czyli ma być adekwatne lub dobrze przystosowane do rzeczywistości), nie może być abstrakcyjne
Rzeczowość	ma być bezbłędne merytorycznie, zgodne ze stanem prawnym, z aktualnymi normami i ustaleniami zwyczajowymi
Dopełnialność	powinno stanowić dopełnienie informacyjne w porównaniu z innymi kryteriami
Zwartość nazwy	nazwa powinna być jak najkrótsza, możliwie jednowyrazowa

Zmienność	powinno odtwarzać zmienne stany przedmiotu analizy w aspekcie odpowiedniego wymagania
Mierzalność	powinno być wielkością mierzalną w całym wykorzystywanym przedziale zmienności
Stopniowalność	powinno być cechą stopniowalną
Przedmiotowość	powinno być dostosowane do potrzeb wynikających z celu analizy
Powtarzalność	powinno być aktualne dla takich samych powtarzających się wymagań dotyczących innych przedmiotów analizy
Poprawność językowa	powinno być zgodne ze współczesnym słownictwem języka, w którym dokonuje się opisu przedmiotu analizy

Źródło: Romuald Kolman *Kwalitologia. Wiedza o różnych dziedzinach jakości*, Placet, Warszawa, 2009, s. 209.

Sam dobór kryteriów i zapewnienie, że będą one dobre pod względem merytorycznym i formalnym nie zamyka sprawy. W przypadku ewaluacji wielokryterialnych (a większość ma właśnie taki charakter) może być tak, że niektóre z kryteriów będą wydawać się ważniejsze od innych. Przykładowo, planowe wykonanie wszystkich zadań wydaje się mniej istotne od tego, czy i na ile program przyczynił się do poprawy sytuacji zamierzonych adresatów. Kolejnym zadaniem jest więc uporządkowanie otrzymanego zbioru według ważności, np. od bardzo ważnych do najmniej ważnych. Skale ważności mogą być mniej lub bardziej złożone, od dwustopniowych (ważne, bardzo ważne) czy trzy stopniowych (mało ważne, średnio ważne, bardzo ważne), do nawet dziesięciostopniowych (ważność najmniejsza, bardzo mała, mała, ograniczona, pośrednia, umiarkowana, dość duża, duża, bardzo duża, największa).¹² Jeżeli dany program działań służb społecznych otrzymuje oceny bardzo dobre w oparciu o wyniki pod względem kryteriów mało ważnych, a dostateczne pod względem tych, które są bardzo ważne, to oczywiście powinno mieć wpływ na ocenę łączną.

Wyróżnia się sześć strategii określania hierarchii ważności kryteriów ewaluacyjnych.¹³

- Zebranie opinii interesariuszy lub klientów (np. za pomocą ankiety ze zbiorem kryteriów i skalą ważności – najważniejsze będą te kryteria, które wskazała jako takie większość, liczba głosów decyduje o ważności).
- Oparcie się na wiedzy wybranych interesariuszy (np. wywiady z najlepiej poinformowanymi z pytaniami dotyczącymi ważności).
- Przegląd literatury (w razie, gdy program jest bardzo złożony, a wiedza interesariuszy ograniczona, np. raporty z badań i ewaluacji podobnych programów).
- Wykorzystanie ocen specjalistów (w zakresie ewaluacji podobnych programów i w związku z tym orientujących się, co zwykle odpowiada za ich sukces lub porażkę).
- Wykorzystanie informacji zgromadzonych w trakcie diagnozy potrzeb (dotyczących tego, zaspokojenie których potrzeb ma mniej lub bardziej istotne znaczenie dla powodzenia lub niepowodzenia programu).
- Wykorzystanie teorii programu i informacji dotyczących związków przyczynowo-skutkowych między głównymi komponentami programu a jego pośrednimi wynikami.

Dla określenia ważności poszczególnych kryteriów przydatne może być zestawienie odpowiedzi na dwa ogólne pytania: 1) jak bardzo negatywne konsekwencje dla programu będzie miała jego bardzo niska ocena pod względem danego kryterium; 2) jak bardzo pozytywne konsekwencje dla programu będzie miała jego bardzo wysoka ocena pod względem danego kryterium. Możliwości odpowiedzi i stopień ważności przedstawione zostały w tabeli.

Tabela 2. Wyznaczanie ważności kryteriów ewaluacji

¹² Kolman, op. cit. s. 231.

¹³ E. Jane Davidson *Evaluation Methodology Basics: The Nuts and Bolts of Sound Evaluation*, SAGE Publications, Thousand Oaks, 2005, s. 105.

		jak bardzo negatywne konsekwencje dla programu będzie miała jego bardzo niska ocena pod względem danego kryterium?		
		Niewielkie	Znaczne	Nie do zaakceptowania
jak bardzo pozytywne konsekwencje dla programu będzie miała jego bardzo wysoka ocena pod względem danego kryterium?	Niewielkie	Umiarkowanie ważne	Ważne	Ważne (ustalenie progu)
	Duże	Ważne	Dość ważne	Bardzo ważne (ustalenie progu)
	Maksymalnie duże	Bardzo ważne	Najważniejsze	Najważniejsze (ustalenie progu)

Źródło: E. J. Davidson, op. cit., s. 108.

Ustalanie progów minimalnych dla przynajmniej niektórych kryteriów najlepiej można zrozumieć w przypadku kosztów, np. ogłaszający konkurs na program pomocy rodzinom problemowym informują, że całkowity koszt nie powinien przekroczyć pewnej kwoty w przeliczeniu na jednego beneficjenta. Jest to maksymalny poziom kosztów akceptowalny przez grantodawcę, a więc każdy program z kosztami większymi zostanie odrzucony w procesie ewaluacji projektów. Ustalanie progów na niektórych kryteriach ma charakter miękki i podlega dyskusji, np. poniżej jakiego odsetka klientów, którzy w wyniku programu odniosą zamierzone korzyści, należy uznać ten program za porażkę, czy ma to być 50%, a może 30%. Wiele zależy w takich przypadkach od wyników programów podobnych.

Po ustaleniu listy kryteriów oraz ich odpowiednim doprecyzowaniu w kontekście ewaluowanego programu, wraz z ustaleniem ich ważności, powinniśmy sporządzić tabelę podobną do wzoru pokazanego niżej.

Tabela 3. Zestawienie kryteriów ewaluacyjnych

Nazwa kryterium	Krótki opis i definicja kryterium w kontekście ewaluowanego programu	Miary i skale do pomiaru wyników programu pod względem danego kryterium	Informacja o wielkości progu minimalnego na danej skali, o ile taki przewidziano	Pozycja danego kryterium na wybranej skali ważności
...
...

Źródło: opracowanie własne

Określenie, doprecyzowanie i „zważenie” kryteriów ewaluacyjnych to podstawowe czynności, które powinny być wykonane w tej fazie przeprowadzania właściwej ewaluacji, tj. oszacowania wartości i jakości działań służb społecznych.

Wyznaczanie stopnia wartości i jakości programu

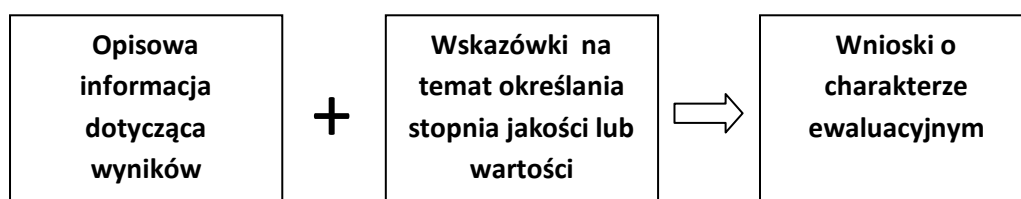
Po zebraniu informacji o ewaluowanym programie działań według opracowanych przez ewaluatorów kryteriów, powinniśmy móc określić wyniki tego programu pod ich względem. Same wyniki nie są jeszcze ocenami. Jak wspomniano wyżej, od tego, że program kosztował 1000 zł na jednego beneficjenta do oceny, że był „tani”, „za drogi”, „umiarkowanie drogi”, „miał dobrą cenę” itp.

nie ma żadnego automatycznego przejścia. Do skali kosztów programu (do każdej z miar i skal) powinny zostać przypisane przedziały ocen i ewentualne progi minimalne (skale oceny), co jest przedmiotem kolejnych metodycznych działań. Nie robi się tego intuicyjnie i z sufitu, stąd istotne znaczenie metodycznego postępowania w tym zakresie.

Dla ewaluacji w pełnym znaczeniu tego słowa jest to najważniejszy krok, gdyż informacji o charakterze opisowym (wynik działania pod pewnym względem) przypisujemy wniosek ewaluacyjny, np. program osiągnął pod względem K wynik x na skali SK, co przy skali ocen SO oznacza, że powinien być on oceniony na ocenę y na tej skali.

Typowym przykładem przejścia od wyników do ocen są praktyki ewaluacji wyników uczniów. Niezależnie od szerszego kontekstu, w którym się to odbywa, praktyczna trudność tego zadania polega na przykład na tym, jak od surowego wyniku z testu wiadomości czy zadań przejść do skali ocen. Maksymalna liczba poprawnych odpowiedzi oznacza maksymalny możliwy wynik testu, łatwo też określić minimum i w ten sposób mamy pełną skalę wyników. Pierwszy krok, to kwestia progu oceny niedostatecznej, która na ogół nie jest stopniowalna, ale możliwe, że jeden uczeń miał wynik tylko nieco niższy niż próg, a dla innego ta luka była bardzo duża. Jeżeli próg wyznaczono na 50% maksymalnego wyniku (co opisowo znaczy, że opanowano połowę zadanego materiału), to w kolejnym kroku powinniśmy przypisać oceny od miernej do celującej lub od dostatecznej do bardzo dobrej wynikom powyżej 50%. Jako, że skala wyników jest ciągła, a skala ocen ma charakter skokowy, ocenom odpowiadać muszą przedziały, np. od 50% do 59% - ocena dostateczna, od 60% do 69% - ocena dostateczna plus, itd. Jeżeli zamiast ocen z testu przyjmiemy, że mowa tu o wynikach pewnego programu, czyli 100% oznacza, że w maksymalnym zakresie pomógł on wszystkim zamierzonym beneficjentom, to zadanie przyjmuje taką samą postać – jaki powinien być próg minimalny, a następnie – jaką przyjąć skalę ocen i jak ją przyporządkować do skali wyników. Ogólną logikę tego postępowania oddaje schemat.

Schemat 2. Od opisu wyników do wniosków ewaluacyjnych



Źródło: E. J. Davidson, op. cit., s. 132.

Powyższe zadanie może być bardziej skomplikowane, gdy mamy więcej miar i skal na podstawie, których dokonuje się pomiaru wyników. Przykładowo, jakość usługi polegającej na pośrednictwie pracy w stosunku do grupy osób może być mierzona szybkością i łatwością z jaką znaleziona została praca, jakością pracy otrzymanej w wyniku pośrednictwa, dopasowaniem miejsca pracy do wykształcenia, zainteresowań i aspiracji korzystających z pośrednictwa. W każdym z tych wymiarów możemy mieć zróżnicowaną informację o wynikach.

W celu wykonania tego kroku należy sporządzić specjalną tabelę, najlepiej wspólnie z głównymi interesariuszami i w partycypacyjny sposób. Jej ogólny wzór zaprezentowany został poniżej.

Tabela 4. Konwersja wyników na oceny ewaluacyjne

Ocena	Wyjaśnienie
Celująca	Przykład najlepszej praktyki nie pozostawiający żadnych wątpliwości; oceny najwyższe pod każdym względem; żadnych słabości
Bardzo dobra	Bardzo dobre wyniki pod wszystkim lub większością względów; ogólnie bardzo dobry, ale nie będący wzorem; brak słabości, które miałyby znaczące konsekwencje
Dobra	Dobre ogólne wyniki; może mieć kilka umiarkowanych słabości, ale żadnej poważnej
Dostateczna	Wyniki zadawalające; kilka poważnych, ale nie krytycznych słabości pod kilkoma względami
Mierna lub niedostateczna	Nie pozostawiające wątpliwości dowody niezadawalającego funkcjonowania; poważne słabości pod prawie wszystkim względami lub w przypadku najważniejszych kryteriów

Źródło: E.J. Davidson, op. cit., s. 136.

Zadanie ewaluatora polega na tym, aby skłonić głównych interesariuszy do zastanowienia się nad tym, jakie wyniki i pod jakim względem będą decydowały o ocenie programu. Być może nietrudno się zgodzić, jakie wyniki powinien mieć program, aby być uznany za doskonały, albo będący całkowitą porażką. Problem jednak w tym, że w rzeczywistości najmniej jest tego rodzaju przypadków. Gdy ewaluacja ma charakter wielokryterialny, dla każdego z nich powinna być wypracowana tabela konwersji. Dodatkową komplikacją są kryteria z kilkoma miarami i skalami, oraz wyniki, które na tych skalach odbiegają od siebie, co utrudnia jednoznaczne przypisanie oceny.

Krótkiego wyjaśnienia wymaga konwersja wyników na oceny, gdy podstawą do tego są wyniki względne, np. ocena celująca, jeżeli jakiś program lub działanie znalazło się w grupie tych, które miały najlepsze wyniki. Polega to na uporządkowaniu całego zbioru ewaluowanych programów według wyników (co wymaga agregowania wyników, jeżeli jest ich więcej dla każdego programu), a następnie wyznaczenie progów w rodzaju 10% najlepszych, kolejne 25%, kolejne 50% itd. Najlepszy w tym przypadku oznacza zawsze – w porównaniu z innymi. Przykładowo, 10% najlepszych wyników z testu może nie osiągnąć 50% maksymalnego wyniku, albo przy wysokim poziomie, 5% najgorszych wyników może przekraczać 50% maksymalnego wyniku. Ocenianie relatywne ma istotne znaczenie w przypadku stosowania metodologii eksperymentalnych (porównuje się wyniki grupy poddanej działaniu z grupą kontrolną) i tam, gdzie stosuje się „benchmarking” (porównanie do najlepszych praktyk, do średniej wyników osiągniętych gdzie indziej).

Przykład tabeli konwersji dla ocen relatywnych pokazany został poniżej. Oceny zostały przypisane do wyników na skali opracowanej na podstawie porównań kosztów programu do jego skuteczności.

Tabela 5. Konwersja wyników na oceny w przypadku ocen relatywnych

Ocena	Wyjaśnienie
Najlepsza praktyka	Alternatywa mająca najlepszy stosunek kosztów do skuteczności
Powyżej średniej	Znacząco lepszy stosunek kosztów do skuteczności niż większość innych alternatyw
Średnia	W przybliżeniu ma taki stosunek kosztów do skuteczności jak większość alternatyw
Poniżej średniej	Znacząco gorszy stosunek kosztów do skuteczności w

	porównaniu z większością alternatyw
Najgorsza praktyka	Bez wątpienia najgorszy stosunek kosztów do skuteczności spośród dostępnych alternatyw

Źródło: E. J. Davidson, op. cit., s. 146.

Agregowanie ocen cząstkowych – metodologia syntezy ewaluacyjnej

Z powyżej przedstawionych informacji powinno jasno wynikać, że programy działań służb społecznych mogą być ewaluowane według więcej niż jednego kryterium z możliwością kilku skal wyników dla każdego. Elementy syntezy muszą więc pojawić się już na tym etapie (sumaryczna ocena kilku wyników, które opisują program pod względem danego kryterium). W przedstawianej tu metodologii robi się nie tylko krok polegający na wyciąganiu wniosków ewaluacyjnych z informacji zebranych o programie i jego wpływie, ale również dokonuje się całkowitej syntezy ewaluacyjnej. Polega ona na wydaniu oceny łącznej programu działania na podstawie ocen cząstkowych, ocen poszczególnych jego wymiarów lub komponentów.

Jeżeli ewaluacji podlega proces wdrażania, wpływu, skuteczności odniesionej do kosztów i przenośności, to mamy co najmniej kilka ocen ewaluacyjnych, a zamawiający może chcieć znać sumaryczną i syntetyczną ocenę łączną. Ewaluator staje więc w obliczu zadania nazywanego agregowaniem ocen lub **syntezą ewaluacyjną**. W jej trakcie odpowiada na główne pytanie ewaluacji: pod jakimi względami program ma się dobrze i gdzie ma braki, jakie są jego ważniejsze i mniej ważne silne i słabe strony, czy jest on najbardziej oszczędnym sposobem osiągnięcia zamierzonego wpływu na dobro klientów, bez wywoływania nadmiernych skutków ubocznych?¹⁴

Synteza ewaluacyjna ma istotne znaczenie, gdy chodzi o wybór najlepszego programu działań służb społecznych spośród kilku konkurentów mających podobne cele. W takich przypadkach konieczne jest tworzenie porównawczych rankingów ewaluacyjnych, z których ma jasno wynikać, która z alternatyw jest lepsza od innych, a która jest gorsza. Trudno przecenić rzetelne rankingi tego rodzaju dla celów praktycznych związanych z wyborem najlepszych projektów do sfinansowania lub najlepszych programów, które mają być kontynuowane czy przedstawiane jako wzory do naśladowania.

Postępujemy więc według następującej procedury złożonej z kilku kroków. Wypracowujemy zestaw kryteriów ewaluacyjnych z uwzględnieniem ich ważności oraz ich miar lub skal (wymiary). Zbieramy i przedstawiamy informację opisową na tych skalach i wyciągamy na tej podstawie wnioski ewaluacyjne. W ten sposób otrzymujemy profil ewaluacyjny danego programu. Może on wyglądać w uproszczeniu tak, jak w kolejnej tabeli.

Tabela 6. Profil ewaluacyjny programu

Kryterium	Wymiary	Oceny				
		M	DST	DB	BDB	C
K1 (krytycznie ważne)	W1K1					
	W2K1					
	W3K1					
K2 (bardzo ważne)	W1K2					
	W2K2					

¹⁴ E. J. Davidson, op. cit., s. 151.

	W3K2					
K3 (umiarkowanie ważne)	W1K3					
	W2K3					
	W3K3					
K4 (pożądane)	W1K4					
	W2K4					
	W3K4					

Skala oceny: M – mierny, DST – dostateczny, DB – dobry, BDB – bardzo dobry. Komórki cieniowane na szaro to hipotetyczny rozkład ocen ewaluacyjnych dla poszczególnych wymiarów. Przyjęto, że na każde kryterium przypadają trzy wymiary, ale ich liczba może być zróżnicowana.

Źródło: E. J. Davidson, op. cit., s. 168 (tabela uproszczona przez autora)

Dla uproszczenia założono, że wymiary przypisane do poszczególnych kryteriów nie różnią się pod względem ważności.

Kolejnym zadaniem jest wyprowadzenie oceny łącznej dla kryteriów na podstawie ocen pod względem ich poszczególnych wymiarów. Wymaga to podobnych do pokazanych wyżej tabel konwersji z regułami w rodzaju, np. celująca ocena łączna wydawana jest, gdy: 1) co najmniej połowa ocen cząstkowych jest równa celującej, 2) żadna z ocen cząstkowych nie jest niższa niż dobra, 3) żadna z ocen cząstkowych nie jest niższa niż dostateczna; bardzo dobra ocena łączna wydawana jest gdy: 1) co najmniej połowa ocen cząstkowych równa jest bardzo dobrej lub jest wyższa, 2) żadna z ocen cząstkowych nie jest niższa niż dobra, 3) żadna z ocen cząstkowych nie jest niższa niż dostateczna; dobra ocena łączna wydawana jest gdy: 1) co najmniej połowa ocen cząstkowych równa jest dobrej lub jest wyższa, 2) mniej niż 35% ocen cząstkowych jest niższa niż dobra, 3) żadna z ocen cząstkowych nie jest niższa niż dostateczna.¹⁵

Kiedy wypracujemy metodę dokonywania syntetycznych ocen na podstawie ocen poszczególnych kryteriów powyższa tabela zostanie uproszczona do dwóch kolumn (kryteria i oceny). Pozostanie do wykonania jeszcze jeden krok polegający na wypracowaniu metody przejścia od ocen pod względem poszczególnych kryteriów do oceny łącznej całego programu. Tym razem trzeba wziąć też pod uwagę różnice pod względem ważności poszczególnych kryteriów. Warunki, jakie muszą spełnić oceny poszczególnych kryteriów mogą być następujące. Żeby program był uznany za najwyższej jakości: 1) żadna z ocen cząstkowych nie może być niższa niż dobra i znacząca większość (np. 80%) powinna być równa celującej; 2) wszystkie krytycznie ważne i ważne kryteria powinny być ocenione jako dobre lub wyżej; 3) spośród umiarkowanie ważnych i pożądanych kryteriów tylko jedno może być ocenione jako dobre, a pozostałe mają mieć oceny bardzo dobre lub celujące. Warunki, jakie muszą być spełnione, aby program był oceniony jako bardzo wysokiej jakości: 1) żadna z ocen cząstkowych nie może być niższa niż dobra i znacząca większość (np. 80%) powinna być bardzo dobra lub celująca; 2) dla kryteriów krytycznie ważnych i bardzo ważnych maksymalnie jedna ocena dobra, wszystkie inne bardzo dobre lub celujące; 3) dla kryteriów umiarkowanie ważnych i pożądanych maksymalnie dwie oceny dostateczne, pozostałe ocenione jako dobre lub wyższe.¹⁶ Na tej podstawie możemy stworzyć tabelę konwersji z tymi zasadami.

Tabela 7. Przykładowe zasady syntezy ewaluacyjnej

Ocena łączna	Warunki, jakie muszą spełniać oceny cząstkowe wydane pod względem
---------------------	--

¹⁵ E. J. Davidson, op. cit., s. 169.

¹⁶ Tamże, s. 170

	kryteriów ewaluacyjnych		
	Wszystkie kryteria	Krytycznie i bardzo ważne kryteria	Umiarkowanie ważne i pożądane kryteria
Najwyższa jakość	Żadna z ocen cząstkowych nie może być niższa niż dobra i znacząca większość (np. 80%) powinna być równa celującej	Ocenione jako dobre lub wyżej	Tylko jedno może być ocenione jako dobre, a pozostałe mają mieć oceny bardzo dobre lub celujące
Bardzo wysoka jakość	Żadna z ocen cząstkowych nie może być niższa niż dobra i znacząca większość (np. 80%) powinna być bardzo dobra lub celująca	Maksymalnie jedna ocena dobra, wszystkie inne bardzo dobre lub celujące	Maksymalnie dwie oceny dostateczne, pozostałe ocenione jako dobre lub wyższe
Wysoka jakość	Znaczna większość (np. 80%) ocen co najmniej dobra lub lepsza	Maksymalnie jedna ocena dostateczna, pozostałe dobre lub wyższe	Nie więcej niż jedna ocena mierna
Umiarkowana jakość	Nie więcej niż dwie oceny mierne	Żadnych ocen miernych	Nie więcej niż jedna ocena mierna
Niska jakość	Jedna lub więcej ocena mierna pod względem kryteriów krytycznie i bardzo ważnych albo więcej niż dwie oceny mierne		

Źródło: E. J. Davidson, op. cit., s. 170

W przypadku ocen relatywnych, czyli przy porównaniach kilku programów, stosowane są dwie metody. Po zebraniu wszystkich informacji opisowych o wynikach programu działań pod względem poszczególnych kryteriów, pierwsza zaczyna się od przypisania liczb wagom i ocenom (np. trzystopniowa skala ważności: 3 najwyższa ważność, 2 średnia ważność, 1 umiarkowana ważność; skali ocen szkolnych przypisuje się liczby od 6 do 1), a metoda agregacji czy syntezy ewaluacyjnej polega na obliczaniu sumy ważonej, czyli sumowaniu ocen pomnożonych przez wagi. Ranking ewaluacyjny powstaje w przypadku wystąpienia różnic takich ważonych sum pomiędzy porównywanymi programami działań. Metoda ta ma sens, gdy 1) liczba kryteriów ewaluacyjnych nie jest duża; 2) uwzględnione zostały progi minimalne, o ile takie ustanowiono dla niektórych kryteriów; 3) określenie ważności kryteriów zostało przeprowadzone w sposób nie budzący wątpliwości np. na podstawie diagnozy potrzeb.¹⁷

Druga metoda ma charakter jakościowy. Skala ważności ma charakter symboliczny i jest wyrażana w kategoriach maksymalnej możliwej wartości przypisanej najlepszemu wynikowi pod względem danego kryterium. Taka skala może mieć trzy stopnie, np. bardzo wartościowy, średnio wartościowy, umiarkowanie wartościowy. Każdemu ze stopni przyporządkujemy trzy różne symbole, np. gwiazdka, trójkąt i plus. Przykładowo, najlepszy wynik pod względem tego, czy program szkoleń był interesujący dla uczestników może być uznany za umiarkowanie wartościowy, najlepszy wynik pod względem dopasowania programu i treści do potrzeb uczestników - jako średnio wartościowy, a najlepszy wynik pod względem kosztów dla organizacji wysyłającej pracowników na szkolenie - jako bardzo wartościowy.

¹⁷ Tamże, s. 172.

Niektóre z kryteriów mogą wymagać ustalenia progu minimalnego, np. ostatnie, w postaci maksymalnego kosztu na uczestnika, powyżej którego cały program uznamy za nie do zaakceptowania, niezależnie od tego, jak jest dobry pod innymi względami. Takiej sytuacji również przyporządkowujemy symbol, np. „x”. Każdy ze stopni dla każdego z kryteriów wymaga opisu, abyśmy wiedzieli w jakich przypadkach i jaki symbol mamy przyznać, np. na skali kosztów ustalamy przedziały do uznania programu za bardzo tani (maksymalny wynik pod względem tego kryterium, przyznaliśmy mu średnią wartościowość, więc i taka będzie jego ważność), umiarkowana cena (umiarkowana wartość), dość drogi, zbyt drogi (przekracza próg akceptowalnych kosztów, czyli przypisujemy symbol nieakceptowalności „x”). Przykład pokazany został w kolejnej tabeli.

Tabela 8. Określanie stopnia wartościowości na przykładzie kosztów

Stopień wartościowości	Symbol	Opis
Średnio wartościowy	Trójkąt	Bardzo tani, koszt poniżej 400 zł za uczestnika
Umiarkowanie wartościowy	Plus	Umiarkowana cena, koszt między 400 a 1500 zł za uczestnika
Bez wartości	[puste miejsce]	Dość drogi, koszt między 1500 a 5000 zł za uczestnika
Nieakceptowalny	Iks	Zbyt drogi, koszt powyżej 5000 zł za uczestnika

Źródło: E. J. Davidson, op. cit., s. 180

Kiedy już wiemy, w jakich sytuacjach i jakie przypisać symbole, przystępujemy do względnej oceny programów, których ranking ewaluacyjny mamy ustalić. Jeżeli pod względem któregoś z kryteriów jeden z programów otrzymał symbol „x”, to eliminujemy go z dalszych porównań. Pomocniczo zliczamy liczbę poszczególnych symboli, jakie otrzymały pozostałe programy. Przykładowy układ symboli pokazany został w kolejnej tabeli.

Tabela 9. Ranking ewaluacyjny

Kryterium ewaluacyjne	Maksymalna wartość	Symbol rzeczywistej wartości	
		Program 1	Program 2
K1	Gwiazdka	Gwiazdka	Trójkąt
K2	Gwiazdka	Trójkąt	Gwiazdka
K3	Trójkąt	Trójkąt	Plus
K4	Trójkąt	Trójkąt	Plus
K5	Plus		Plus
<i>Łącznie</i>		-	-
Gwiazdki		1	1
Trójkąty		2	1
Plusy			3

Źródło: E. J. Davidson, op. cit., s. 185 (tabela uproszczona przez autora)

Przy jakościowej metodzie ustalania rankingów ewaluacyjnych sytuacja nie jest tak klarowna, jak w przypadku obliczania sum ważonych. Możliwych rozkładów poszczególnych symboli może być wiele, większość może pozostawiać wątpliwości, co do tego, jaki powinien być ostateczny ranking ewaluacyjny, który program jest najlepszy. W szczegółowej metodologii syntez ewaluacyjnych powinniśmy szukać odpowiedzi na przynajmniej niektóre z nich.

Zakończenie

Dla czytelnika powinno być już jasne, że pytanie, czy i jaki wpływ miał program działań służb społecznych na sytuację (dobro) zamierzonych klientów ma charakter przedewaluacyjny. Dokładny pomiar tego wpływu jest oczywiście potrzebny, ale daje jedynie podstawę do wydania oceny, co wymaga – jak widzieliśmy wyżej – metodycznego postępowania o innym charakterze. Praktyka ewaluacyjna umożliwia nabycie biegłości w stosowaniu powyższych zaleceń, a w zależności od treści zamówienia czy zlecenia mogą być one mniej lub bardziej użyteczne. Trudno jednak przecenić wiedzę z zakresu metod stricte ewaluacyjnych dla ludzi, którzy chcieliby być profesjonalnymi ewaluatorami, albo przynajmniej pełniej rozumieć to, co oni będą robić, gdy zostanie zamówiona ewaluacja.

Jest wiele modeli oceniania ewaluacyjnego i w związku z tym możemy zapytać, w którym z nich w największym stopniu kładzie się nacisk na metody stricte ewaluacyjne. W odpowiedzi na to pytanie powinniśmy wskazać na Michaela Scrivena, którego idee wywierają bardzo duży wpływ na to, jak się rozumie ewaluację jako taką (nacisk na określenie wartości i jakości programów działań i innych ewaluowanych obiektów) i który jest autorem jednego z modeli ewaluacji oraz listy kontrolnej do niego (Key Evaluation Checklist). W tej książce ze szczegółami została omówiona lista kontrolna opracowana do nieco innego podejścia (patrz rozdział o ewaluacji jako procesie).

Lista haseł do skorowidza przedmiotowego

Analiza ewaluacyjna
Synteza ewaluacyjna
Ważność kryteriów ewaluacyjnych, skala ważności
Próg minimalny oceny pozytywnej
Tabela konwersji wyników na oceny
Agregowanie ocen
Prakryteria
Ranking ewaluacyjny
Ocena relatywna
Suma ważona
Skala oceny

Ewaluacja jako proces

Wstęp

Ewaluowanie jest działaniem złożonym z czynności dokonywanych przez jednego lub większy zespół ludzi (zespół ewaluacyjny). Podzielenie całego procesu ewaluacyjnego na fazy i rozłożenie ich na uporządkowaną sekwencję grup zadań dostarcza nam informacji praktycznych i analitycznych dotyczących procesu ewaluacji. Najbardziej użytecznym wyrazem tego podejścia są listy kontrolne ewaluacji. Zawierają one zestawy zadań do wykonania w odpowiedniej kolejności i niekiedy z podziałem na zadania ewaluatora oraz zadania zamawiającego ewaluację.

Poniżej zostanie przedstawiona jedna z takich list zaproponowana i rozwijana przez Daniela Stufflebeama.¹⁸ Od lat 60. proponuje on całościowe podejście do ewaluacji znane jako model CIPP (skrót od angielskich słów: *context, input, process, product*). Ewaluacja w tym ujęciu ma cztery główne komponenty odpowiadające logice każdego działania: kontekst odnosi się do jego celów, wkład - do planów ich osiągnięcia, proces - do realizacji tych planów, a produkt - do wyników jakie zostały w wyniku działania uzyskane.

Sekwencja zadań ewaluacyjnych w modelu CIPP

Lista kontrolna dla modelu CIPP i dla większości innych podejść do całościowej ewaluacji składa się z kilku grup zadań, które powinny następować po sobie. W tym punkcie krótko omówione zostanie dziesięć takich grup.

Pierwsza dotyczy **uzgodnień kontraktowych**, gdyż ewaluacja z praktycznego punktu widzenia jest przedmiotem umowy między zamawiającym ją a tym, kto zobowiązuje się do jej dokonania (oczywiście możliwa jest też ewaluacja niezamawiana). Jak każda umowa powinna ona określać prawa i obowiązki obu stron. Od jakości umowy zależy również jakość samej ewaluacji, a że nie wszystko da się przewidzieć przy jej podpisywaniu, powinna być przewidziana w niej możliwość późniejszego skorygowania.

Druga grupa działań wiąże się z uwzględnieniem **kontekstu** tego, co ma być ewaluowane i polega, ogólnie rzecz ujmując, na diagnozie potrzeb i problemów, zasobów i możliwości w określonym środowisku. Jeżeli więc mamy poddać ewaluacji działalność służb społecznych w zakresie integracji społecznej danej grupy klientów, to właśnie ich potrzeby, problemy, zasoby i możliwości stanowią istotny dla ewaluacji kontekst. Bez dobrej znajomości kontekstu trudno dokonać oceny użyteczności celów.

Kolejny, trzeci zestaw zadań dotyczy **wkładu**, czyli planowanego sposobu zaspokojenia potrzeb czy rozwiązania problemu klientów. Wybór właściwego sposobu realizacji celów jest równie ważny jak odpowiednie zdefiniowanie celów działania. Jeżeli uda nam się przekonać zainteresowanych działaniem i jego wynikami, że nasz sposób rozwiązania będzie skuteczny i efektywny, stanowić to będzie istotną informację dla ewaluacji całego przedsięwzięcia. Ogólna strategia działania powinna być uzupełniona dokładnym planem zawierającym harmonogram działań oraz dostosowany do niego budżet.

Gdy już zaplanowane działanie zostanie uruchomione skupiamy uwagę na procesie jego realizacji, a inaczej - **implementacji**. Wiąże się to z zadaniami takimi jak: monitorowanie, dokumentowanie i ocenianie aktywności podejmowanych w trakcie danego przedsięwzięcia. Postępowanie według wcześniej opracowanego planu może napotykać na wiele różnorodnych trudności, które mogą być związane z błędami planisty, ale również z nieprzewidzianymi wcześniej okolicznościami, np. opóźnieniami w przekazywaniu pieniędzy przez podmiot finansujący ewaluację. Czynności ewaluacyjne podejmowane w trakcie realizacji działań służb społecznych mogą posłużyć do ich bieżącego korygowania jeżeli odbiegają od planu, lub gdy plan (budżet, harmonogram) okazał się nie przystawać do rzeczywistości w niektórych miejscach.

¹⁸ D. Stufflebeam *CIPP Evaluation Model Checklist*, drugie wydanie, 2007: <http://www.wmich.edu/evalctr/checklists/cippchecklist_mar07.pdf>

W piątej grupie zadań koncentrujemy się na tym, czy i na ile podjęte działania (usługi) dotarły do zamierzonych adresatów (beneficjentów) i miały na nich **wpływ**. Jak dobrze wiadomo, część zakładanych adresatów działań służb społecznych w ogóle z nich nie korzysta, np. osoby ubogie, które nie znają swoich praw czy wstydzące się prosić o pomoc. Ze względu na trudności w rekrutacji klientów może się też zdarzyć, że wśród tych z nich, którzy korzystają znajdują się osoby lub grupy, które mniej lub bardziej odbiegają od zamierzonych beneficjentów, np. mieliśmy pomagać tylko długotrwale bezrobotnym z problemami zdrowotnymi, ale pomogliśmy również kilku osobom, które były bezrobotne, ale nie długotrwale lub nie miały problemów ze zdrowiem.

Po ewaluacji wpływu przychodzi czas na ewaluację jakości i wartości działania ze względu na zakres, w jakim zaspokoiło ono potrzeby zamierzonych adresatów i szersze konsekwencje dla ich bliższych i dalszych kręgów społecznych, np. rodzin, społeczności lokalnych. W tym kroku należy brać pod uwagę zarówno zamierzone, jak i niezamierzone skutki ewaluowanego działania, które powstają w jego trakcie i po zakończeniu. Z tego też względu mowa tu o ewaluacji **skuteczności** działania.

Siódma grupa zadań ewaluacyjnych związana jest z poszukiwaniem odpowiedzi na pytania o ocenę **trwałości** osiągnięć, które przyniosło działanie. Nie tylko zależy nam na tym, żeby było ono ocenione jako sukces, ale też na tym, aby jego osiągnięcia były trwałe. Przykładowo, stwierdzenie, że większość uczestników zaraz po zakończeniu udzielania im pomocy była zintegrowana ze społeczeństwem nie musi oznaczać, że tak samo było po roku i w dłuższym okresie czasu. Celem działań służb społecznych jest trwała poprawa sytuacji i funkcjonowania jednostek i rodzin, a nie nawet spektakularne osiągnięcia, które po chwili przestaną istnieć.

Kolejna grupa zadań dotyczy oceny **przenośności** i ma charakter opcjonalny (w toku czynności ewaluacyjnych nie jest konieczna), ale może być istotna w pewnych kontekstach. Jeżeli działanie ocenione zostało jako duży i trwały sukces, z perspektywy dbałości o interes publiczny ważne jest, aby sprawdzić, czy gdzie indziej nie dałoby się tego sukcesu powtórzyć. Gdy już dojdzie do naśladowania pewnej dobrej praktyki w innym miejscu i czasie, ewaluacji podlega również to, czy osiągnięte zostały równie dobre wyniki.

Działania ewaluacyjne również podlegają ewaluacji, którą nazywa się **metaewaluacją**. Zadania z tym związane stanowią przedostatnią grupę na naszej liście kontrolnej. Jakość oceniania ewaluacyjnego może być różna, a od niej zależy wiarygodność ocen uzyskanych w ten sposób. Żeby przekonać się o tym, z jakiej jakości ewaluacją mamy do czynienia, a pośrednio na ile wiarygodne są jej wyniki, zamawiamy meta ewaluację, najlepiej u ewaluatora zewnętrznego wobec zespołu ewaluującego nasz program.

Sporządzenie **raportu końcowego** z przeprowadzonej ewaluacji to też złożony zestaw czynności i zajmuje on na naszej liście kontrolnej ostatnie miejsce. Zamawianym produktem ewaluacji jest rzetelna i obiektywna ocena tego, co było jej przedmiotem. Raport końcowy z przeprowadzonej ewaluacji powinien ją zawierać, ale spełniać też winien wiele innych kryteriów, żeby zadowolić zamawiającego i nie tylko jego.

Powyższe etapy w prowadzeniu czynności ewaluacyjnych przedstawione zostaną w skrócie w poniższej tabeli.

Tabela 1. Ewaluacja jako sekwencja dziesięciu grup zadań

Faza ewaluacji	Syntetyczna charakterystyka
Uzgodnienia kontraktowe	Ewaluacja powinna być oparta na umowie z klientem z możliwością jej korygowania w trakcie ewaluacji
Ewaluacja kontekstu	Oszacowanie i analiza potrzeb, zasobów, problemów w określonym środowisku
Ewaluacja wkładu	Porównywanie konkurencyjnych strategii, planów działań i budżetów wybranego podejścia do programu

Ewaluacja procesu	Monitorowanie, dokumentowanie i ocenianie działań podejmowanych w ramach programu
Ewaluacja wpływu	Ocena na ile program dotarł do zamierzonych beneficjentów
Ewaluacja skuteczności	Dokumentacja i ocena jakości i znaczenia wyników programu
Ewaluacja trwałości	Ocena czy wyniki programu są dostatecznie zinstytucjonalizowane i mają zapewnione podstawy do trwania
Ewaluacja przenośności (opcjonalna)	Ocena na ile program może być przyjęty i zastosowany gdzie indziej
Metaewaluacja	Ocena, czy ewaluacja została przeprowadzona zgodnie ze standardami i/lub zasadami dobrej ewaluacji
Raport końcowy	Przygotowanie raportu, który zawiera m.in. Wyniki ewaluacji

Źródło: na podstawie D. Stufflebeam *CIPP Evaluation Model Checklist*

Poniżej każdy z tych etapów zostanie przedstawiony z większą liczbą szczegółów i problemów, które mogą się pojawić w praktyce.

Uzgodnienia kontraktowe

Lista kontrolna CIPP rozpoczyna się od zadań związanych z przygotowaniem umowy, co jednak pomija pewien element w działalności ewaluacyjnej. Publiczne, jak i niepubliczne podmioty często zamawiają ewaluację poprzez konkursy ofertowe lub przetargi publiczne, stąd pojawiają się zadania ewaluatora związane z przygotowaniem oferty, a ze strony zamawiających – zadania polegające na przygotowaniu i rozstrzygnięciu konkursu. Już na tym etapie pojawiają się czynności ewaluacyjne po stronie zamawiającego, gdyż wybór najlepszej oferty zakłada, że dokonano oceny wszystkich ofert i uporządkowano je od najlepszej do najgorszej. W przypadku skomplikowanych zadań ewaluacyjnych taka ocena nie jest łatwa i wymagać może zaangażowania zewnętrznych ewaluatorów (nazywanych w tej roli i w kontekście konkursów o środki UE asesorami). Po rozstrzygnięciu konkursu dochodzi o właściwego etapu uzgodnienia treści umowy.

Zawarcie umowy pisemnej jest standardowym wymogiem w kontaktach handlowych, w szczególności zaś gdy przedmiot umowy jest kosztowny, co oznacza zwiększone ryzyko strat dla stron zawierających umowę. W przypadku ewaluacji złożonych działań służb społecznych (projekty, programy społeczne) konieczne jest zaangażowanie różnego rodzaju zasobów i kapitału, aby zapewnić odpowiednią jakość procesu ewaluacji i wiarygodność jej wyników. Zadania związane z zawieraniem umowy na ewaluację z podziałem na te dla ewaluatora i jego klienta (zamawiającego) wyszczególnia poniższa tabela.

Tabela 2. Uzgodnienia kontraktowe

1. UZGODNIENIA KONTRAKTOWE	
Działania ewaluatora	Działania klienta
Wypracuj jasne zrozumienie pracy ewaluacyjnej, która ma być wykonana	Wyjaśnij z ewaluatorem co ma być przedmiotem ewaluacji, w jakim celu się ją przeprowadza, według jakich kryteriów i dla jakiej publiczności
Zapewnij , żeby umowa gwarantowała dostęp do wymaganych informacji	Wyjaśnij z ewaluatorem jaka informacja jest zasadnicza dla ewaluacji i jak ułatwisz mu do niej dostęp
Uświadom klientowi, jakie jakościowe i ilościowe analizy będą potrzebne dla dokonania ewaluacji	Uzgodnij z ewaluatorem, jakie analizy będą najbardziej odpowiadać pytaniom klientów
Wyjaśnij istotę, ogólną zawartość, terminarz przedstawiania raportów z ewaluacji ex ante, interim i ostatecznej	Zapewnij, aby treść i terminarz raportów były zgodne z potrzebami i fazami programu

Wyjaśnij kanały komunikacji i pomocy ze strony klienta/interesariusza	Zapewnij, aby plan ewaluacji był zgodny ze statutem organizacji
Zabezpiecz w umowie odpowiednie środki i daty wpłat	Zapewnij, aby budżet ewaluacji był jasny i gwarantujący jej powodzenie

Źródło: D. Stufflebeam, op. cit., s. 3.

Można sobie wyobrazić różne scenariusze dotyczące zawarcia umowy o ewaluację. Zamawiający nie musi być bezpośrednio związany z działaniami służb społecznych, może być tylko za nie odpowiedzialny przed społeczeństwem lub innymi interesariuszami, np. organ władzy samorządowej zamawia ewaluację jednego z programów prowadzonych przez służby społeczne w ramach Ośrodka Pomocy Społecznej lub Powiatowego Centrum Pomocy Rodzinie. W takiej sytuacji kluczowe są kwestie zapewnienia dostępu do informacji, które są podstawą do prowadzenia analiz i wyciągnięcia wniosków ewaluacyjnych.

W zamówieniu mogą być już określone pytania, na które ma odpowiedzieć ewaluator lub zespół ewaluacyjny. Wówczas częścią negocjacji w sprawie zawarcia umowy może być ich doprecyzowanie lub skorygowanie, aby ewaluator nie musiał odpowiadać na pytania, które nie mają charakteru ewaluacyjnego, albo żeby pytania zostały tak sformułowane, aby można było na nie w ogóle odpowiedzieć.

W początkowej fazie procesu ewaluacji ewaluatorzy powinni zadać kilka ważnych pytań różnym grupom interesariuszy ewaluowanego programu działań (w tym jego klientom i ich rodzinom). Zestaw takich pytań został przedstawiony w ramce.

Ramka 1.

- Jakie są ich ogólne spostrzeżenia dotyczące obszaru danego programu czy praktyki?
- Jakie według nich są cele danego programu czy praktyki?
- W jakim zakresie rozumieją oni obszar programu czy praktyki, jak miały być one realizowane, jak są wdrażane w rzeczywistości?
- Jakie problemy czy kwestie są najważniejsze do poruszenia w procesie ewaluacji?
- Dlaczego właśnie te, a nie inne, kwestie są dla nich ważne?
- Jakie nadzieje każdy z nich wiąże z przeprowadzeniem ewaluacji i jej wynikami?
- Jak chcieliby wykorzystać wyniki takiej ewaluacji?
- Jakie korzyści może przynieść ewaluacja klientom programu?
- W jaki sposób ewaluacja przyczyni się do dobra klientów programu?
- Jak należy zastosować wyniki ewaluacji, aby w większym stopniu zwiększyła się jakość życia klientów programu?

Źródło: James. R. Dudley *Social Work Evaluation: Enhancing What We Do*, Lyceum Books, Chicago, 2009, s. 87, 88.

Kontakt i komunikacja z interesariuszami przed przeprowadzeniem ewaluacji, w jej trakcie i po są ważne z wielu względów. Program działań służb społecznych ma służyć klientom, a ewaluacja ma służyć programowi, rola ewaluatorów, jak i personelu programu jest służebna, co powinno się wyrażać również w sposobie traktowania i komunikacji z zamierzonymi beneficjentami programu.

Ewaluacja kontekstu

Każde działanie zachodzi w pewnym kontekście, który stanowi też dla niego źródło uzasadnień i oceny. Jednym z najważniejszych kontekstów dla działań służb społecznych są potrzeby ich docelowych klientów (adresatów, beneficjentów), a ogólniej ich dobro. Dlatego też w wielu ujęciach podkreśla się znaczenie jasnego zidentyfikowania ich oraz diagnozy ich potrzeb dla właściwego zaprojektowania działania i jego sukcesu. Odgrywa to zasadniczą rolę dla oceny użyteczności projektowanych i podejmowanych działań, czyli tego, czy i na ile przyczynić się one mogły i przyczyniły się do dobra klientów.

Szczegółowy wykaz zadań w tej fazie ewaluacji pokazany został w tabeli.

Tabela 3. Ewaluacja kontekstu

2. EWALUACJA KONTEKSTU	
Działania ewaluatora	Działania klienta/interesariusza
Zbierz i oceń informacje dotyczące potrzeb i zasobów zamierzonych beneficjentów programu	Wykorzystaj wyniki ewaluacji kontekstu do wyboru i/lub określenia zamierzonych beneficjentów programu
Przeprowadź wywiady z liderami programu i przedyskutuj z nimi ich punkt widzenia na potrzeby beneficjentów i problemy, które ma rozwiązywać program	Wykorzystaj wyniki ewaluacji kontekstu do przeglądu celów programu, aby zapewnić, że odpowiadają one stwierdzonym potrzebom
Przeprowadź wywiady z innymi zainteresowanymi programem by zdobyć dodatkowe informacje o potrzebach i zasobach zamierzonych beneficjentów i potencjalnych problemach dla programu	Zastosuj wyniki ewaluacji kontekstu, aby program wykorzystywał zidentyfikowane zasoby społeczności
Oceń cele programu w świetle potrzeb i zasobów beneficjentów programu	Wykorzystaj wyniki ewaluacji kontekstu pomocniczo do określenia skuteczności i znaczenia programu dla zaspokajania zidentyfikowanych potrzeb beneficjentów
Zaangażuj specjalistę od gromadzenia danych o środowisku programu (np. potrzeby i zasoby na obszarze działania programu)	
Wymagaj aby personel programu regularnie udostępniał zespołowi ewaluacyjnemu informacje o beneficjentach i środowisku	
Przedstawiaj interesariuszom w odpowiednim czasie wstępne raporty dotyczące związanych z programem potrzeb, zasobów, problemów wraz z oceną celów i priorytetów programu	
Periodycznie dyskutuj o wynikach ewaluacji kontekstu z klientem i zamierzonymi odbiorcami na specjalnych sesjach (<i>feedback sessions</i>)	
Zakończ przygotowywanie raportu końcowego z ewaluacji kontekstu i przedstaw go klientowi i uzgodnionym interesariuszom	

Źródło: D. Stufflebeam, op. cit., s. 4.

Istotne jest, że ewaluacja kontekstu może być pomocna już w fazie projektowania i przygotowywania programu przy weryfikacji celów oraz wyborze i rekrutacji zamierzonych adresatów. Niepomijanie problemów, barier oraz zasobów i możliwości w społeczności najbliższej

klientom też trudno przecenić dla powodzenia programów społecznych, a w tym działań służb społecznych.

Już na tym etapie w ewaluacji wykorzystuje się zróżnicowane metody zbierania i analizowania danych. Diagnozowanie potrzeb (ang. *needs assessment*) stanowi odrębny obszar refleksji metodologicznej rozwijany również w kontekście działania służb społecznych. Mamy tu istotne kwestie różnej natury od najprostszych, czyli poziomu diagnozy – czy jest nim jednostka lub mała grupa rodzinna, czy też cała społeczność, a jeżeli ta ostatnia, to w jakiej skali, np. domu wielorodzinnego, osiedla, dzielnicy, wsi, miasta, gminy itd. W pierwszym przypadku większe znaczenie mają podejścia znane z psychologii i psychoterapii, w drugim ważne jest ujęcie statystyczne, np. klasyfikacja i liczenie przypadków, aby zaplanować odpowiednią ilość i jakość usług służb społecznych.

W ramce przedstawiono przykładowy zestaw pytań, na który powinna odpowiedzieć diagnoza potrzeb w obszarze wychowywania dzieci w ubogim środowisku.

Ramka 2. Przykład pytań do diagnozy potrzeb

- W jakim zakresie i jak rodzice (biologiczni, zastępczy) są zaangażowani w wychowywanie swoich dzieci?
- W jakim zakresie i czy rodziców stać na zaspokajanie potrzeb ich dzieci?
- Jakie problemy mają matki i ojcowie w wypełnianiu ról rodzicielskich?
- Jakie są trendy w zakresie zjawisk takich, jak nieletnie macierzyństwo, dzieci pozamażeńskie, samotne rodzicielstwo, rozwody? Czy SA to trendy wzrostowe, czy spadkowe?
- W jakim zakresie w społeczności są usługi wspomagające rodziców w wypełnianiu ról rodzicielskich?
- Jakie są cechy demograficzne (np. wiek) ojców i matek, którzy korzystają z tych usług? Jakie są cechy demograficzne tych rodziców, którzy z nich nie korzystają?
- Jakie są bariery w dostępie i w korzystaniu z usług?
- Jakie problemy wydają się oczywiste w przypadku usług?
- Jakie nowe inicjatywy polityczne mogłyby pomóc tym rodzicom?

Źródło: J. R. Dudley, op. cit. s. 110.

Cele i funkcje diagnozy potrzeb są zróżnicowane, np.: ocena stopnia w jakim ewaluowany program czy praktyka są potrzebne (dotyczy działań dopiero planowanych); określenie tego, na ile istniejące programy odpowiadają zdiagnozowanym potrzebom; określenie tego, na ile istniejące programy odpowiadają nieformalnym zasobom (np. dalszy krąg rodzinny, organizacje społeczne, instytucje wyznaniowe, instytucje publiczne o charakterze niepomocowym); zidentyfikowanie tego, co zniechęca lub odstrasza klientów od korzystania z usług; wspomaganie procesu określania celów programu, które mogą za bardzo odpowiadać perspektywie usługodawców w porównaniu z potrzebami klientów.¹⁹

Z perspektywy zadań ewaluatora może okazać się konieczne zatrudnienie dodatkowych specjalistów w zakresie badań terenowych, gdyż każdy program działań służb społecznych jest realizowany w konkretnych lokalizacjach i to między innymi one decydują o jego kontekście.

Ewaluacja wkładu

Zdiagnozowanie potrzeb i problemów oraz możliwości i zasobów zamierzonych beneficjentów oraz ich środowiska społecznego powinno dać nam informację o tym, czy potrzebna jest zmiana oraz pomoc w zdefiniowaniu celów działania. W większości sytuacji nie jest jednak

¹⁹ J. R. Dudley, op. cit. s. 111-113.

oczywiste, w jaki sposób należy je zrealizować, a jeżeli już to wiadomo, możemy mieć do czynienia z konkurencyjnymi rozwiązaniami i problemem wyboru spośród nich najlepszego.

Sposób realizacji celów dotyczy tego, co zamierzamy zrobić, żeby je osiągnąć. Może być to opisane ogólnie, ale dla celów praktycznych wymagane są szczegóły, które zawierają przede wszystkim harmonogramy (co i kiedy ma być zrobione) oraz budżet (kiedy i jaka suma będzie potrzebna na sfinansowanie zadania określonego w harmonogramie).

Właśnie tej fazy działania dotyczy ewaluacja wkładu, zadania ewaluatora i zamawiającego przedstawia kolejna tabela.

Tabela 4. Ewaluacja wkładu

3. EWALUACJA WKŁADU	
Działania ewaluatora	Działania klienta/interesariusza
Zidentyfikuj i zbadaj istniejące programy, które mogą służyć jako modele dla ewaluowanego programu	Wykorzystaj wyniki ewaluacji wkładu w celu opracowania takiej strategii programu, która będzie ekonomicznie, politycznie, społecznie, naukowo i technologicznie uzasadniona
Oceń proponowaną strategię programu pod względem stwierdzonych potrzeb i wykonalności	Wykorzystaj wyniki ewaluacji wkładu, aby zapewnić, że strategia programu zaspokoi oszacowane potrzeby docelowych beneficjentów
Oceń budżet programu pod względem jego wystarczalności na potrzebne działania	Wykorzystaj wyniki ewaluacji wkładu, aby uzasadnić wystąpienia o finansowanie zaplanowanych działań
Oceń strategię programu w świetle badań i odpowiedniej literatury	Wykorzystaj wyniki ewaluacji wkładu, aby zapoznać personel programu z kwestiami kluczowymi dla powodzenia realizacji programu
Oceń wartość strategii na tle alternatywnych strategii w podobnych programach	Wykorzystaj wyniki ewaluacji wkładu przedstawiając uzasadnienie dla wybranej strategii programu i planu operacyjnego
Oceń plan działań programu i ich harmonogram pod względem wystarczalności, wykonalności i poparcia politycznego	
Wstępną wersję raportu z ewaluacji wkładu prześlij klientowi i uzgodnionym adresatom	
Przedyskutuj wyniki ewaluacji wkładu na specjalnym spotkaniu (warsztacie)	
Zakończ raport z ewaluacji wkładu i przedstaw go klientowi oraz uzgodnionym adresatom	

Źródło: D. Stufflebeam, op. cit., s. 5.

Odnalezienie programów, które miały podobne cele do naszego jest naturalnym źródłem informacji o tym, jak je osiągać. Mając dostęp do ich harmonogramów i budżetów możemy je naśladować dostosowując do konkretnego kontekstu. Niezależnie opracowane plany działania mogą być też oceniane na tle planów podobnych przedsięwzięć, które zostały z sukcesem zrealizowane.

Wyniki ewaluacji kontekstu już w tej fazie posłużą ewaluatorowi do oceny strategii realizacyjnej oraz planów, np. czy uwzględni ona zidentyfikowane potrzeby i inne ważne dla powodzenia programu elementy. Dobre plany powinny być realistyczne (wykonalne) przy danych zasobach, ale to nie wyczerpuje ich pożądanych cech. Pozostałe to m.in.: 1) celowość (plan zawiera właściwe środki do realizacji założonego celu); 3) zgodność wewnętrzna (niesprzeczność, koordynacja, współdziałanie elementów planu); 4) operatywność (zrozumiałość, komunikatywność,

łatwość przejścia do realizacji); 5) racjonalność (oparcie na rzetelnej wiedzy), 6) elastyczność (dopuszczanie i wskazywanie ewentualnych zmian podczas realizacji); 7) optymalność szczegółowa (równowaga między szczegółowością i ogólnością, dopuszcza niedookreślenie tam, gdzie nie ma dostatecznej wiedzy o warunkach działania); 9) określoność czasowa (np. określenie harmonogramu, terminu wykonania); 10) kompletność (obejmuje wszystkie czynności realizacji).²⁰

Należy pamiętać, że dobry plan nie jest wystarczającym warunkiem osiągnięcia sukcesów w działaniu służb społecznych i nie tylko ich, gdyż zawsze odnosi się on do założonych celów. Jeżeli nie odpowiadają one potrzebom lub mamy z nimi inne problemy (nadmiar celów, sprzeczności między celami), nawet doskonały plan nic nie pomoże.

Cele działań służb społecznych są formułowane w różny sposób, przy czym mogą pojawić się problemy takie, jak trudność w pomiarze ich realizacji, nadmierna ogólność, przeteoretyzowanie, zbyt duża liczba i sprzeczności między nimi. Można wyróżnić wiele rozmaitych klasyfikacji celów, np. strategiczne i operacyjne, ogólne i szczegółowe, ostateczne i instrumentalne. W odniesieniu do problemów społecznych jeden z zasadniczych podziałów dotyczy odróżnienia zapobiegania od łagodzenia, co też łatwo przełożyć na cele, w których kładzie się nacisk na zapobieganie i takie, które mają kierować działaniem na zmniejszanie dolegliwości skutków. Jedną z klasyfikacji przedstawia kolejna tabela.

Tabela 5. Cztery typy celów.

<i>Czy to mamy?</i>	<i>Czy chcemy to mieć?</i>	
	<i>Tak</i>	<i>Nie</i>
<i>Tak</i>	Cel: zachować (np. zachować jak najdłużej sprawność)	Cel: usunąć (np. zlikwidować bezrobocie)
<i>Nie</i>	Cel: osiągnąć (np. doprowadzić do tego, że dziecko wróci do rodziców biologicznych)	Cel: unikać (np. unikać zachowań ryzykownych)

Źródło: Michael Q. Patton *Utilization-Focused Evaluation*, czwarte wydanie, SAGE Publications, Thousand Oaks, 2008, s. 236, (z uzupełnieniami autora).

Ewaluacja procesu

Kluczowe dla powodzenia zaplanowanych działań służb społecznych jest to, w jaki sposób będą one realizowane. Nietrudno sobie wyobrazić sytuację, w której nawet bardzo szlachetne i odpowiadające potrzebom cele i dobry plan ich osiągnięcia nie prowadzi do sukcesu ze względu na różnorakie przeszkody, np. brak poparcia u lokalnych polityków, niekompetencja realizatorów, niechęć klientów do uczestnictwa, opóźnienia w finansowaniu, zawodzący kooperanci, nieoczekiwane problemy z dostępem do kluczowych informacji itp. Tych problemów jest tak wiele, że czasem zadaje się pytanie, jak to w ogóle możliwe, że jakiegokolwiek programy społeczne są realizowane i mają sukcesy.

Wynika stąd co najmniej tyle, że bardzo istotne jest monitorowanie postępów i przebiegu działań służb społecznych - czy są zgodne z tym, jak je zaplanowano pod względem treści, harmonogramu i kosztów, nie mówiąc już o bardziej podstawowych kwestiach typu zgodności z prawem, zasadami rachunkowości i innymi standardami. Ewaluacja procesu ma też istotne znaczenie, gdy jej celem jest bieżące rozwiązywanie problemów realizacyjnych i ogólniej – zarządzanie programem.

²⁰ Tadeusz Kotarbiński *Sprawność i błąd*, Państwowe Zakłady Wydawnictw Szkolnych, Warszawa 1970, s. 157-164.

Tabela 6. Ewaluacja procesu

4. EWALUACJA PROCESU	
Działania ewaluatora	Działania klienta/interesariusza
Zaangażuj członka zespołu ewaluacyjnego do monitorowania, obserwowania i okresowego raportowania postępów w realizacji programu	Użyj ewaluacji procesu by koordynować i wspierać działania personelu programu
	Zastosuj wyniki ewaluacji procesu, aby poprawić konstrukcję programu
We współpracy z personelem programu prowadź rejestr zdarzeń, problemów, kosztów i alokacji związanych z programem	Wykorzystaj wyniki ewaluacji procesu w celach dokumentacji realizacji programu
Przeprowadzaj okresowe wywiady z beneficjentami, kierownictwem programu i personelem, aby uzyskać ich ocenę postępów w realizacji programu	Wykorzystaj wyniki ewaluacji procesu w celach dokumentacji kosztów programu
Na bieżąco aktualizuj i prowadź dokumentację realizacji programu	Na podstawie wyników ewaluacji procesu przedstawiaj raporty realizacyjne sponsorom programu, zarządowi organizacji, członkom społeczności, projektantom programu itp.
Przygotuj wstępne wersje raportu z ewaluacji procesu i przedstawiaj je klientowi i uzgodnionym adresatom	
Przedstaw i przedyskutuj wyniki ewaluacji procesu na specjalnym spotkaniu (warsztacie)	
Zakończ raport z ewaluacji procesu i przedstaw go klientowi oraz uzgodnionym adresatom	

Źródło: D. Stufflebeam, op. cit., s. 6.

Ewaluacja procesu wymaga szczegółowego monitorowania i dokumentowania tego, co się działo i co robiono w trakcie realizacji programu, a także jego kosztów. Przewodnikami dla tego rodzaju działalności ewaluacyjnej są strategie, plany, harmonogramy i inne dokumenty programu, z których wynika, co ma być monitorowane. Typy ewaluacji procesu i zasadnicze dla nich pytania przedstawia kolejna tabela.

Tabela 7. Typy ewaluacji procesu

Typ ewaluacji procesu	Ogólne pytanie ewaluacyjne
Powiązania pomiędzy problemami klienta a podejściem reprezentowanym przez program	Czy podejście reprezentowane przez program może odnieść sukces w rozwiązywaniu problemów klienta?
Realizacja programu zgodnie z zamierzeniami	Czy program jest realizowany, tak jak to zamierzano i projektowano?
Jakość programu	Czy poziom jakości programu jest akceptowany przez interesariuszy?
Dostępność programu	Czy program jest dostępny dla wszystkich jego zamierzonych klientów?
Zadowolenie klienta	Czy klienci są zadowoleni z bieżącego działania programu?
Badania personelu	Jakie poglądy, postawy i praktyki ma personel programu i jego wolontariusze w związku z jego realizacją?

Źródło: J. R. Dudley, op. cit., s. 161.

Pierwsze z pytań przedstawionych w tabeli mocno wiąże się z tym, co nazywane jest logiką, matrycą logiczną, modelem, a jeszcze inaczej – teorią programu. Nie zawsze jest ona bezpośrednio wyrażona, więc jednym z zadań ewaluatora może być jej odkrycie czy odtworzenie. Teoria programu wyjaśnia, w jaki sposób i dlaczego działania służb społecznych podjęte w ramach programu rozwiązują problemy jego zamierzonych klientów. Inaczej mówiąc, które przyczyny tych problemów i jak zmniejszają bądź usuwają te działania.

Pytania zadawane w trakcie ewaluacji procesu przedstawia kolejna ramka.

Ramka 3. Przykładowe pytania w ewaluacji procesu

- W jaki sposób powstał ewaluowany program? Jakie były źródła oporu wobec niego lub poparcia dla niego? W jaki sposób uwzględniono tę informację?
- Jaka była teoria programu i czy została zastosowana tak, jak planowano?
- Co zrobiono, aby zwiększyć poparcie programu w realizującej go instytucji i w społeczności?
- Jak reklamowano program? Jakie grupy potencjalnych klientów były rekrutowane?
- Czy program odniósł sukces w udzielaniu usług potencjalnym klientom, dla których był przeznaczony?
- Jakie dodatkowe usługi dla klientów mogły zwiększyć prawdopodobieństwo, że program odniesie sukces?
- Dlaczego niektórzy klienci odeszli z programu przedwcześnie?
- Kto realizował program? Czy wystąpiły opóźnienia w realizacji programu, a jeżeli tak, dlaczego do nich doszło?
- Jakie problemy realizacyjne mogły mieć wpływ na ewentualne powodzenie programu?
- Czy program został odpowiednio zaplanowany?
- Czy kadra zarządzająca wykonywała wszystkie niezbędne zadania zarządcze na czas?
- Czy program był odpowiednio monitorowany i czy wprowadzano zmiany w odpowiedzi na zgromadzone dane?
- Jakie nieplanowane i niezgodne z teorią programu zmiany wystąpiły?
- Czy program zrekrutował i utrzymał profesjonalny i pomocniczy personel? Czy płynność personelu była problemem?
- Jakie normy grupowe, tradycje i wzory interakcji ustanowiono między personelem oraz między nim a klientami programu?
- Jakie problemy były z personelem? Czy klimat organizacyjny programu był odpowiedni?
- Jak program współpracował z innymi programami w zakresie wymiany informacji, wydawania i przyjmowania skierowań itp.?
- Czy budżet programu był odpowiedni do realizacji jego celów? Jakie zmiany w finansowaniu w trakcie realizacji programu mogły mieć wpływ na jego potencjalną skuteczność?
- Jakie wskaźniki powinny być zastosowane w celu przeprowadzenia uczciwej ewaluacji programu?
- Co powinno zostać zrobione inaczej, jeżeli podobny program miałby być uruchomiony ponownie?

Źródło: Robert W. Weinbach *Evaluating Social Work Services and Programs*, Pearson, Allyn & Bacon, Boston, 2005, s. 168.

Ogólne pytanie zadawane podczas ewaluacji procesu może być też sformułowane następująco: co się wydarzyło w czasie realizacji programu i dlaczego miało to wpływ na jego potencjał w osiągnięciu zamierzonych skutków?

Ewaluacja wpływu

Zasadniczo każdy program społeczny skierowany jest do pewnej kategorii klientów. Nic jednak nie gwarantuje tego, że tylko oni z niego skorzystają. W skrajnych przypadkach może się okazać, że tylko mniejszość obsłużonych przez program klientów była jego zamierzonymi adresatami. Ewaluowanie programu bez określenia tych informacji nie ma sensu, np. to, że program dla trudniejszych przypadków ukończyło z sukcesem 80% jego klientów ma dopiero wtedy znaczenie, gdy stwierdzimy, że wszyscy oni byli rzeczywiście jego zamierzonymi adresatami. Po ewaluacji wpływu może się okazać, że z tych 80% tylko część była właściwymi beneficjentami, a do tego, że wśród 20% porażek większość stanowiły właśnie przypadki trudniejsze, co może zasadniczo zmienić pozytywny obraz ewaluowanego programu.

Sytuacja idealna ma miejsce wtedy, gdy wszyscy obsłużeni przez program byli jego zamierzonymi adresatami i nikt poza nimi z niego nie korzystał. W praktyce taki ideał trudno osiągnąć w szczególności, gdy beneficjentami mają być osoby z różnych powodów niechętnie lub obawiające się korzystania z oferowanych im usług, np. niepełnosprawni w obawie przed utratą renty z tytułu niezdolności do pracy mniej chętnie mogą uczestniczyć w programach prozatrudnieniowych; osoby z mniejszości etnicznej nieufne wobec usługodawców zatrudnionych przez państwo zdominowane przez grupę większościową.

Tabela 8. Ewaluacja wpływu

5. EWALUACJA WPLYWU	
Działania ewaluatora	Działania klienta/interesariusza
Zaangażuj personel programu i konsultantów i/lub członków zespołu ewaluacyjnego by utrzymywać rejestr obsłużonych osób i grup, ich potrzeb i usług jakie otrzymują z programu	Zastosuj wyniki ewaluacji wpływu, aby zapewnić, że program dociera do zamierzonych beneficjentów
	Zastosuj wyniki ewaluacji wpływu, aby określić z programu korzystali niewłaściwi odbiorcy
Oszacuj i oceń zakres w jakim obsługiwane jednostki i grupy są zgodne z zamierzonymi beneficjentami	Na podstawie wyników ewaluacji wpływu oszacuj zakres w jakim program służy zamierzonym beneficjentom
Okresowo przeprowadzaj wywiady z interesariuszami takimi, jak przywódcy społeczności, pracodawcy, personel szkolnych i programów socjalnych, kler, policja, sędziowie i właściciele domów, aby poznać ich punkt widzenia na wpływ programu na społeczność	Na podstawie wyników ewaluacji wpływu określ czy program zaspokoił lub zaspokaja ważne potrzeby społeczności
	Zastosuj wyniki ewaluacji wpływu w celach rozliczenia programu z tego, czy odniósł sukces w docieraniu do zamierzonych beneficjentów
Zamieszczaj zebrane informacje i oceny ewaluacyjne w profilu programu	
Określ zakres, w którym program objął właściwą grupę beneficjentów	
Oszacuj zakres, w którym program dostarczył usług grupie niebędącej jego zamierzonymi beneficjentami	
Przygotuj wstępne wersje raportu z ewaluacji wpływu i przedstawiaj je klientowi i uzgodnionym adresatom	
Przedstaw i przedyskutuj wyniki ewaluacji wpływu na specjalnej sesji	

Zakończ raport z ewaluacji wpływu i przedstaw go klientowi oraz uzgodnionym adresatom	
---	--

Źródło: D. Stufflebeam, op. cit., s. 7.

Informacje zawarte w raporcie z ewaluacji wpływu powinny być wykorzystane przez zarządzających nim do poprawy w zakresie rekrutacji i utrzymania zamierzonych klientów w programie. Może to dotyczyć między innymi zmian w sposobie rozpowszechniania informacji o programie (reklama), w kryteriach i procedurach przyjmowania, w systemie skierowań czy w układzie bodźców, które zachęcają i zniechęcają zamierzonych adresatów do korzystania z oferowanych usług.

Ewaluacja skuteczności

Jak już wiemy, skuteczność jest jednym z głównych kryteriów ewaluacyjnych programów społecznych. Im lepiej sformułowane cele programu (jego zamierzone skutki dla klientów), tym łatwiejsza będzie też ocena na ile udało się je osiągnąć. Zwraca się uwagę na częste mylenie usług udzielanych przez personel programu, czyli służby społeczne, z jego skutkami dla klientów. W ocenie skuteczności należy jednak odróżnić informację o usłudze, np. jakie usługi szkoleniowe oferowano, ile osób z nich skorzystało, ile to trwało, ile kosztowało itd., od informacji o tym, jaki miały one wpływ na klienta, np. czy usługa szkoleniowa doprowadziła do zdobycia nowych umiejętności, które z kolei sprawiły, że uzyskująca je osoba bezrobotna znalazła szybko pracę.

Teoria programu powinna wyjaśniać, w jaki sposób usługa udzielana przez służby społeczne spowoduje zamierzony wpływ, czyli osiągnięcie celów określonych w odniesieniu do dobra czy korzyści klientów w krótkim, średnim i dłuższym okresie.

Tabela 9.

6. EWALUACJA SKUTECZNOŚCI	
Działania ewaluatora	Działania klienta/interesariusza
Przeprowadź wywiady z kluczowymi interesariuszami takimi, jak przywódcy społeczności, beneficjenci, kierownictwo i personel programu i inne zainteresowane strony, aby ustalić ich ocenę pozytywnych i negatywnych wyników programu	Zastosuj wyniki ewaluacji skuteczności, aby zmierzyć pozytywne i negatywne skutki programu na beneficjentów
	Zastosuj wyniki ewaluacji skuteczności, aby zmierzyć pozytywne i negatywne skutki programu na społeczność i otoczenie
Jeżeli jest to wykonalne przeprowadź kilka pogłębionych studiów przypadków z wybranymi beneficjentami	Zastosuj wyniki ewaluacji skuteczności, aby wyodrębnić i ocenić ważne skutki uboczne programu
Zaangażuj członków zespołu ewaluacyjnego i personel programu, aby dostarczyć dokumentację potrzebną do określenia stopnia, głębokości, jakości i znaczenia wpływu programu na beneficjentów	Użyj ewaluacji skuteczności, aby odpowiedzieć na pytanie, czy działania i plany programu wymagają zmiany
Zbierz i oceń informacje o wpływie programu na społeczność	Użyj ewaluacji skuteczności, aby przygotować i opublikować raporty rozliczające program z wyników
Zaangażuj ewaluatora od ewaluacji wolnej od celów (ma on nie wiedzieć jakie są cele programu) do określenia tego co program rzeczywiście daje i pełnego zakresu jego efektów pozytywnych i negatywnych, zamierzonych i niezamierzonych	Użyj ewaluacji skuteczności, aby ocenić czy program odniósł sukces

Zdobądź informacje o istocie, kosztach i sukcesach podobnych programów przeprowadzonych gdzie indziej i oceń skuteczność danego programu w świetle osiągnięć jego „kluczowych konkurentów”	Na podstawie danych z analizy potrzeb (wyniki ewaluacji kontekstu) i ewaluacji skuteczności określ znaczenie programu w porównaniu z podobnymi programami
Przygotuj wstępne wersje raportu z ewaluacji skuteczności i przedstawiaj je klientowi i uzgodnionym adresatom	
Przedstaw i przedyskutuj wyniki ewaluacji skuteczności na specjalnej sesji	
Zakończ raport z ewaluacji skuteczności i przedstaw go klientowi oraz uzgodnionym adresatom	
Włącz wyniki z raportu o ewaluacji skuteczności do profilu programu i do ostatecznego raportu	

Źródło: D. Stufflebeam, op. cit., s. 8.

Skutki podobne do zamierzonych mogą wywoływać inne czynniki, np. mimo uzyskanych w programie umiejętności, bezrobotny znalazł pracę w zawodzie, który ich nie wymagał, a więc mimo uczestnictwa w szkoleniu nie przyczyniło się ono do uzyskania zamierzonego wpływu. Z tego względu zagadnienia związane z oceną skuteczności muszą opierać się najpierw na badaniach, których celem jest odróżnienie skutków programu od podobnych do nich konsekwencji innych czynników. Na tym polega zasadnicza trudność w przeprowadzaniu ewaluacji skuteczności i to jest również przedmiotem wielu kwestii o charakterze metodologicznym, np. eksperymenty z grupą kontrolną i pomiarem wyników przed i po interwencji.

Istotne dla oceny skuteczności są również kwestie dotyczące skutków ubocznych. Powiedzenie: „operacja się udała, ale pacjent zmarł” dobrze ilustruje problem związany z pomijaniem skutków ubocznych interwencji społecznych (w szczególności tych negatywnych dla klientów). Warto, aby ewaluator był kimś w rodzaju adwokata diabła i w tej roli może krytykować program poszukując jego szkodliwych konsekwencji.

Wąsko rozumiana skuteczność nie wystarcza do oceny programu, gdyż może być ona osiągnięta różnym kosztem, stąd w określaniu jakości i wartości programów społecznych istotne jest porównanie ich kosztów z osiąganą skutecznością (efektywność lub ekonomiczność w osiąganiu skuteczności). Wymaga to co najmniej tego, aby poziom realizacji celów miał charakter mierzalny i abyśmy mieli dane o nim, a także o całkowitych kosztach programu. Przykładowo, 80 klientów osiągnęło sukces, a koszty całkowite programu wyniosły 80 tys. zł, co oznacza, że za każdy sukces zapłaciliśmy 1000 zł. Może być to akceptowalna suma, ale niekoniecznie musi tak być, gdy mamy ograniczone zasoby finansowe. W innym programie, który z sukcesem zakończyło 50 klientów, ale kosztował 40 tys. zł, jednostkowy sukces kosztował 800 zł, czyli o 200 zł mniej. Wyraźnie widać, że stosunek skuteczności do kosztów był lepszy w drugim programie.

Ewaluacja trwałości

Jeżeli program został oceniony jako w wysokim stopniu skuteczny i efektywny, to istotne jest, aby był kontynuowany. Poza tymi, którzy na nim najpierw skorzystali może być wiele innych osób, którym mógłby pomóc. Stąd też w pełnej jego ewaluacji istotne są również pytania o jego trwałość, albo inaczej możliwość i potrzebę jego kontynuacji.

Część programów może być z zasady krótkoterminowa, np. w związku z cyklem konkursów grantowych i rocznego finansowania. W takich sytuacjach ewaluacja trwałości może dostarczać argumentów za tym, aby program był finansowany w kolejnym cyklu grantowym, o ile taki w ogóle

przewidziano, lub starań o nowe źródła finansowania. W przypadku stałych działań służb społecznych również może być potrzebne wykazywanie, że powinny być one kontynuowane, szczególnie w czasach oszczędności budżetowych. Jeżeli już trzeba ciąć wydatki, cięcia powinny dotyczyć tych działań służb społecznych, które mają niskie oceny trwałości, np. ani personel, ani też klienci nie bardzo chcą jego kontynuacji.

Tabela 10. Ewaluacja trwałości

7. EWALUACJA TRWAŁOŚCI	
Działania ewaluatora	Działania klienta/interesariusza
Przeprowadź wywiady z kierownictwem i personelem programu, aby zidentyfikować ich opinie o tym, które z sukcesów programu mogą i powinny być podtrzymane	Wykorzystaj wyniki ewaluacji trwałości, aby określić czy personel programu i jego beneficjenci chcą jego kontynuacji
Przeprowadź wywiady z beneficjentami, aby zidentyfikować ich opinie o tym, które z sukcesów programu mogą i powinny być podtrzymane	Wykorzystaj wyniki ewaluacji trwałości, aby określić czy istnieje potrzeba i przekonujące argumenty za kontynuowaniem programu
Dokonaj przeglądu danych ewaluacyjnych na temat skuteczności programu, jego kosztów i potrzeb beneficjentów, aby ocenić które z sukcesów programu mogą i powinny być utrzymane	Zastosuj wyniki ewaluacji trwałości, aby opracować cele i plany kontynuacji programu
	Uwzględnij wyniki ewaluacji trwałości, aby określić, jak najlepiej podzielić odpowiedzialność za kontynuację programu
Przeprowadź wywiady z beneficjentami, aby określić ich opinie na temat kontynuacji usług programu	
Uzyskaj i zbadaj plany, budżety, zadania personelu, i inne istotne informacje, aby zmierzyć prawdopodobieństwo przetrwania programu	
Okresowo przeglądaj program, aby ocenić w jakim stopniu jego sukcesy są trwałe	
Działania związane z przygotowaniem raportu (jw.)	

Źródło: D. Stufflebeam, op. cit., s. 9.

Część zadań ewaluatora w tej fazie polega na zebraniu danych dotyczących opinii kluczowych interesariuszy na temat tego, co z programu zasługuje na kontynuację. Może być tak, że kierownictwo i personel wyrażają opinie bardzo pozytywne, jako że włożyli w działanie programu wysiłek, w związku z tym mogą się z nim utożsamiać, bagatelizując przy tym problemy i pomijając inne punkty widzenia. Przewidywanie tego, czy i jak stronnicze mogą być opinie klientów o programie jest już trudniejsze. Jeżeli usługa oferowana jest bezpłatnie dla klienta, to może być on wobec niej mniej wymagający w porównaniu z sytuacją, gdyby musiał za nią płacić.

Niezależnie od opinii interesariuszy o sensowności kontynuowania programu działań służb społecznych, ewaluator dokonuje niezależnej oceny tego zagadnienia. Informacje zebrane podczas całego procesu ewaluacji mogą mu w tym wydatnie pomagać. Trzeba jednak wziąć pod uwagę fakt, że określenie długoterminowego wpływu programu na dobro klientów wymaga czasu, a decyzje o kontynuacji lub zaniechaniu programu mogą być podejmowane w pośpiechu.

W przypadku, gdy z ewaluacji wynikają wnioski o konieczności zakończenia programu, ewaluator może spotkać się z negatywną reakcją wszystkich tych, którzy czerpią z działania programu korzyści, a więc kierownictwa, personelu oraz klientów i ich bezpośredniego otoczenia. Inni interesariusze, np. lokalni politycy, też mogą negatywnie reagować na sugestie, iż ich ulubiony program może nie być wart kontynuowania.

Ewaluacja przenośności (opcjonalna)

Od dłuższego już czasu w dyskusjach o programach społecznych wspomina się o dobrych praktykach, dobrych przykładach itp. Tworzone są bazy z danymi dotyczącymi zrealizowanych programów i projektów. Naśladownictwo tego, co uznane zostało zasadnie za dobre jest jednym ze sposobów uczenia się i nie wyklucza innowacji oraz wprowadzania ulepszeń do tego, co stanowiło pierwowzór.

W ramach modelu ewaluacji CIPP uznano tę grupę zadań za niekonieczną, ale może być ona wymagana, tak jak to było w projektach finansowanych w ramach Inicjatywy Wspólnotowej EQUAL. Na wynikach ewaluacji przenośności powinno zależeć m.in. władzom publicznym, które mogą w ten sposób przyczynić się do upowszechniania dobrych praktyk i eliminowania złych.

Tabela 11. Ewaluacja przenośności

8. EWALUACJA PRZENOŚNOŚCI	
Działania ewaluatora	Działania klienta/interesariusza
Spraw, aby personel programu dokumentował przypadki pytań o program z zewnątrz, wizyt w celu zapoznania się z nim oraz adaptacji programu gdzie indziej w celu zidentyfikowania potencjalnych i rzeczywistych naśladowców	Wykorzystaj wyniki ewaluacji przenośności w celu określenia potrzeby rozpowszechniania informacji o programie
Jeżeli będzie to uzasadnione zbadaj opinie potencjalnych naśladowców. Poproś ich o: 1) przegląd opisu programu i streszczenia wyników ewaluacji; 2) ocenę odpowiedniości programu do ich sytuacji; 3) ocenę jakości, znaczenia i możliwości powtórzenia programu; 4) informację, czy zastosowali cały lub niektóre tylko elementy programu lub planują to zrobić	Wykorzystaj wyniki ewaluacji przenośności w celu określenia do kogo przede wszystkim powinny być kierowane informacje o programie
	Wykorzystaj wyniki ewaluacji przenośności w celu określenia jakie informacje o programie powinny być rozpowszechniane
	Wykorzystaj wyniki ewaluacji przenośności w celu oceny jak program sprawuje się gdzie indziej
Dokonaj wizytacji i oceny adaptacji programu w innych miejscach	
Działania związane z przygotowaniem raportu (jw.)	

Źródło: D. Stufflebeam, op. cit., s. 10.

Zwiększanie prawdopodobieństwa, że program zostanie zaadaptowany gdzie indziej polega też na tym, że informacje o nim umożliwiają łatwe naśladownictwo. Plany, harmonogramy i budżety powinny być załączone w wersjach elektronicznych do skrótu informacji o programie i o wynikach jego ewaluacji, o ile pozwala na to prawo. Decydujący się na replikację programu mają wówczas wszystkie dane i szablony potrzebne do adaptacji programu w innym czasie i miejscu. Nie ma jednak gwarancji, że program, który w jednym miejscu uznany został za sukces sprawdzi się gdzie indziej, kierownictwo, personel, klienci i inni interesariusze, a także kontekst mogą być mniej lub bardziej odmienne.

Dodatkowe działania związane z ewaluacją przenośności mogą być kosztowne. Upowszechnianie informacji o programie, wizytacja i ocena jego adaptacji w innych miejscach oraz wykonanie innych zadań w ramach ewaluacji przenośności wymaga czasu i pieniędzy. Z tego też względu konieczność przeprowadzenia ewaluacji tego rodzaju powinna być omawiana w fazie uzgodnień kontraktowych.

Metaewaluacja

Ze względu na to, że ewaluacja ma istotne znaczenie dla powodzenia i losów programów społecznych i działań służb społecznych ważne jest zapewnienie, aby miała ona wysoką jakość. Wydawanie wyroków na konkretne programy lub uznawanie innych za dobre praktyki na podstawie wyrwykowych i nierzetelnych danych oraz wątpliwej ich analizy jest szkodliwe i powinno być zmniejszane.

Ewaluacja ma również charakter programu i może być ewaluowana tak jak inne działania. Publicznym i prywatnym sponsorom powinno zależeć na tym, żeby mieć programy ewaluacyjne, których jakość i wartość jest na odpowiednim poziomie.

Źródłem kryteriów dla oceny ewaluacyjnej działań ewaluacyjnych są zwykle standardy ewaluacji opracowywane przez środowisko profesjonalnych ewaluatorów.²¹ W uzgodnieniach kontraktowych może się znaleźć zapis o tym, według jakich zasad i standardów będzie postępował ewaluator. Poza tym stosuje on wybrany model czy podejście ewaluacyjne, które może mieć też własne procedury, czego przykładem są listy kontrolne, jak ta omawiana w tym rozdziale.

Tabela 12. Metaewaluacja

9. METAEWALUACJA	
Działania ewaluatora	Działania klienta/interesariusza
Uzgodnij z klientem standardy prowadzenia i oceny ewaluacji (mogą to być standardy i zasady zaproponowane przez krajową organizację ewaluacyjną, np. Polskie Towarzystwo Ewaluacyjne)	Przejrzyj standardy i zasady dobrej ewaluacji wypracowane przez krajową organizację ewaluacyjną i uzgodnij z ewaluatorem standardy i zasady jakie powinna spełniać jego ewaluacja Rozważ zatrudnienie niezależnego ewaluatora do metaewaluacji
Zachęcaj i wspieraj klienta w wysiłkach by plan, proces i/lub raporty ewaluacyjne zostały poddane niezależnej ocenie	Zachowuj informacje, które będą istotne dla oceny ewaluacji wg uzgodnionych standardów i/lub zasad
Dokumentuj proces i wyniki ewaluacji, aby mogła być ona rzetelnie zbadana i ewaluowana	Dostarczaj informacji i pomagaj we właściwym przeprowadzeniu metaewaluacji
Konsekwentnie stosuj uzgodnione z klientem standardy i/lub zasady, aby zapewnić, że ewaluacja będzie solidna i odpowiedzialna	Zadawaj pytania i podejmuj odpowiednie kroki, aby zapewnić, że ewaluacja będzie prowadzona według ustalonych zasad i standardów
Okresowo stosuj wyniki metaewaluacji, aby poprawić jakość ewaluacji	Weź pod uwagę wyniki metaewaluacji przy decydowaniu o dalszym wykorzystaniu wyników ewaluacji
Sporządzaj notatki o tym, na ile ewaluacja spełniała poszczególne standardy i lub zasady i zamieść je w aneksie do raportu końcowego	

Źródło: D. Stufflebeam, op. cit., s. 11.

Jeżeli któryś z interesariuszy ma wątpliwości, co do wniosków, które płyną z ewaluacji, może zamówić niezależną metaewaluację, która może potwierdzić lub podważyć jego opinię. Jak jednak wynika z zamieszczonych w tabeli zadań dla ewaluatora on sam powinien wspierać metaewaluację, która, tak jak w przypadku innych programów, może być prowadzona na bieżąco i służyć do poprawy i usprawniania procesu prowadzenia ewaluacji.

²¹ Polskie Towarzystwo Ewaluacyjne zaproponowało Standardy Ewaluacji w 2008 r.: <http://www.pte.org.pl/repository/files/PTE/Standardy_ewaluacji_PTE.pdf>.

Raport końcowy

Wyniki całego procesu ewaluacji prezentowane są w raporcie końcowym, który jest też najważniejszym produktem tego procesu. Ewaluator ma już mniejszy wpływ na to, co z treścią tego raportu zrobią interesariusze. O raporcie końcowym można myśleć w ten sam sposób, w jaki traktuje się produkty programów społecznych w odróżnieniu do ich wpływu na klienta (zamawiającego ewaluację) oraz jego zamierzonych adresatów. Proces przeprowadzania całej ewaluacji, zgodnie z listą kontrolną i standardami ewaluacji zwiększa prawdopodobieństwo, że jej wyniki zostaną uznane za wiarygodne i najważniejsi interesariusze zechcą je wykorzystać w swoich postawach i działaniach wobec ewaluowanego programu i jemu podobnych.

Tabela 13. Raport końcowy.

10. RAPORT KOŃCOWY	
Działania ewalatora	Działania klienta/interesariusza
Przygotuj raport tak, aby zaspokoić oczekiwania różnych publiczności, np. opracuj trzy raporty w jednym, czyli raport dotyczący faktów poprzedzających program, implementacji programu i jego wyników	Zapewnij, aby planowany raport był zrozumiały i użyteczny dla różnych rodzajów publiczności
W raporcie dotyczącym faktów poprzedzających program umieść rozdziały na temat organizacji sponsorującej program, pochodzenia programu i jego otoczenia	Zapewnij, żeby raport o faktach poprzedzających realizację programu był dokładny, zwięzły, zrozumiały i użyteczny dla różnych odbiorców
W raporcie o implementacji programu zamieść rozdziały szczegółowo opisujące, jak główne części składowe programu były planowane, finansowane i wyposażane w personel i realizowane, zrób to w taki sposób, by zainteresowani zastosowaniem programu gdzie indziej wiedzieli, jak to zrobić. Ta część ma mieć głównie charakter opisowy, a ewaluacja sprowadza się do zamieszczenia stosownych ostrzeżeń (co może pójść źle)	Zapewnij, żeby raport z implementacji programu był wystarczająco szczegółowy, aby był zrozumiały dla innych i pomógł im w ewentualnym zastosowaniu procedur programu w innym miejscu
W raporcie o wynikach programu zamieść szczegółowe informacje o formie ewaluacji, o jej wynikach z podziałem na wyżej określone punkty (kontekst, wkład, proces, wpływ, skuteczność itd.) i konkluzje ewaluacji (z podziałem na silne i słabe strony programu, wnioski z doświadczeń, ocenę jakości, wartości i ważności programu). Zestaw wyniki programu z jego zamierzeniami, z tym czego potrzebowali beneficjenci, z tym ile program kosztował i porównaj to z informacjami na temat podobnych programów gdzie indziej	Zastosuj raport z wyników programu do określenia tego, co zostało osiągnięte; do zidentyfikowania porażek i ograniczeń; do oceny na ile program był etyczny; do porównania z wynikami podobnych programów i lekcji, jakie należałoby uwzględnić w przyszłych programach tego typu
Pod koniec każdego z trzech raportów zamieszczaj zdjęcia i graficzne elementy powtarzające i podkreślające główne wątki	Zachowaj raport ewaluacyjny jako instytucjonalną pamięć programu i informację dla zainteresowanych nim
Uzupełnij treść raportu odpowiednimi cytatami; we wstępie napisz, jak ewaluacja została zainicjowana; w zakończeniu opisz potrzeby dotyczące dalszej ewaluacji; opracuj streszczenie; zamieść informacje o ewaluatorach; w aneksach technicznych zamieść kwestionariusze, programy	

sesji dyskusyjnych, tabele z danymi itp.	
--	--

Źródło: D. Stufflebeam, op. cit., s. 12.

Zadania związane ze sporządzeniem raportu końcowego nie ograniczają się jedynie do tego, jak ma go sporządzić ewaluator, jaką ma mieć strukturę i co ma on zawierać. Istotne zadania ma też sam zamawiający ewaluację, gdyż w tej postaci otrzymuje właśnie to, co zamówił i za co zapłacił. Ze względu na koszty ewaluacji niekorzystanie z jej wyników jest marnotrawstwem. W badaniach nad ewaluacją programów społecznych zauważono jednak, że jest to istotny problem. Należy unikać tego, co może sprawić, że raport zostanie przyjęty przez decydujących interesariuszy z niedowierzeniem i nieufnością, co zasadniczo zmniejszy prawdopodobieństwo wpływu tego, co zostało w nim zawarte na program i praktykę.

W szczególności dotyczy to przekazywania w raporcie informacji, które mogą być przykre dla sponsorów, kierownictwa czy personelu programu. Jednym z użytecznych zaleceń w tym względzie jest przedstawianie silnych stron ewaluowanego programu najpierw. Zakładamy, że każdy program ma jakieś silne strony, nawet te ocenione najgorzej, chociażby zaangażowanie, poświęcenie i ciężka praca personelu. W przypadku ustnego przekazywania negatywnych informacji należy je umieszczać w jak najbardziej pozytywnym kontekście, np. zapewnić aby spotkanie odbywało się w przyjacielskiej atmosferze. W pierwszych wersjach raportu powinno się przedstawiać negatywne informacje w jak najbardziej rzeczowy sposób, ale z pozytywnej perspektywy. Potrzebna jest bezpośrednia dyskusja z klientem nad wersją lub kilkoma wersjami wstępnymi raportu, w której może przedstawić on fakty nieznanne ewaluatorowi, a mające wpływ na sposób interpretowania informacji negatywnej. Należy je umieścić w raporcie końcowym w zestawieniu z rzeczową informacją o niezadawalających wynikach. Takie działania ułatwiają klientowi zaakceptowanie treści końcowego raportu ewaluacyjnego.²²

W ocenie jakości końcowego raportu ewaluacyjnego może być przydatna lista jego pożądanych cech, która może być przekształcona w kolejną listę kontrolną służącą do oceny jakości raportu (patrz ramka).

²² Van Mondfrans, za: Jody L. Fitzpatrick, James R. Sanders, Blaine R. Worthen *Program Evaluation: Alternative Approaches and Practical Guidelines*, trzecie wydanie, Pearson, Allen & Bacon, Boston 2004, s. 396.

Ramka 4. Pożądane cechy raportów końcowych

- Wersje wstępne i końcowa przygotowane zostały na czas.
- Treść dostosowana do rodzaju adresata lub adresatów.
- Forma i styl dostosowane do rodzaju adresata.
- Włączenie adresatów do wyboru formy i stylu .
- Zawiera streszczenie.
- Ma odpowiednie wprowadzenie.
- Informuje o ograniczeniach przeprowadzonej ewaluacji.
- Odpowiednia prezentacja planu i procedury ewaluacji.
- Odpowiednio zorganizowana prezentacja wyników.
- Wszystkie istotne informacje techniczne są obecne (najlepiej w załącznikach).
- Zawiera specyfikację standardów i kryteriów ewaluacji.
- Zawiera oceny ewaluacyjne.
- Wylicza zidentyfikowane zarówno mocne, jak i słabe strony.
- Zawiera rekomendacje działań.
- Chroni interesy klienta i interesariuszy.
- Uwzględnia to, że wyniki ewaluacyjne mogą bezpośrednio dotyczyć ludzi (wrażliwość pod tym względem).
- Uwzględnia zdanie mniejszości i zdania odrębne.
- Prezentacja jest dokładna i niestronnicza.
- Skuteczna komunikacja poprzez „opowiadanie historii”.
- Odpowiedni poziom szczegółów.
- Unikanie żargonu technicznego.
- Stosowanie poprawnego, nieskomplikowanego i interesującego języka.
- Stosowanie przykładów i ilustracji.
- Atrakcyjna forma wizualna.

J. L. Fitzpatrick i in. op, cit., s. 400.

Z listy przedstawionej w ramce wynika, że nie tylko samo zbieranie danych i ich analiza są istotne, ale również przyjazna i dostępna forma prezentacji wniosków ewaluacyjnych ma duże znaczenie dla powodzenia całego przedsięwzięcia ewaluacyjnego.

Zakończenie

Przedstawiony powyżej proces ewaluacji ma w dużej mierze idealny charakter. Nie bez racji poszczególne podejścia do ewaluacji nazywa się modelami, a więc upraszczającymi rzeczywistość i traktowanymi jako wzory czy wskazówki, a nie ścisłe reguły, do których należy się ślepo dostosować. Pamiętając o tym, lista kontrolna przedstawiona powyżej może nam posłużyć zarówno jako materiał ułatwiający zrozumienie procesu ewaluacji, jak i praktyczny przewodnik po czynnościach ewaluacyjnych.

Model CIPP ma charakter całościowy, a niektóre z wyżej przedstawionych modułów są opcjonalne (ewaluacja przenośności). W praktyce zlecane zadania ewaluacyjne mogą mieć charakter częściowy, np. zamówienie na metaewaluację, czy ewaluację skuteczności. Modułowy charakter listy kontrolnej sprawia, że można ją stosować również w takich przypadkach.

Należy pamiętać, że poza modelem CIPP są też inne podejścia do ewaluacji całościowej, są to m.in. ewaluacja zorientowana na zastosowanie wyników Michaela Q. Pattona, czy ewaluacja zorientowana na potrzeby klienta Michaela Scrivena. Każda z nich ma własną listę kontrolną. W

każdym przypadku jednak wiele wątków powtarza się, gdyż ewaluacja ma cechy uniwersalne, które nie zależą od zastosowanego podejścia.

Lista haseł do skorowidza przedmiotowego

Lista kontrolna ewaluacji
Ewaluacja kontekstu
Ewaluacja procesu
Ewaluacja wkładu
Typy celów
Ewaluacja wpływu
Ewaluacja skuteczności
Ewaluacja trwałości
Ewaluacja przenośności
Uzgodnienia kontraktowe
Raport końcowy

Przykłady ewaluacji projektów

Wstęp

Każda ewaluacja ma swoją specyfikę, która zależy od ocenianego projektu czy programu lub projektów i programów działań, jak i od kontekstu, w którym go lub je realizowano. Poniżej zostanie opisanych kilka ewaluacji na podstawie podsumowujących je raportów końcowych. Dwie z nich dotyczyły pojedynczych projektów z zakresu reintegracji społecznej z wykorzystaniem pomocy w zdobyciu własnego mieszkania lub domu, a także budowy społeczności wolnej od problemów społecznych (przykład polski i amerykański).

Dwie kolejne ewaluacje odnosiły się do całych zbiorowości projektów, które miały służyć podobnym celom. W pierwszym przypadku były to projekty szkoleniowe z zakresu technologii informatyczno-komunikacyjnych (ICT) dla osób bezrobotnych. Drugi przykład ewaluacji wieloprojektowej dotyczył wsparcia dla instytucji rynku pracy i pomocy społecznej. Obie grupy projektów były realizowane w ramach Sektorowego Programu Operacyjnego Rozwój Zasobów Ludzkich (SPO RZL).

Każdy przykład zostanie przedstawiony według następującego schematu. Najpierw zaprezentowane zostaną informacje na temat projektu lub projektów, a następnie wykorzystane metody i techniki ewaluacji oraz wnioski, jakie przedstawiono w raportach końcowych. Każdy z przykładów zostanie też opatrzony wnioskami ogólniejszymi, w których poruszone zostaną zagadnienia dotyczące praktyki ewaluacji.

Ewaluacja Wolskiego Programu Reintegracji Społecznej „Druga Szansa”

Informacje o projekcie

Program polegał na tym, że wytypowanym samotnym osobom bezdomnym przydzielono zdewastowane mieszkania socjalne (pustostany), które zostały przy ich udziale wyremontowane. Następnie osoby te były monitorowane i wspierane w razie potrzeby w ciągu dwóch lat. Realizację i finansowanie projektu powierzono organizacji pozarządowej.²³

We wrześniu 2004 r. Rada Dzielnicy Wola w Warszawie przyjęła do realizacji projekt „Druga szansa” - Wolski Program Reintegracji Społecznej. Beneficjentami zamierzonymi programu było 10 osób bezdomnych objętych indywidualnymi planami wychodzenia z bezdomności. Cele programu określono następująco.

1. Wyremontowanie 10 mieszkań z zasobów komunalnych Dzielnicy Wola (posiadających status lokali socjalnych, wymagających remontu i adaptacji).
2. Nabycie przez uczestników projektu:
 - a. kwalifikacji i umiejętności w wykonywaniu zawodów: malarza, glazurnika, sprzątacza.
 - b. umiejętności związanych z pracą w zespole oraz samodzielnym życiem i prowadzeniem gospodarstwa domowego.
3. Realizacja zadań wynikających z ustawy o pomocy społecznej.

W założeniach programu było zakwalifikowanie najpierw 20 osób, które miały zgłaszać organizacje społeczne zajmujące się pomocą dla bezdomnych (jednym z warunków kwalifikacji było życie w trzeźwości od co najmniej 6 miesięcy), a następnie wybór spośród nich ostatecznych adresatów. Ci mieli otrzymać umowy najmu wyremontowanych mieszkań na 2 lata z możliwością

²³ Szczegóły dotyczące projektu znajdują się w załączniku do Uchwały Nr X/33/2004 Rady Dzielnicy Wola: <<http://bip.warszawa.pl/NR/rdonlyres/39ED6177-68AE-4603-AF33-72E9459C227B/577900/X33210904.doc>>.

przedłużenia po spełnieniu kilku warunków, np. regulowania na czas wszystkich opłat związanych z użytkowaniem mieszkania. W tym okresie adresaci projektu mieli otrzymywać kompleksowe wsparcie. Kluczowe decyzje w sprawie rekrutacji do programu oraz jego oceny podejmował Wolski Zespół Reintegracji Społecznej (WZRS, przedstawiciele władzy i administracji lokalnej, instytucji pomocy społecznej oraz Stowarzyszenia „Otwarte Drzwi”). Autorem i liderem projektu był Naczelnik Wydziału Spraw Społecznych i Zdrowia Dzielnicy Wola.

Program miał trwać od 1 września do 31 grudnia 2004 r., ale, po pierwsze, uchwała o jego realizacji miała datę 24 września, a po drugie, osoby którym przyznano mieszkania jeszcze dwa lata po tym fakcie były monitorowane i wspierane.

Wśród rezultatów przewidywano, poza wyremontowaniem dziesięciu zdewastowanych lokali będących własnością dzielnicy, m.in. wyjście z bezdomności i usamodzielnienie osób objętych programem oraz stworzenie podstaw do systemowego rozwiązywania problemu wykluczenia społecznego.

Całość środków (np. materiały budowlane) miało pozyskać z różnych źródeł Stowarzyszenie „Otwarte Drzwi”, które było realizatorem projektu.

Metody i techniki monitoringu i ewaluacji

W dokumencie, który był oficjalną podstawą działania programu była tylko jedna wzmianka o ewaluacji: „Prowadzenie monitoringu realizacji programu (ewaluacja)”. Z kolei sam monitoring był ostatnim z zaplanowanych działań, które miało nastąpić po podpisaniu umów najmu i wręczeniu kluczy uczestnikom.

Monitoring sytuacji beneficjentów programu polegał głównie na przeprowadzanych co pół roku wywiadach środowiskowych, comiesięcznych spotkaniach uczestników (za każdym razem w innym z mieszkań) oraz niezapowiedzianych wizytacjach o charakterze kontrolnym. W wyniku tych ostatnich dostrzegano problemy, na które starano się na bieżąco odpowiadać, np. za pomocą dodatkowego wsparcia.

Informacje o metodach ewaluacji opiszę na podstawie tekstów Julii Wygnańskiej *Ewaluacja „Drugiej Szansy”* oraz *Członkowie WZRS o „Drugiej Szansie”*²⁴. Ewaluatorzy byli niezależni, co według deklaracji, umożliwiło zdobycie im dodatkowych informacji, które nie byłyby ujawnione w innym przypadku. Zbieranie danych do analizy polegało na przeprowadzeniu z osobami zakwalifikowanymi do programu trzech wywiadów w 2005, 2006 i 2007 r. (pierwotnie planowano cztery wywiady w odstępach półrocznych). Kwestionariusze do wywiadu różniły się nieco, gdyż z części pytań z I etapu zrezygnowano, a w ostatnim etapie dodano kilka nowych pytań.²⁵

Pytania z kwestionariusza zostały podzielone na kilka części tematycznych i służyły do uzyskania danych na temat stanu zdrowia (np. czy Pani/Pan się leczy: tak, nie bo nie trzeba, nie chociaż należałoby; czy w ciągu ostatnich 6 m-cy stan Pani/Pana zdrowia: polepszył się, pozostał bez zmian, pogorszył się), gospodarstwa domowego (np. ile osób liczy Pani/Pana gospodarstwo domowe (dorośli, dzieci), co Pani/Pan zrobił ostatnio w swoim mieszkaniu (konkrety)?), relacji społecznych (np. jak Pani/Pan opisze stosunki z osobami, z którymi Pani/Pan mieszka? Proszę opisać swoje

²⁴ Oba artykuły ukazały się w publikacji: M. Starzyński, J. Wygnańska (red.) „*Druga szansa*” Wolski Program Reintegracji Społecznej, Wydział Spraw Społecznych i Zdrowia dla Dzielnicy Wola Urzędu m.st. Warszawy, Warszawa 2007: <<http://www.wola.waw.pl/page/index.php?str=451>>.

²⁵ Z informacji autorki wynika, że opublikowano raporty ewaluacyjne z pierwszego i drugiego etapu.

stosunki z sąsiadami), kondycji psychicznej, zatrudnienia (np. jaka była Pani/Pana podstawa zatrudnienia w momencie zgłoszenia się do Programu - rodzaj umowy, na jaki okres, brak zatrudnienia, praca bez umowy, samozatrudnienie itd.; jaka jest obecnie Pani/Pana podstawa zatrudnienia... ?), sytuacji materialnej (Jaka część (procent lub kwoty) w Pani/Pana miesięcznym budżecie stanowią wydatki na: czynsz i opłaty, wyżywienie, lekarstwa/leczenie, odzież, kultura/rekreacja, wydatki na dzieci, alimenty, edukacja, szkolenia, oszczędności, inne?, jak Pani/Pan ocenia swoją sytuację materialną - bardzo dobra, dobra, dostateczna, niedostateczna), pomocy instytucji (Czy korzysta Pani/Pan z oferty Ośrodków Pomocy Społecznej/Rodzinie? Jeżeli tak, to z jakiego rodzaju pomocy Pani/Pan korzysta?), planów i marzeń oraz samego programu „Druga Szansa” (np. Co by Pani/Pan zmienił/a w Programie, w przyszłości dla następnych uczestników? Jakie są Pani/Pana oczekiwania wobec osób prowadzących Program po jego zakończeniu?).

Kwestionariusz wywiadu przygotował wolontariusz, przeprowadzali je również wolontariusze. Treść wypowiedzi respondentów była spisywana.

J. Wygnańska potraktowała odrębnie standaryzowane wywiady przeprowadzone z członkami Wolskiego Zespołu Reintegracji Społecznej, który reprezentował kluczowe instytucje i organizacje zaangażowane w realizację programu. Jednym z głównych tematów wywiadów były jego efekty – sukcesy i porażki. W ten sposób uzyskano dodatkowe dane, które również powinny być podstawą ewaluacji „Drugiej Szansy”.

Na podstawie monitoringu i wywiadów z beneficjentami oraz dodatkowych wywiadów z zarządzającymi i realizatorami uzyskano więc dane do oszacowania wpływu programu na życie jego zamierzonych beneficjentów oraz oceny jego wartości i jakości. Część wywiadów z członkami WZRS dała też informacje o procesie jego realizacji.

Wyniki ewaluacji

Ogólne wnioski ewaluatora zostały sformułowane po skrótowym przedstawieniu wyników wywiadów z kolejnych etapów. Bez żadnych wątpliwości program skutecznie pomógł ośmiu osobom, kolejnym trzem z wątpliwościami (zmarły w trakcie programu w późnej fazie jego realizacji, jedna nie zgodziła się na udział w III etapie, ale były inne przesłanki do wniosku pozytywnego). Dwie osoby nie rozwiązały swoich problemów mimo wzięcia udziału w programie, tzn. otrzymania lokalu i wsparcia. W przypadku pozostałych pięciu osób okres uczestnictwa we wszystkich fazach programu był zbyt krótki, ale część z nich dobrze rokowała.

Członkowie WZRS za sukcesy programu uznali skuteczną pomoc dla jego zamierzonych adresatów, a poza tym wskazywali na jego pośrednie oddziaływanie na pozostałych bezdomnych (wzbudzenie motywacji do wyjścia z bezdomności), pracowników, którzy zajmują się pomocą dla nich (przełamanie dotychczasowych bezskutecznych wysiłków pomocowych), władz lokalnych zajmujących się polityką społeczną (poznanie problemów walki z wykluczeniem w praktyce), ogół społeczeństwa (zaprzeczenie negatywnym stereotypom dotyczącym bezdomnych i możliwości wyjścia z bezdomności), instytucji i organizacji pomocy społecznej (zawiązanie platformy współpracy i koordynacji działań).

Do porażek zaliczono zaś wyraźne niepowodzenie programu w dwóch przypadkach, a ponadto problemy z otrzymaniem lokalu przez jednego z najbardziej aktywnych uczestników (ze względu na zbyt wysokie zarobki), kłopoty z doprowadzeniem mieszkań do odpowiedniego standardu (obawy o to, że będzie zbyt wysoki jak na mieszkania socjalne) oraz niski poziom zainteresowania programem poza dzielnicą, w której był realizowany.

Na podstawie ewaluacji sformułowano rekomendacje dotyczące: 1) sposobu ewaluacji podobnych programów; 2) samego programu jako rozwiązania, które mogłyby być upowszechniane. Do pierwszych należało m.in. zalecenie, aby ewaluację rozpoczynano wraz z przystąpieniem do realizacji programu.²⁶ Pierwsze wywiady ewaluacyjne z uczestnikami przeprowadzono kilka miesięcy po tym, jak otrzymali mieszkania, więc na ich podstawie nie można było porównać ich sytuacji z okresu przed ich wytypowaniem do programu. Podobna uwaga dotyczyła zebrania danych o uczestnikach dopiero po zakończeniu okresu monitoringu, który miał charakter nadzorczo-kontrolny. Wówczas to dopiero można powiedzieć o wpływie programu na uczestników w dłuższym okresie, czyli oceniać trwałość jego wyników.

Z rekomendacji dotyczących programów tego typu zalecano m.in. dodatkowy etap przygotowawczy dla osób bezdomnych, które mają otrzymać lokal socjalny; rezygnację z zasady, że finansowanie remontów w całości spoczywa na organizacji pozarządowej.

Wnioski ogólniejsze

Silną stroną tej ewaluacji jest było to, że jej koszty były minimalne lub żadne dla publicznego budżetu. Można w związku z tym uznać, że sukcesem było, iż w ogóle została przeprowadzona. Część rekomendacji zapewne nie zostałyby sformułowane, gdyby nie podjęto wysiłku ewaluacyjnego.

Opisana ewaluacja prawdopodobnie nie była integralną częścią programu. W związku z tym nie przewidziano na nią żadnych środków finansowych i dlatego była przeprowadzona w ramach wolontariatu. Nie wykonano badań przed rozpoczęciem programu, ani po jego zakończeniu. Błędnie oddzielono różne źródła danych, które łącznie powinny stanowić podstawę do wydania pełnej oceny ewaluacyjnej. Analizie powinny podlegać dane z wywiadów z kluczowymi interesariuszami, członkami WZRS, z wywiadów ewaluacyjnych oraz z monitoringu. Powinien też powstać raport końcowy spełniający podstawowe standardy w tym względzie (patrz rozdział o ewaluacji jako procesie).

Nie przeprowadzono analizy kosztów programu w związku z czym nie ma w nim żadnych ocen dotyczących stosunku kosztów do osiągnięć. Z drugiej strony można utrzymywać, że dla podatników ten koszt był zerowy: środki na remonty zebrała organizacja pozarządowa, a usługi remontowe były udzielane przez samych bezdomnych. Ten aspekt podlegał jednak krytyce w rekomendacjach, stąd informacje o kosztach byłyby bardzo użyteczne dla przyszłych naśladowców.

Ewaluacja programu KE AKA HO'ONA

Informacje o projekcie

Program polegał na tym, że ubogie rodziny z odpowiednią pomocą wybudowały małe osiedle domów jednorodzinnych, w którym następnie zamieszkały tworząc tam nową i zdrową społeczność.

Projekt był opracowany i finansowany przez Consuelo Foundation i miał charakter wieloletni. Rozpoczął się w 1992 r. i zakończył w 2001 r. Omówiony zostanie tutaj na podstawie raportu końcowego z ewaluacji, opublikowanego w 2002 r.²⁷ Fundacja zakupiła niewielki teren w biedniejszej części jednej z wysp Hawajów (Oahu) z zamiarem zrealizowania tam programu samopomocy

²⁶ Nie wszyscy uczestnicy otrzymali lokale w jednym momencie, stąd nie jest też pewne, że powinniśmy uznać za koniec programu upływanie dwóch lat od wydania kluczy pierwszym beneficjentom.

²⁷ A. Gullickson, D. Stufflebeam, L. Wingate *The Spirit of Consuelo: An Evaluation of KE AKA HO'ONA*, Evaluation Center Western Michigan University, kwiecień 2002: <<http://www.wmich.edu/evalctr/pubs/consuelo/SpiritOfConsueloEvalFull.pdf>>.

mieszkańców. Początkowo celem była pomoc dla najbiedniejszych, ale szybko się okazało, że nie mają oni szans na uzyskanie kredytów hipotecznych. W związku z tym ostatecznymi beneficjentami były rodziny w nieco lepszej sytuacji, ale przewidziano też udział kilku rodzin biedniejszych oraz specjalne rozwiązania finansowe dla nich.

Cele programu były następujące: 1) zbudowanie społeczności niezamożnych rodzin wolnej od przemocy i uzależnień, oddanych idei wzajemnej pomocy; 2) powiększenie zasobu finansowo dostępnych domów jednorodzinnych; 3) stworzenie zdrowego i opartego na wartościach podejścia do samopomocy mieszkaniowej i rozwoju społeczności.

Do uczestnictwa w projekcie niekonieczne były umiejętności budowlane, trzeba było jednak spełnić kilka warunków, m.in.: brak własnego domu, co najmniej jedno dziecko poniżej 18 roku życia, dochody na poziomie nie więcej niż 80% mediany dochodów rodzin na wyspie, gotowość do pracy w projekcie co najmniej dwóch osób dorosłych, pozytywny stosunek do rodziny, zdolność do uzyskania kredytu hipotecznego. Procedury upowszechniania informacji o programie były różnorodne, a proces rekrutacji dość złożony, utworzono też listy oczekujących. W założeniach chciano wspomóc 79 ubogich rodzin.

Samopomoc mieszkaniowa przy budowie domów była połączona z fachowym nadzorem i wsparciem budowlanym (np. w zakresie kładzenie podłóg, instalacji kanalizacyjnych i elektrycznych). Zapewniono też okresowe kontrole inspekcji budowlanej. Rodziny (dwoje dorosłych) miały pracować co najmniej dziesięć godzin przy budowie domu w soboty i niedziele przez 9-10 miesięcy, zapewniając w tym czasie opiekę innym dorosłym nad swoimi dziećmi.

Po każdym okresie budowy następował przydział domów na zasadzie loterii, żeby zapewnić równe bodźce do pracy. Oddanie kolejnych domów do użytku było uroczyste celebrowane, a lokatorzy zobowiązywali się przestrzegać formalnie określonych zasad ich użytkowania.²⁸ Fundacja zapewniła m.in. budynek użyteczności publicznej, w którym miało być zlokalizowane centrum społeczności (z salą konferencyjną na spotkania całej społeczności, małym biurem itp.). Dla uczestników nie był to projekt bezpłatny. Po pierwsze, musieli pracować przy budowie, po drugie, musieli spłacać zaciągnięte kredyty hipoteczne oraz płacić Fundacji za dzierżawę działek. Wiele dodatkowych instrumentów miało jednak zmniejszyć obciążenia finansowe dla rodzin, np. za 1 zaoszczędzonego przez rodzinę dolara Fundacja dodawała 3 dolary, po czym cała zebrana w ten sposób kwota (maksymalnie 5 tys.) była przeznaczana na poczet łagodzenia kosztów kredytu hipotecznego.

Mieszkańcom udzielono pomocy w założeniu stowarzyszenia, które miało za zadanie zarządzać całą społecznością. Ponadto w trakcie projektu zapewniono rodzinom doradztwo, nauczanie, wycieczki, imprezy integracyjne, kursy dla rodziców i dzieci, granty edukacyjne, pomoc w uzyskaniu kredytów hipotecznych, szkolenia z organizowania społeczności, pomoc w rozwiązywaniu konfliktów, pomoc w zarządzaniu finansami.

W trakcie ośmiu etapów wybudowano 75 domów, w których zamieszkało około 390 osób, zapewniono też infrastrukturę i obiekty użyteczności publicznej (budynek centrum społeczności i plac zabaw dla dzieci).

²⁸ Przed przystąpieniem do programu każdy uczestnik podpisywał zasady, które miały obowiązywać w całej społeczności zawarte w *Ke Aka Ho'ona Declaration of Covenants, Conditions, and Restrictions*. Treść tej swoistej konstytucji była nowelizowana w wyniku dyskusji między mieszkańcami a zarządem Fundacji. W końcowej fazie projektu ustanowiono też kilkusobową komisję złożoną z mieszkańców, która co kwartał sprawdzała stan przestrzegania zasad.

Metody i techniki ewaluacji

Projektodawcy chcieli swój program doskonalić w oparciu o zamówioną ewaluację, więc zespół ewaluacyjny towarzyszył projektowi od początku, przedstawiając co roku raporty ewaluacyjne. Miały one pomóc w rozwiązywaniu problemów realizacyjnych i usprawnić cały proces. Poza tym, na zakończenie przedstawiono raport podsumowujący.

Ewaluatorzy zastosowali podejście ewaluacyjne CIPP opracowane przez Daniela Stufflebeama, omówione w rozdziale dotyczącym ewaluacji jako procesu. W związku z tym sformułowali pytania dotyczące kontekstu, wkładu, procesu oraz wpływu, skuteczności, trwałości i przenośności tego programu.

Wśród metod zbierania danych o programie znalazły się: analiza środowiska (np. warunki ekonomiczne i demograficzne regionu, potrzeby i problemy grupy docelowej), tworzenie profilu programu (aktualizowana baza danych o programie), stała obserwacja (wynajęci obserwatorzy na miejscu, mający za zadanie m.in. dokumentowanie i przeprowadzanie wywiadów), studia przypadków (okresowo powtarzane wywiady z tymi samymi uczestnikami), wywiady z interesariuszami, ewaluacja zaspokojenia potrzeb niezależnie od celów programu („ewaluacja wolna od celów” prowadzona przez ewaluatora nie znającego programu), raporty z wykonania zadań oraz warsztaty do zbierania opinii na ich temat (dyskusje nad wstępnymi wersjami raportów ewaluacyjnych²⁹), synteza wyników ewaluacji częściowych i raport końcowy. Nie wszystkie metody stosowano w każdym roku, ale zawsze co najmniej trzy z nich były wykorzystane.

Do ograniczeń w przeprowadzeniu ewaluacji autorzy raportu końcowego zaliczyli m.in. dużą odległość między siedzibą ewaluatora a miejscem realizacji programu (stąd potrzeba wynajęcia obserwatorów na miejscu) oraz cięcia finansowe w późniejszej fazie ewaluacji co sprawiło, że zaprzestano dokonywania ewaluacji kontekstu (analiza środowiska) oraz aktualizacji profilu programu.³⁰

Ewaluatorzy zadeklarowali, że ich ewaluacja spełniała prawie wszystkie kryteria dobrej ewaluacji ustalone przez środowisko amerykańskie w standardach ewaluacji (łącznie 30, problem był tylko z obowiązkiem upublicznienia wyników ewaluacji ze względu na umowę z Fundacją). W aneksie do raportu końcowego autorzy zawarli tabelaryczne zestawienie standardów z krótkimi opisami, jak je wypełnili. Była to mini raport z autometaewaluacji.

Wyniki ewaluacji

Oceniono, że projekt dobrze odpowiadał na potrzeby ubogich rodzin, które poza dostępnym finansowo mieszkaniem, dotyczyły również wielu innych wymiarów życia, w tym społecznej integracji w społeczności opartej na wartościach.

Pierwotne plany projektu oparte były na wiedzy, którą nabył zarząd Fundacji wizytując programy samopomocy mieszkaniowej w Kalifornii. Plany te okazały się mieć kilka wad, które zostały jednak usunięte w trakcie realizacji, np. początkowo planowano zbudować 79 domów, a zbudowano 75; początkowo działki przyznane rodzinom różniły się rozmiarem, a opłaty dzierżawy były takie same, więc zmieniono ich pierwotny rozkład; początkowo domy miały być dwóch typów, ale to miało zły wpływ na pracę i atmosferę wśród budowniczych. Oznacza to, że proces był dobrze

²⁹ Lista kontrolna takich sesji A. Gullickson, D. Stufflebeam *Feedback Workshop Checklist*, grudzień 2001: <<http://www.wmich.edu/evalctr/checklists/feedbackworkshop.pdf>>.

³⁰ Pełny wynegocjowany koszt tej ewaluacji wynosił ponad 940 tys. dolarów.

przygotowany i elastyczny na tyle, żeby korygować plany i dostosowywać je do pojawiających się problemów i potrzeb.

Wyniki ewaluacji procesu realizacji projektu były prezentowane za pomocą bezpośrednich cytatów z wypowiedzi zamierzonych beneficjentów. W ich opinii realizacja była dobrze przeprowadzona. Proces rekrutacji do programu trwał długo, ale faza ta była istotna dla dalszego powodzenia projektu. Dobrze też oceniono fazę budowy, gdy wynajęci przez Fundację wykwalifikowani budowlańcy uczyli uczestników fachu i nadzorowali ich pracę zgodnie z wcześniej ustalonymi zasadami bezpieczeństwa. Nie zanotowano żadnych wypadków poza jednym, który wynikał ze naruszenia zasady, że dzieci nie będą się pojawiać na budowie, efektem czego była złamana noga. Ewaluatorzy uchylili się od oceny wysiłków podjętych w ostatniej fazie programu. Miały one na celu upodmiotowienia społeczności poprzez założenie stowarzyszenia, które miałyby nią zarządzać. Z jednej strony wskazywano na niepewność projektodawcy w tym względzie, z drugiej na obawy mieszkańców. Głosy negatywne na temat procesu realizacji projektu również prezentowano, ale zwykle była to mniejszość.

Kolejne oceny dotyczące wpływu programu na zamierzonych beneficjentów oraz jego skuteczność w zaspokojeniu ich potrzeb również były pozytywne. Z osiągnięć programu w tym ostatnim zakresie wyliczono m.in.:

- Poczucie wielkiego osiągnięcia w postaci zbudowania domu.
- Funkcjonalne i dobrze położone domy.
- Dostępne kredyty i opłaty za dzierżawę.
- Duma z posiadanej własności.
- Życie w społeczności kierującej się wartościami i opartymi na nich statutami i przepisami.
- Bezpieczne i wolne od narkotyków osiedle.³¹
- Zwiększona wiedza z zakresu gospodarowania budżetem.
- Nabycie umiejętności związanych z utrzymaniem domu.
- Nowe znajomości i przyjaźnie członków kolejnych etapów budowy.
- Dostęp do szerokiego zakresu usług fundacji.

Na dowód tego, że sytuacja finansowa mieszkańców poprawiła się podano różnice między tym, ile wynosiły ich opłaty za czynsz przed wprowadzeniem się do domów oraz miesięczne raty spłat kredytu hipotecznego (w większości te pierwsze były większe o kilkadziesiąt do kilkuset dolarów).

Ewaluatorzy uchylili się od oceny trwałości wyników projektu, ze względu na to, że zakończyli zbieranie danych kilka miesięcy po wprowadzeniu się do nowych domów budowniczych ostatniego etapu. Wspomniano jednak o propozycjach zrobienia ewaluacji w 2005 r., która miałyby odpowiedzieć na tego rodzaju pytania.³²

W przypadku oceny przenośności znowu była ona dość pozytywna, wskazywano na wiele ważnych doświadczeń, które projekt przyniósł i z których można skorzystać gdzie indziej. Zasugerowano Fundacji, aby upowszechniała wiedzę o projekcie i jego osiągnięciach oraz doświadczeniach z nim związanych.

³¹ Wspomniano jednak o przypadku handlarza narkotyków należącego do społeczności z adnotacją, że liderzy społeczności wraz z policją się tym zajęli.

³² Najprawdopodobniej jednak do niej nie doszło. W późniejszej książce D. Stufflebeama i A. J. Shinkfielda *Evaluation Theory, Models, and Application* z 2007 r. zaproponowano kilku przedstawicielom najważniejszych modeli ewaluacyjnych, aby odpowiedzieli na hipotetyczną ofertę przeprowadzenia ewaluacji tego rodzaju.

W podsumowaniu przytoczono aż czterdzieści silnych stron programu (np. orientacja na pomoc dla samopomocy a nie na zasiłki; nacisk na utrzymanie społeczności bez narkotyków i przemocy; finansowe doradztwo udzielane rodzinom), kilka jego słabości oraz ponad dwadzieścia wniosków, jakie można na jego podstawie wyciągnąć (np. diagnoza potrzeb powinna być pierwszorzędnym zadaniem; oparcie na wartościach wprowadza nową jakość; uważna selekcja rodzin wpływa dodatnio na wyniki). Do słabości zaliczono to, że projekt nie służył jednak najbardziej z biednych, czyli rozmiął się z misją Fundacji; przez pierwsze pięć lat trwania projektu nie było jasności co do tego, jak Fundacja ma zamiar przekazać społeczności zarządzanie osiedlem; dla niektórych intersariuszy osiągnięty poziom spójności społecznej był zbyt niski; brak zaangażowania członków społeczności w organizowanie pomocy dla sąsiadujących z nimi osiedlami i społecznościami; zagrożenie rosnącym zadłużeniem i niemożnością wykupienia działek po spłaceniu kredytów hipotecznych.

Raport końcowy był bogato ilustrowany zdjęciami i zawierał kilka aneksów, w tym podręczniki dla miejscowych obserwatorów, scenariusze wywiadów, listę kontrolną ewaluacji według modelu CIPP.

Wnioski ogólniejsze

Przedstawiona powyżej w skrócie ewaluacja miała charakter jakościowy. W wydawaniu ocen opierano się przede wszystkim na cytatach z wywiadów bez wprowadzania mniej lub bardziej zaawansowanych analiz o charakterze statystycznym, chociaż natura programu i liczba obsłużonych osób pozwalały na to.

W dochodzeniu do syntezy wniosków ewaluacyjnych nie określono ważności, ani też metody agregowania poszczególnych ocen. Tych ostatnich nie próbowano zresztą sprowadzić do jakiejś wspólnej skali. Może to oznaczać, że nie zastosowano metod specyficznych dla syntezy ewaluacyjnej (patrz rozdział o metodach oceniania).

Brak ewaluacji programu w dłuższym okresie po jego zakończeniu sprawia, że trudno poddać ocenie jego długoterminowy wpływ na życie mieszkańców osiedla oraz na spójność stworzonej tam społeczności. Bez odpowiedzi pozostaje też pytanie, co się stało lub by się stało, gdyby zrealizowano radykalny projekt upodmiotowienia tej społeczności.

Ewaluacja projektów szkoleniowych z zakresu ICT dla osób bezrobotnych

*Informacje o projektach*³³

Wiele projektów finansowanych ze środków Europejskiego Funduszu Społecznego polega na bezpłatnym udzieleniu osobom bezrobotnym usług szkoleniowych. Jedną z popularniejszych w dobie idei rozwoju społeczeństwa informacyjnego, czy gospodarki opartej na wiedzy, są kursy z zakresu obsługi komputera i oprogramowania. Używa się w tym kontekście angielskiego skrótu ICT od wyrażenia technologie informacyjno- lub informatyczno-komunikacyjne, czasem po prostu mowa o „nowych technologiach”.

³³ Informacje w tym punkcie opierają się na raporcie J. Drażkiewicza i in. *Ocena stopnia rozwoju społeczeństwa informacyjnego w ramach Sektorowego Programu Operacyjnego Rozwój Zasobów Ludzkich 2004-2006*: <http://www.efs.2004-2006.gov.pl/NR/rdonlyres/44AFE056-C063-4708-A863-E4C9585D1117/48944/MRR_SPORZL_ICT_Raport_koncowy_final.pdf>. □

Ze względu na to, że wiele projektów koncentrowało się na szkoleniach komputerowych, ale również miało taki komponent poza innymi, była to spora zbiorowość statystyczna (ponad 380 projektów). Włączając w to jeszcze kategorię projektów, które powinny zawierać tego rodzaju szkolenia, a miały tylko zadania związane z wykorzystaniem nowych technologii, otrzymujemy kolejne setki projektów (ponad 290).

Przedmiotem ewaluacji nie jest już w tym przypadku jeden projekt, ale cała grupa projektów o podobnym charakterze lub zawierających podobne działania. Na ogół są to projekty, które mają realizować jeden z priorytetów czy celów programu szerzej zakrojonych oddziaływań na większe zbiorowości ludzi, np. zwiększenie szans na zatrudnienie osób bezrobotnych czy rozwój społeczeństwa informacyjnego w Polsce, rozpisane dodatkowo na konkursy wojewódzkie. W przypadku omówionym tutaj sytuacja była bardziej złożona, gdyż analizowano projekty, które w założeniu miały realizować różne cele szczegółowe (priorytety SPO RZL).

Metody i techniki ewaluacji

Jedno z zasadniczych pytań w przypadku omawianej ewaluacji było takie: na ile działania podejmowane w trakcie realizacji programu (SPO RZL) przyczyniły się do rozwoju społeczeństwa informacyjnego. Poprzez działania rozumiano to, co zostało zrobione przez realizatorów projektów, które miały zadania ze wskazanego zakresu, czyli przede wszystkim szkolenia ICT. Zamierzonym celem i przewidywanym wpływem bezpośrednim każdego szkolenia jest podniesienie poziomu wiedzy i umiejętności jego uczestników w zakresie, który był przedmiotem szkolenia. Stąd jedno z pytań ewaluacyjnych dotyczyło tego, jaki w rzeczywistości był ten wpływ. Jeżeli przy tym szkolenia były adresowane do bezrobotnych, to spodziewano się, że pomogą im w znalezieniu pracy. Kolejnym problemem do zbadania było więc, czy i na ile tak się stało. Warunkiem zwiększenia zatrudnienia wśród uczestników tych szkoleń było to, że przekazywane w ich wyniku wiedza i umiejętności będą przydatne pracodawcom. Tego właśnie dotyczyło trzecie pytanie szczegółowe.

Bezpośrednio określono kryteria ewaluacyjne: trafność (to, co oferowano w projektach a potrzeby rozwoju społeczeństwa informacyjnego), skuteczność (osiągnięte rezultaty a cele programu) i użyteczność (udzielone wsparcie a potrzeby beneficjentów).

Ewaluatorzy postanowili zebrać dane, które pozwolą im odpowiedzieć na pytania szczegółowe oraz dokonać ewaluacji poprzez dotarcie do czterech zbiorowości: 1) beneficjentów ostatecznych; 2) projektodawców; 3) pracodawców i 4) społeczeństwa. Do tego celu wykorzystali wywiady telefoniczne ze wspomaganiami komputerowym (skrót ang. CATI), ankiety internetowe (skrót ang. CAWI), indywidualne wywiady pogłębione oraz analizę źródeł wtórnych zastanych, czyli informacji zawartych w raportach, artykułach czy książkach na temat interesujący ewaluatorów.³⁴

W przypadku tego zadania ewaluacyjnego występowały dodatkowe kwestie natury metodologicznej, czyli wybór odpowiednich projektów z większego zbioru (dobór próby), wybór beneficjentów ostatecznych z którymi przeprowadzono wywiady, wybór pracodawców i w końcu też wybór literatury, którą chcieli przeanalizować. W analizach statystycznych istotnym jest, aby próby były reprezentatywne, czyli miały wystarczającą liczebność i podobną strukturę do całej zbiorowości. Jest to warunek konieczny, aby na podstawie analizy próby wyciągać wnioski o całej zbiorowości.

Ewaluatorzy natrafili na problemy właściwego doboru próby projektów oraz beneficjentów, gdyż udostępniona im baza danych o nich nie pozwalała na zidentyfikowanie zakresu wsparcia. W

³⁴ Podejście to zastosowano, żeby odpowiedzieć na pytania dotyczące poziomu całego społeczeństwa.

związku z tym zastosowali tzw. dobór celowy, który jednak nie zapewniał reprezentatywności, co też uniemożliwiło wyprowadzanie wniosków uogólniających. Podobna kwestia pojawiła się w przypadku pracodawców, gdyż wielu beneficjentów nadal było bezrobotnych, a zatrudnieni często nie chcieli podawać danych aktualnego zatrudniającego.

W celu zbadania skuteczności zbudowano bardziej złożone wskaźniki, np. syntetyczną miarę postępów beneficjentów ostatecznych. W analizie danych przy orzekaniu o determinantach skuteczności projektów wykorzystano metody analizy statystycznej (analizę regresji i jednoczynnikową analizą wariancji lub jedną z jej wersji).

Wyniki ewaluacji

Wnioski ewaluacyjne przedstawiano w treści raportu oraz w jego podsumowaniu. Stwierdzono m.in., że dobór odpowiednich beneficjentów jest kluczowym warunkiem osiągnięcia zamierzonych celów, większość badanych zdobywa umiejętności ICT przez samodzielną naukę, według opinii większości beneficjentów ich umiejętności ICT nie zostały nabyte dzięki wsparciu, a nieco dalej mamy wniosek, że jednak przyczyniło się ono do nabycia i rozwoju kompetencji z zakresu ICT (również na podstawie deklaracji samych beneficjentów). Kolejne wnioski dotyczyły tego, że największe postępy odnotowali ci beneficjenci, których początkowe kompetencje były najmniejsze, a większość z nich była zadowolona z otrzymanego wsparcia.

Tabelaryczna prezentacja wyników i rekomendacji zasługuje na większą uwagę. Pierwsza kolumna zawierała informacje dotyczące stwierdzonych problemów i miała tytuł „Problem: obserwacja, wynik”; druga miała w nagłówku dwa pytania: „Jak powinna wyglądać rzeczywistość? Jaka zmiana powinna zostać dokonana?”; trzecia - również dwa pytania: „Co należy zrobić, aby osiągnąć pożądany stan? Co należy zrobić, aby przeprowadzić pożądane zmiany?”; w ostatniej zamieszczano informację o „odbiorcy rekomendacji” ze wskazaniem nazwy instytucji lub grupy, np. „Instytucja zarządzająca”, „projektodawcy EFS”.

Wśród problemów wymieniano zarówno te, które utrudniły ewaluację (wadliwa baza danych o projektach), jak i inne, np. słabości w uzasadnieniach przeanalizowanych projektów; oparcie diagnozy uzasadniającej projekt wyłącznie na danych zastanych; brak zastosowania przez projektodawców jednolitych narzędzi badania umiejętności z zakresu ICT; niski poziom zatrudnienia w postaci telepracy w badanych firmach; rozbieżności w przypisywaniu wagi umiejętnościom ICT oraz wyobrażeniami o miejscu pracy dla takich pracowników między pracodawcami i beneficjentami; stosowanie przez projektodawców przy selekcji klientów wyłącznie kryteriów formalnych; wydawane na zakończenie szkoleń świadectwa i dyplomy potwierdzają jedynie udział, a nie nabycie konkretnych umiejętności; duże różnice umiejętności w grupach szkolonych; brak związku szkoleń z umiejętnościami poszukiwania pracy itd.

Wnioski ogólniejsze

Przedstawiona powyżej ewaluacja w większym stopniu niż poprzednie opiera się na bardziej zaawansowanej analizie statystycznej. Nie ograniczono się wyłącznie do zestawień tabelarycznych lub graficznych zebranych danych, poddano je również agregacji (budowa złożonych wskaźników) oraz analizie w celu wyciągnięcia bardziej zaawansowanych wniosków, np. jakie cechy beneficjentów ostatecznych i jak wpływały na postęp w poprawianiu ich umiejętności z zakresu ICT .

Wadą raportu końcowego z ewaluacji było nieumieszczenie we wnioskach bezpośrednich ocen dotyczących trafności, skuteczności i użyteczności wybranych do analizy projektów. Czytelnik

musi je znaleźć samodzielnie w tych częściach raportu, które znalazły się przed tabelarycznym zestawieniem wniosków i rekomendacji. Znajdują się tam pogrubione zdania w ramkach, które jednak dotyczą różnych aspektów badania, więc dodatkowo i je trzeba przeszukiwać. Nie zadbane też, aby wyjaśniać pojawiające się w wnioskach częściowych sprzeczności.

Na tym przykładzie możemy zobaczyć niektóre z praktycznych trudności, na jakie napotykały zespoły ewaluacyjne (baza danych wyjściowych z poważnymi lukami) i ich konsekwencje (niemożność dokonywania uogólnień). Poza tym, znowu przekonujemy się o dużej różnorodności metod zbierania i analizowania danych w celach ewaluacyjnych. W przypadku analiz o charakterze ilościowym konieczne są wiedza i umiejętności z tego zakresu (metodologia badań naukowych, metody analizy statystycznej).

Ewaluacja projektów wsparcia dla instytucji rynku pracy i pomocy społecznej

*Informacje o projektach*³⁵

Zamierzonym adresatem czy beneficjentem działań projektowych są również instytucje, których celem jest udzielanie pomocy osobom bezrobotnym, ubogim i zagrożonym wykluczeniem ze społeczeństwa. Wspieranie tych, którzy mają za zadanie pomagać innym powinno również przynieść korzyści zamierzonym adresatom działań służb społecznych. Przykładowo, jeżeli kierownicy i pracownicy tych służb otrzymają odpowiadające ich potrzebom szkolenia, to poziom ich wiedzy i umiejętności zawodowych wzrośnie. To zaś powinno zaowocować zwiększeniem jakości usług udzielanych przez nich obywatelom potrzebującym pomocy.

Zakładając ponadto, że instytucje rynku pracy i pomocy społecznej tworzą odrębne systemy współpracujące ze sobą, można częściowo oddzielić od siebie pomoc dla systemu instytucjonalnego i pomoc dla poszczególnych jego jednostek. Projekty mogły mieć różny zasięg od centralnego, przez regionalny, do lokalnego. Do ich klasyfikacji według tego kryterium posłużył ich zasięg terytorialny (np. obejmujące do 5 powiatów zaklasyfikowano jako lokalne, a kilka województw, ale nie wszystkie – jako ponadregionalne).

Ewaluacji poddano ponad 380 projektów, które miały realizować cele określone w działaniach 1.1. i 1.5 w schemacie a) SPO RZL. Już podczas ewaluacji podzielone one zostały na kilka typów: systemowe i badawcze, szkoleniowe, ukierunkowane na rozwój instrumentów i usług, wspierające partnerstwo i współpracę oraz promocyjne.

Główne pytania ewaluacyjne dotyczyły tego, czy i w jakim stopniu działania projektowe przyczyniły się do realizacji wybranych celów SPO RZL. Były nimi m.in. poprawa jakości i produktywności lub efektywności pracy pracowników instytucji pomocy społecznej i instytucji rynku pracy, zarządzania i administracji, działań tych instytucji wzmacniających spójność i integrację społeczną.

Metody i techniki ewaluacji

³⁵ Opis na podstawie: *Ewaluacja działań podejmowanych na rzecz wsparcia systemu instytucjonalnego rynku pracy oraz pomocy społecznej w ramach EFS. Raport końcowy*, Instytut Badań Strukturalnych, czerwiec 2008: <http://www.efs.2004-2006.gov.pl/NR/rdonlyres/44AFE056-C063-4708-A863-E4C9585D1117/47500/raport_koncowy_FINAL.pdf>.

Dane zbierano za pomocą kilku narzędzi. Analizowano dokumenty, czyli wnioski projektowe, aby podzielić je na kilka grup istotnych z perspektywy pytań ewaluacyjnych. Wywiady internetowe (CAWI) zastosowano wobec projektodawców oraz instytucji rynku pracy i pomocy społecznej. Pracownikom służb społecznych, którzy brali udział w działaniach z zakresu doskonalenia zawodowego zadano pytania za pomocą wywiadów telefonicznych wspomaganych komputerowo (CATI). Zestandaryzowane wywiady pogłębione (według tego samego scenariusza) zastosowano do zebrania danych od specjalnie wytypowanych przedstawicieli wspieranych instytucji.

W ramach tej ewaluacji dokonano diagnozy problemów (potrzeb) instytucji rynku pracy i pomocy społecznej na podstawie przeglądu literatury oraz badań własnych. W ten sposób uzyskano niezależny od diagnozy zawartej w SPO RZL obraz potrzeb beneficjentów wsparcia. Z perspektywy metodologii ewaluacyjnej jest to bardzo istotne posunięcie, gdyż pozwala później oceniać zgodność podjętych działań (interwencji) ze zdiagnozowanymi problemami i potrzebami. Główną metodą zbierania danych do diagnozy były wywiady internetowe, najczęściej z kierownikami wybranych instytucji (próba nie była reprezentatywna). Wśród zadanych pytań były: „W jakim stopniu brak wystarczających środków finansowych stanowił istotne ograniczenie w funkcjonowaniu Państwa instytucji?”, „W jakim stopniu poniższe problemy stanowią istotne ograniczenie w prawidłowym funkcjonowaniu Państwa instytucji?”. Wśród zaproponowanych odpowiedzi na to drugie pytanie znalazły się m.in.: niedostateczna liczba pracowników, brak uniwersalnych standardów jakości świadczonych usług, niewystarczający poziom współpracy z pozostałymi instytucjami, brak wiedzy na temat obszaru działania instytucji, niedostateczne wyposażenie w sprzęt, niedostateczne kwalifikacje pracowników, niska jakość organizacji i zarządzania. Badano opinie przedstawicieli publicznych, jak i niepublicznych instytucji rynku pracy i pomocy społecznej.

Ewaluacje prowadzono w kilku modułach, które odpowiadały typom projektów wskazanym wyżej. Każdy moduł poprzedzała prezentacja ocenianych projektów pod względem liczby, struktury (np. według rodzaju beneficjentów, rodzaju podejmowanej działalności, obejmowanego obszaru) i poziomu dofinansowania. Otrzymane w ten sposób zestawienia porównywano następnie z danymi od projektodawców czy instytucji, które otrzymano po zastosowaniu wymienionych wyżej metod.

W przypadku projektów systemowych i badawczych projektodawcom zadawano między innymi pytania: „Czy rezultatem tej analizy [zaproponowanej w projekcie badawczym] będą/są rekomendacje dla polityki społeczno-gospodarczej, zakładające zmianę obowiązujących rozwiązań prawnych?”; „Czy rekomendacje zmian rozwiązań prawnych, wypracowane w ramach projektu, zostały dotychczas w jakimkolwiek stopniu wprowadzone w życie?”. W ten sposób zbierano dane do oceny potencjalnego wpływu przeprowadzonych badań na rzeczywistość instytucjonalną.

Przy ewaluacji projektów o charakterze szkoleniowym ważnym elementem było porównywanie wsparcia ze zdiagnozowanymi potrzebami, zarówno tych instytucji, które je otrzymały, jak i tych, które nie były wspierane. Wśród zadanych pytań na temat potrzeb było m.in. takie: „W jakim stopniu rozwój kompetencji pracowników w poniższych obszarach przyczyniłby się do większej skuteczności realizacji zadań Państwa instytucji?”. Obszary, o które pytano to m.in. zarządzanie, umiejętności miękkie, zagadnienie administracyjno-prawne, praca z klientami, praca z komputerem, znajomość języków obcych. Pytano też o oceny dopasowania otrzymanych usług do potrzeb np.: „Czy wsparcie było dopasowane do Pana/Pani potrzeb?”; „Czy przed rozpoczęciem udziału w szkoleniu/studiach podyplomowych odczuwał Pan potrzebę doskonalenia w problematyce, którą ono obejmowało?”; „Jakie tematy szkoleń byłyby najbardziej przydatne z punktu widzenia poprawy Pana(i) kompetencji zawodowych?”; „jak ocenia Pan/Pani przydatność wiedzy i umiejętności przekazywanych podczas [szkolenia/kursu/studiów podyplomowych] w Pana(i) pracy zawodowej?”.

Projekty ukierunkowane na wsparcie rozwoju instrumentów i usług w wyniku analizy zostały podzielone na siedem typów (w tym mające na celu m.in. wyposażenie instytucji w sprzęt komputerowy i biurowy oraz oprogramowanie, promowanie projektów pilotażowych i innowacyjnych, standaryzację usług rynku pracy). Zbiorowość tych projektów była więc bardzo zróżnicowana. W związku z tym ewaluatorzy zastosowali dodatkową metodę – analizę przypadków, czyli bardziej szczegółowy opis 6 projektów wybranych ze względu na największe koszty. W tym module pytano respondentów (beneficjenci z instytucji) m.in. o ocenę stopnia, w jakim wsparcie w zakresie modyfikacji lub oferowania nowych usług odpowiada potrzebom ich klientów. Inne pytanie dotyczyło trwałości wprowadzonych zmian: „Czy modyfikacje/nowe instrumenty stosowane w Państwa instytucji mają trwały charakter, tj. są nadal stosowane pomimo zakończenia projektu/będą stosowane po zakończeniu projektu?”.

Kolejny typ ocenianych projektów miał na celu wspieranie współpracy i partnerstwa. Poza szczegółową analizą tych projektów, które były podejmowane w partnerstwie lub były ukierunkowane na inicjowanie lub wzmacnianie współdziałania, pytano ich koordynatorów m.in. o to: „Jak ocenia Pan/Pani współpracę w ramach partnerstwa?” oraz „W jakim stopniu realizacja projektów w ramach SPO RZL zwiększyła zaangażowanie Państwa instytucji we współpracę z instytucjami rynku pracy i pomocy społecznej (ogólnie, nie tylko w obszarze wspólnej realizacji projektów finansowanych z EFS)?”.

Projekty należące do ostatniego rodzaju miały upowszechniać informacje o konkretnych usługach, o dobrych praktykach (w większości wypracowanych w danym projekcie), a także o ogólnej działalności instytucji, polegały też na prowadzeniu kampanii wizerunkowych.

Ewaluacji podlegały również projekty łącznie, czyli niezależnie od wyróżnionych siedmiu typów i ewentualnie dodatkowych podtypów. W tym przypadku za najważniejsze uznano pytanie skierowane do beneficjentów instytucjonalnych. Pytano ich o ocenę stopnia poprawy skuteczności realizacji zadań instytucji w wyniku wsparcia otrzymanego dzięki ewaluowanym projektom. Już samo przedstawienie rozkładu odpowiedzi według obszaru przynależności instytucjonalnej (rynek pracy, pomoc społeczna) oraz rodzaju sektora (publiczny, niepubliczny) dostarczyło interesującego materiału do interpretacji. Porównanie rozkładu odpowiedzi udzielonych na to pytanie przez projektodawców oraz beneficjentów uwidoczniło różnice, które ewaluator musiał dodatkowo wyjaśniać. Pytano również o obszar, w którym nastąpiła największa poprawa w wyniku otrzymanego wsparcia.

Dane zbierano także od instytucji, które nie otrzymały wsparcia. Możliwe były porównania rozkładów odpowiedzi między nimi oraz tymi, które je otrzymały. Przypomina to już metodę quasi eksperymentalną (patrz odpowiedni rozdział w tym poradniku), gdy wykryte w ten sposób podobieństwa i różnice stanowią podstawę do orzekania o wpływie wsparcia.

Wnioski ewaluacji

W podsumowaniu wyników ewaluacji sformułowano ogólne oceny wsparcia, które było wynikiem realizacji analizowanych projektów. Stwierdzono, że było ono zgodne z założeniami, przyczyniało się do poprawy działania instytucji, osiągnęło zamierzone rezultaty, a ankietowani projektodawcy i beneficjenci oceniali je co najmniej dobrze.

Ponadto wskazano na nierównomierność wsparcia: dominowały projekty szkoleniowe i skierowane do instytucji rynku pracy, a przy tym ich umiejscowienie geograficzne nie pokrywało się z geografiami bezrobocia. Problem ten miał wynikać z braku spójnego systemu diagnozy potrzeb i centralnego zarządzania SPO RZL.

Większe znaczenie w ewaluowanej zbiorowości projektów miały szkolenia dla pracowników instytucji. Prowadziło to ewaluatorów do wniosku, że była to przyczyna uzyskania silnego wpływu wsparcia na jakość ich pracy. Szkolenia oceniono jako dobrze dopasowane do potrzeb pracowników, co jednak niekoniecznie dokładnie pokrywało się z potrzebami instytucji.

Wpływ dużych projektów systemowych oceniono jako mniejszy niż oczekiwany. W przypadku działań, które miały na celu rozwój instrumentów i usług stwierdzono „niewielki stopień wykorzystywania wypracowanych narzędzi” oraz „niewielki zakres rozpowszechniania uzyskanych wyników”. W odniesieniu do tych i innych problemów sformułowano 8 rekomendacji, w rodzaju: „Konieczne jest zwiększenie intensywności kompleksowych badań systemowych obszaru pomocy społecznej”.

Wiele szczegółowych wniosków zawarto w podsumowaniach ewaluacji projektów zaklasyfikowanych do poszczególnych modułów.

Wnioski ogólniejsze

Tym razem poziom zróżnicowania projektów był większy niż w poprzednim przykładzie. Jest to kolejne wyzwanie dla ewaluatorów. Cel ogólny zawarty w programie typu SPO RZL może być złożony, a jego realizacja powierzona setkom projektodawców, którzy zaproponowali działania różnej skali i typu. Ewaluacja musi się rozpocząć od klasyfikacji projektów i poddania ewaluacji ich bardziej jednorodnych grup. I tu zasadnicze znaczenie ma jakość bazy danych o projektach.

Ponownie wykorzystano wiele metod zbierania danych, chociaż tym razem nie próbowano budować bardziej złożonych wskaźników statystycznych³⁶ i nie wykorzystywano zaawansowanych metod analizy statystycznej.

Wnioski formułowane były na ogół w oparciu o opinie i oceny koordynatorów projektów oraz przedstawicieli wspieranych instytucji. W pierwszym przypadku ze świadomością, iż mogą oni zawyżać oceny wsparcia, za którego jakość sami byli odpowiedzialni.

Nie sformułowano jasnych kryteriów ewaluacyjnych i nie wyszczególniono żadnych metod oceny (patrz odpowiedni rozdział w tym poradniku). Z tego powodu wnioski stricte ewaluacyjne są formułowane w sposób mało precyzyjny i niekiedy trudno je odróżnić od wniosków badawczych dotyczących różnych aspektów analizowanych projektów.

Zakończenie

Zapewne w działaniach służb społecznych w poszczególnych instytucjach bardziej przydatne są nauki, jakie możemy wynieść z przykładów ewaluacji pojedynczych projektów. Jeżeli jednak pracownicy służb społecznych będą mieli zadania związane z ewaluacją na poziomie ponadlokalnym, to bardzo prawdopodobne, że będzie wchodziła w grę ewaluacja wielu projektów.

Wszystkie przedstawione wyżej ewaluacje miały swoje mocne i słabe strony. Część tych drugich nie zależała od samych ewaluatorów, ale była powodowana czynnikami będącymi poza ich kontrolą. W założeniach projektowych nie uwzględniono ewaluacji, więc nie było na nią pieniędzy. W innym przypadku zmniejszono poziom finansowania, bo sponsor ewaluacji stracił w wyniku kryzysu finansowego. W udostępnionych bazach danych o projektach nie było istotnych dla ewaluacji informacji. Pytania zadane przez zamawiającego w małym stopniu miały charakter ewaluacyjny.

³⁶ Zaproponowano je jednak dla następcy SPO RZL, czyli Programu Operacyjnego Kapitał Ludzki.

Części problemów można było jednak uniknąć i miejmy nadzieję, że upowszechnianie w Polsce wiedzy o ewaluacji, m.in. przy pomocy tego poradnika, przyczyni się do poprawy jakości pracy ewaluacyjnej.

Lista haseł do skrowidza przedmiotowego

Analiza środowiska

Wywiady z interesariuszami

Obserwacja na miejscu

Studium przypadku

Profil programu

Ewaluacja wolna od celów

Autometaevaluacja

Ewaluacja wieloprojektowa

CATI

CAWI

Indywidualne wywiady pogłębione

Analiza dokumentów

Dobór próby

Próba reprezentatywna

Analiza regresji

Analiza wariancji